

# Structure / Nomenclature Guide

A GUIDE TO THE GRAPHIC  
REPRESENTATION AND NOMENCLATURE  
OF CHEMICAL FORMULAE IN THE  
EUROPEAN PHARMACOPOEIA

EUROPEAN PHARMACOPOEIA

**European Directorate for the Quality of Medicines**



**November 2006**

**1<sup>st</sup> Edition**

© Council of Europe, 67075 Strasbourg Cedex, France - 2006  
All rights reserved

Making copies of this file for commercial purposes or posting this file on a web site that is open to public consultation is strictly prohibited.

# NOMENCLATURE AND GRAPHIC REPRESENTATION OF CHEMICAL FORMULAE

## CONTENTS

PREAMBLE

SECTION A – General rules for graphic representation

SECTION B – Graphic rules specific to the Ph. Eur.

SECTION C – Main structural classes

SECTION D – Nomenclature and application of IUPAC rules

SECTION E – Frequently asked questions (FAQ)

REFERENCES

## PREAMBLE

The guide on nomenclature and graphic representation of chemical formulae has been prepared to reply to a number of questions from the European Pharmacopoeia Commission and users of the Ph. Eur.

### I. CHEMICAL NAME OR GRAPHIC REPRESENTATION?

In principle, a chemical structure or name alone can be used to define a chemical compound. However, the Ph. Eur. uses both to facilitate checking and to remove ambiguities. Each system has its advantages and disadvantages, which are summarised below.

#### 1. STRUCTURES

Advantages: molecules are immediately recognisable and their structures are easily compared.

Limits: there is a risk of some inaccuracy with any representation of a chemical structure because it involves drawing a molecule with a 3-dimensional structure in 2 dimensions; bond angles and lengths are not necessarily depicted accurately.

#### 2. NAMES

Advantages: stereochemistry is specified directly with no need to interpret the structure.

Limits: nomenclature rules are very complex, numerous and difficult to master; at present, there are no rules that can be used to determine which name is officially preferable.

### II. SCOPE OF THE GUIDE

Four types of compounds in the Ph. Eur. involve structures and nomenclature: parent substances, reagents, chemical reference substances and impurities.

1 *1. PARENT SUBSTANCES*

2 These are drawn by the scientific officers responsible for structure/nomenclature; they  
3 are then validated and named by the expert in structures and nomenclature (S/N expert).

4  
5 *2. REAGENTS*

6 Trivial names are given to reagents because they are shorter and do not contain special  
7 characters (prime symbols, Greek symbols), which complicate electronic searches when  
8 the CD-ROM and online versions are used. When a reagent corresponds to an impurity,  
9 its trivial name is placed in parentheses after the name of the impurity. In the test for  
10 related substances, the name of an impurity corresponding to a reagent may be specified  
11 in parentheses to establish a link to the prescribed limit.

12 *3. CHEMICAL REFERENCE SUBSTANCES*

13 These substances are named systematically, being given:

- 14 – either the name of the substance if it is the subject of a monograph (first priority), for  
15 example, *cetobemidone CRS*;  
16  
17 – or the name of an impurity of a given substance, in which case only the active  
18 part is mentioned (neither the salt nor the solvate), for example, *cetobemidone*  
19 *impurity A CRS* rather than *cetobemidone hydrochloride impurity A CRS*.

20  
21 *4. IMPURITIES*

22 Impurities are represented in a similar manner to the parent substance to make it clear  
23 that they are structurally analogous. They are drawn by the scientific officers responsible  
24 for structure/nomenclature. They are then validated and named by the S/N expert. They  
25 are grouped together, whenever permitted by the structures, to save space on the printed  
26 page (to reduce shipping costs) and to show their structural analogy, which makes the  
27 transparency list easier to understand. For editorial reasons, it has been decided not to  
28 indicate salts, counter-ions and solvates for impurities, unlike for parent substances.

29  
30 **III. STRUCTURE OF THE GUIDE**

31 This guide consists of 5 sections that deal with distinct aspects of chemical structures  
32 and nomenclature.

33  
34 *1. SECTION A*

35 This section describes the rules for drawing structures based on WHO recommendations  
36 but omits rules that are not applied by the Ph. Eur.

37 *2. SECTION B*

38 This section is intended for use as a reference to reply to most of the questions on  
39 structures and nomenclature raised during public enquiries or during sessions of the  
40 Commission, by setting out in print the editorial rules observed by the S/N expert.

41  
42 *3. SECTION C*

43 This section refers to the main structural classes of compounds whose structures are  
44 clearly analogous.

45 *4. SECTION D*

46 This section describes the nomenclature rules used as references by the Ph. Eur. to name  
47 substances.

## 5. SECTION E

This section consists of Frequently Asked Questions (FAQ) concerning structures and nomenclature and the standard answers to these questions.

## SECTION A - General rules for graphic representation

## INTRODUCTORY NOTE

It is important for chemical names and graphic representations in the Ph. Eur. to be defined as precisely and as unambiguously as possible. It is possible to draw many different 2-dimensional representations of a given molecule, all of which are entirely correct from a chemical standpoint. For editorial reasons and for consistency, conventions should be applied so that 2 persons, working independently, will draw a chemical structure in the same way. In this context, the Ph. Eur. follows the recommendations of the World Health Organisation on the drawing of structures [1]. If there is any ambiguity in the structure, the systematic name established in accordance with the rules of the International Union of Pure and Applied Chemistry [3,13] is used to remove this ambiguity. This guide is intended to lay out the rules followed by the Ph. Eur. and to illustrate them with examples from the monographs.

## A-1. GRAPHIC CONVENTIONS

## I. ORIENTATION OF THE STRUCTURES

Wherever possible, structures are drawn horizontally rather than vertically. They are orientated so that the atom with the highest number is on the left with the numbering of atoms decreasing from left to right [1].

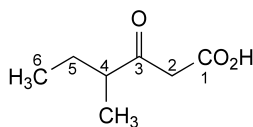


Figure A-1-1

## II. NUMBERING OF RINGS

Rings are numbered according to the rules of chemical nomenclature. Wherever possible, rings are numbered in a clockwise direction [1].

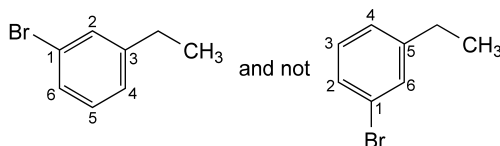


Figure A-1-2

## III. REPRESENTATION OF CHEMICAL GROUPS

## 1. General rule

Links between atoms are represented by dashes. Structures are shown in full. Polyatomic groups are drawn so that the atoms are shown as close as possible to the dashes representing the links [1].

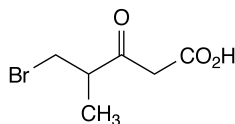


Figure A-1-3

## 2. Condensed form

Certain groups, however, are shown in condensed form [1].

Table A-1-1. Groups shown in condensed form

—CH <sub>3</sub>	methyl	—CHO	formyl
—CN	cyano	—NC	isocyano
—CO <sub>2</sub> H	carboxy	—CO <sub>2</sub> <sup>-</sup>	carboxylate
—OH	hydroxy	—OCH <sub>3</sub>	methoxy
—NH <sub>2</sub>	amino	—NO <sub>2</sub>	nitro
—SO <sub>3</sub> H	sulpho	—SO <sub>3</sub> <sup>-</sup>	sulphonate
—N <sub>3</sub>	azido		

It is strongly recommended to avoid using the following symbols for condensed representations, except in the case of polypeptides [1]:

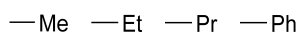
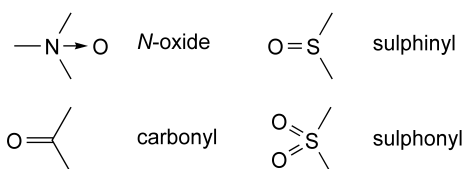


Figure A-1-4

## 3. Expanded form

A choice has been made on how to depict characteristic groups. The various classes of compounds are assembled from such groups [1].

Table A-1-2. Groups represented in expanded form



## IV. SPECIAL CASE OF *N*-OXIDES

At present there is no consensus on how to depict *N*-oxides, and the following representations co-exist:

- A simple N-O bond with a + charge on the N and a – charge on the O. However, this representation needlessly complicates structures and names.

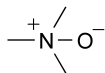


Figure A-1-5

- An arrow pointing from the N to the O.

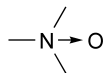


Figure A-1-6

1 – An N=O double bond.



5 Figure A-1-7

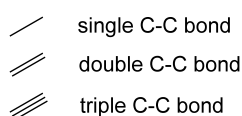
6 It has been decided to represent *N*-oxides in the Ph. Eur. with an arrow until a consensus  
7 is reached. Comments have already been made on this matter (see section E).  
8

9 **A-2. ACYCLIC STRUCTURES**

10 **I. REPRESENTATION OF COVALENT BONDS**

11 **1. Convention for different types of bonds**

12 In acyclic structures, a single bond is shown as a single dash, a double bond as a double  
13 dash, and a triple bond by a triple dash [1].  
14



19 Figure A-2-1

20 **2. Distinctive characteristic of double bonds**

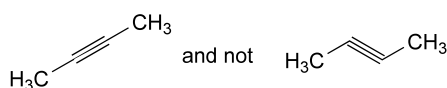
21 In linear structures, double bonds are shown as centered between the two atoms (in cyclic  
22 structures, double bonds are shown inside the ring; see section A-3).  
23



27 Figure A-2-2

28  
29 **3. Distinctive characteristic of triple bonds**

30 Triple bonds impose a fixed linear structure. The two remaining bonds of the carbon  
31 atoms are drawn along the axis of the triple bond.  
32



35 Figure A-2-3

36  
37 **II. REPRESENTATION OF CARBON CHAINS**

38 **1. Angular representation**

39 In the Ph. Eur., the entire carbon chain is shown. The chains are drawn as lines at angles to  
40 one another since this representation makes it easier to describe the centres of asymmetry.  
41 The carbon and hydrogen atoms are intentionally omitted from the chain to simplify the  
42 appearance of the main structure; the chain is simply represented by a series of links at an  
43 angle to one another. Only the terminal groups are shown. Groups on the left-hand end of  
44 the formula are inverted, with the hydrogen atoms to the left of the carbon atom. [1].  
45

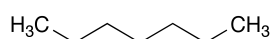


Figure A-2-4

## 2. Chains containing repeated groups

### Parentheses

Parentheses are not usually used to show that several identical groups are linked to the same atom or to a group linked to the principal chain [1] except for certain very complicated structures and for the definition of substituent groups in the case of grouped impurities (see section B-1).

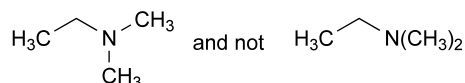


Figure A-2-5

### Square brackets

Square brackets are used to indicate the repetition of a large number of identical groups. If one of these groups terminates the chain or bears a heteroatom, it is shown outside the square brackets [1].

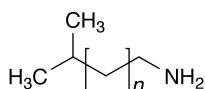


Figure A-2-6

In the Ph. Eur., square brackets are used only when a structure is too large to fit into a single column or when substituent groups are defined, as in the case of grouped impurities (see section B-1). It is aesthetically preferable, but not obligatory, to use square brackets rather than fold back the structure in the plane. The repeating unit may be linear (Figure A-2-7) or branched (Figure A-2-8).

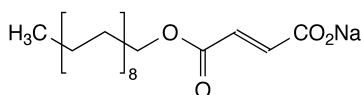


Figure A-2-7. – Sodium stearyl fumarate (1567)

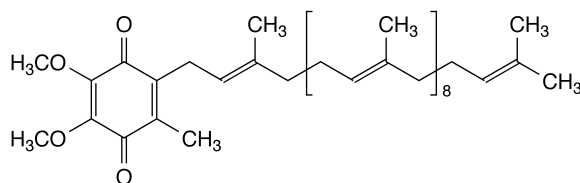


Figure A-2-8. – Ubidecarenone (1578)

### III. STEREOCHEMICAL INDICATIONS

Only the carbon atoms and hydrogen atoms at the end of a chain are shown systematically. However, hydrogen atoms can be shown if this is necessary to specify the stereochemistry.

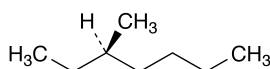


Figure A-2-9

It should be noted that stereochemical indications are given in the Ph. Eur. wherever necessary (see section A-7). Whenever there is ambiguity, this should be clarified.

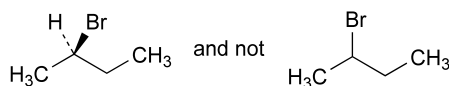


Figure A-2-10

5  
6  
7  
8  
9

To specify the stereochemistry of a centre of asymmetry, a single bond can be replaced by a broken line (bond projecting behind the plane of the paper) or a filled wedge (bond projecting in front of that plane) [1]. The rules for stereochemical representation are described in more detail in section A-7 of this guide.

10  
11  
12

Double bonds and triple bonds cannot be represented spatially; if such a bond is attached to a centre of asymmetry, it is shown in the plane of the paper and the single bonds are depicted to show the stereochemistry [1].

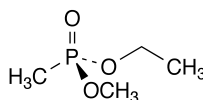


Figure A-2-11

17  
18

### 19 A-3. CYCLIC STRUCTURES

20

#### 20 I. BASIC GEOMETRIC FORMS

21  
22

##### 21 1. Monocyclic systems

23  
24

Rings are shown as regular polygons when they consist of up to 8 carbon atoms [8].

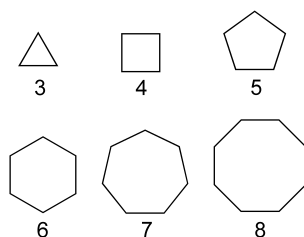


Figure A-3-1

33  
34  
35

Rings with more than 8 vertices are often shown with re-entrant angles. The Chemical Abstract Service (CAS) recommends that they should be drawn like amalgamated rings with 5, 6 or 7 vertices [8].

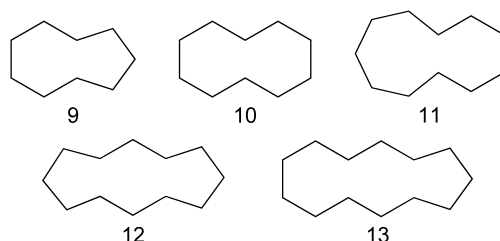


Figure A-3-2

43  
44  
45

##### 45 2. Polycyclic systems

46  
47

Wherever possible, the regularity of the polygons is maintained in the drawing of fused cyclic compounds [8].

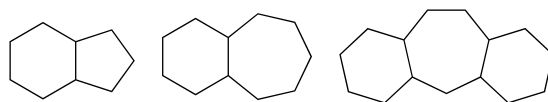


Figure A-3-3

However, in fused polycyclic systems the polygons may often be distorted in order to maintain the symmetry of the structure [8].

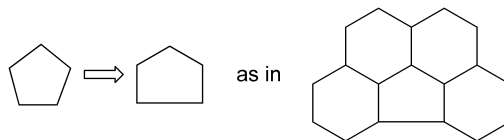


Figure A-3-4

## II. HETEROCYCLES

Rings are shown in full. The symbols of the carbon atoms that form the ring are not shown. The hydrogen atoms attached to them are not represented unless they are needed to show stereochemistry. Heteroatoms are shown with all the hydrogen atoms attached to them but without linking dashes [1].

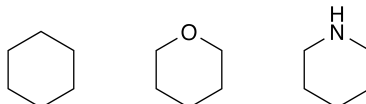


Figure A-3-5

Trivial names exist for heterocycles that can be used to establish a nomenclature.

## III. UNSATURATED SYSTEMS

In aromatic systems, all the double bonds are shown [1]. The Ph. Eur. does not use a circle to depict delocalised electrons.

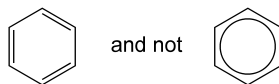


Figure A-3-6

### 1. Monocyclic systems

Double bonds are shown inside the ring rather than being centred between 2 carbon atoms.

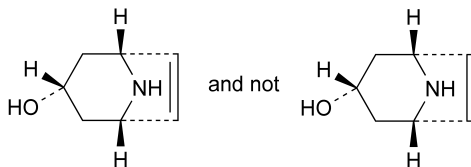


Figure A-3-7

In monocyclic compounds, by convention, double bonds should be arranged to have the lowest possible numbering [1].

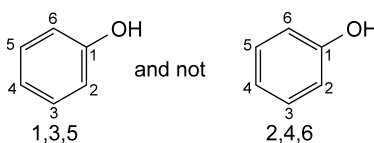
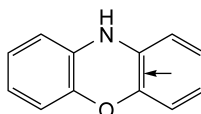


Figure A-3-8

1 **2. Polycyclic systems**

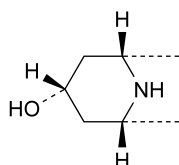
2 In fused polycyclic systems a double bond should form the fusion bond nearest to the  
3 right-hand side, where there is a choice possible [8].



8  
9 Figure A-3-9

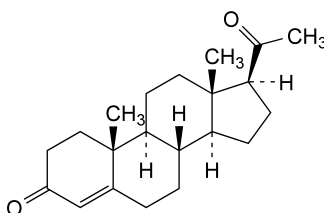
10 **IV. SUBSTITUENTS AND INDICATED HYDROGENS**

11 Substituents and indicated hydrogens are normally placed outside monocyclic or polycyclic  
12 systems if the configuration of the rings permits this (sufficient space).



18  
19 Figure A-3-10

20 In the case of steroids, terpenes and alkaloids and of crowded structures, substituents  
21 and indicated hydrogens attached at bridgeheads can be displayed inside the rings of  
22 polycyclic structures [8].



28  
29 Figure A-3-11

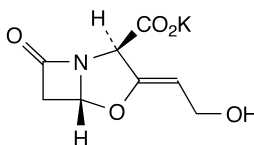
30 **A-4. IONIC STRUCTURES**

31 **I. GENERAL RULES OF PRESENTATION**

32 **1. Order of appearance of ions**

33 The following rules apply to the drawings of parent substances but do not apply either to  
34 impurities, which are drawn as conjugate acids and bases, or to quaternary ammonium  
35 salts, which are drawn without showing the counter-ion.

36 In most cases, the ion can be represented in condensed form (see section A-1); the  
37 structure is then drawn as if it were not an ion. This is done in particular for salts of  
38 alcoholates, carboxylates, and phosphoric and sulphuric esters



45  
46 Figure A-4-1. – *Potassium clavulanate, diluted (1653)*

47 In general, in ionic structures, the cationic part is placed on the left and the anionic part  
on the right [1]. Cations and anions are not separated by a comma or a dot.

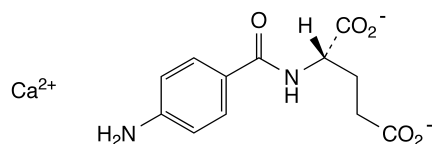
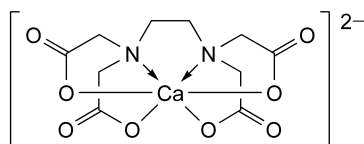


Figure A-4-2

7 **2. Symbolising the charge**

8 Ionic charges are not encircled and are shown as superscripts on the right of the charged  
9 atom.



15 and not

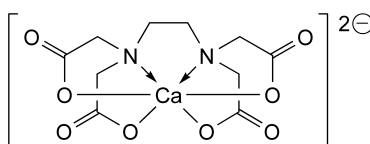
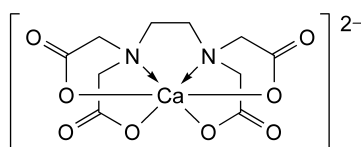


Figure A-4-3

21 Multiple charges are indicated by writing  $n+$  or  $n-$  and not by writing the + or - symbol  $n$   
22 times [1].



29 and not

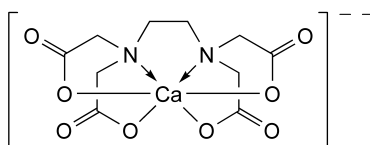


Figure A-4-4

34  
35 **3. Placing the charge on the molecule**

36 A terminal charge is shown as a superscript on the right of the group concerned, unless  
37 the order of atomic symbols in the group is reversed, when the charge is shown as  
38 a superscript on the left.

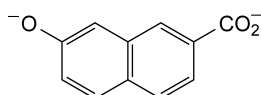


Figure A-4-5

44 In structures where the positions of the charges cannot be indicated precisely, the  
45 structure is put in square brackets, with the overall charge placed outside them as a  
46 superscript on the right.  
47

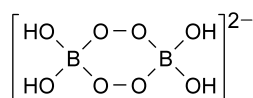
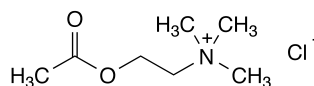
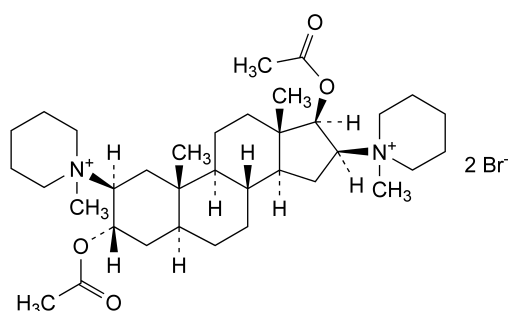


Figure A-4-6

In a horizontal acyclic chain, if there is no space for a superscript on the right of the atom concerned, the charge can be shown immediately above that atom.

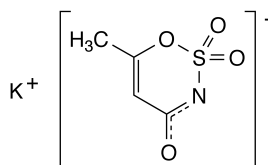
Figure A-4-7. – *Acetylcholine chloride (1485)*

When a ring is involved, the charge is usually placed outside the ring, unless it is impossible to place the charge without ambiguity [1].

Figure A-4-8. – *Pancuronium bromide (0681)*

## II. DELOCALISED CHARGE

In structures with delocalised charge, the structure is put in square brackets, with the charge sign outside them as a superscript on the right [1].

Figures A-4-9. – *Acesulfame potassium (1282)*

## III. SUBSTANCES WITH SEVERAL GROUPS THAT CAN FORM SALTS

### 1. Case of several acid groups forming different salts

When substances contain several acid groups to which the various cations cannot easily be attributed, ionic forms may be used [1].

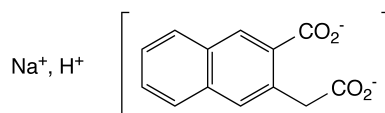


Figure A-4-10

### 2. Zwitterions

For zwitterions, the + and – charges are drawn and placed according to the rules described above [1], provided that these charges are indispensable.

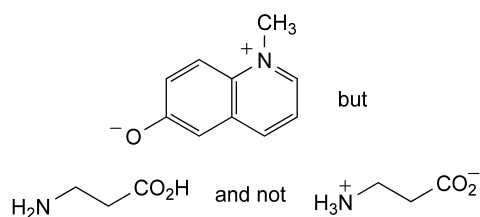


Figure A-4-11

8  
9 **IV. SOLVATE SALTS**

10 The Ph. Eur. does not follow the IUPAC recommendation to use a dot to symbolise  
11 solvates of inorganic compounds. It uses a comma for both organic and inorganic  
12 compounds. It should be noted that for molecular formulae, the Ph. Eur. does not put  
13 spaces between the comma, the number of molecules of water and the molecular formula  
14 for water. However, these are separated by spaces in structural formulae.

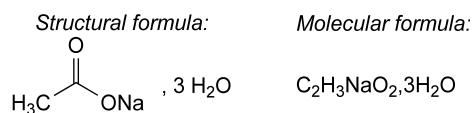


Figure A-4-12

20  
21 **V. METAL SALTS**

22 **1. Inorganic salts**

23  
24 Metal salts (for example, KMnO<sub>4</sub>), like inorganic acids (for example, HClO<sub>4</sub>), are shown  
25 without charges or bonds. Structural formulae are not given for these substances in the  
26 monographs since the molecular formula is considered to be sufficient, except in the  
27 case of co-ordination compounds for which the stereochemistry of the complexes must  
28 be specified (see section A-6). If metal salts of inorganic acids include several metals, the  
29 symbols for the metals are shown in alphabetical order (for example, K<sub>2</sub>NaPO<sub>4</sub>). In salts of  
30 inorganic acids, the metal precedes the hydrogen (for example, NaH<sub>2</sub>PO<sub>4</sub>) [1]. Molecules  
31 of water of crystallisation or of substances of solvation follow the formula of the salt,  
32 from which they are separated by a comma, without a preceding or following space (for  
33 example, H<sub>3</sub>PO<sub>4</sub>·5H<sub>2</sub>O or Na<sub>2</sub>SO<sub>4</sub>·<sup>3</sup>/<sub>2</sub>H<sub>2</sub>O).

34 **2. Organic salts**

35  
36 In the metal salts of organic acids and the metal compounds of alcohols, phenols (and  
37 their sulphur, selenium and tellurium analogues), amines and amides, the metal symbol  
38 usually replaces the acid hydrogen [1].

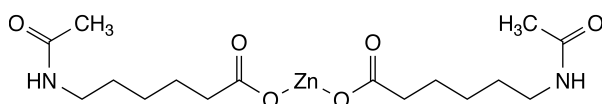
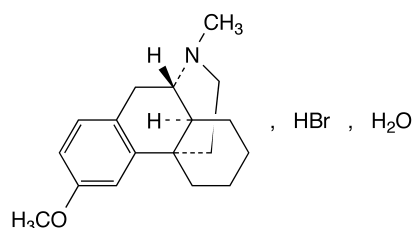


Figure A-4-13. – Zinc acexamate (1279)

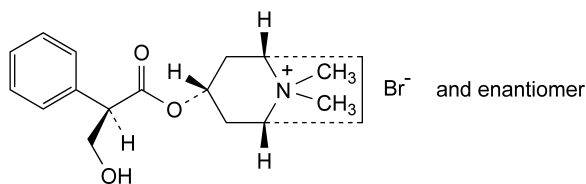
44  
45 **VI. AMINE SALTS**

46 Amine salts are shown with the structure of amine on the left (as if it were in the  
47 ammonium form) and, after a comma, the formula of the inorganic acid on the right [1].



7  
8 Figure A-4-15. – *Dextromethorphan hydrobromide (0020)*

9 Quaternary ammonium salts and other compounds with a positive charge on a heteroatom  
10 (P, As, Sb, O, S, Se, Te) are shown in ionic form (with + and – charges), the 2 ions being  
11 separated by a space [1].



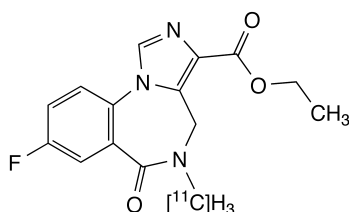
17  
18 Figure A-4-16. – *Methyldatropine bromide (0511)*

19 **A-5. ISOTOPICALLY MODIFIED COMPOUNDS**

20 In most monographs on radiopharmaceutical substances in the Ph. Eur. a structural or  
21 molecular formula is not given. However, it may be necessary to indicate them in certain  
22 monographs.

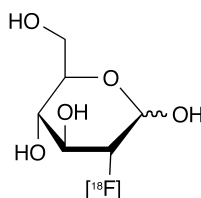
23 In an isotopically modified compound, the isotope used is indicated by its mass number  
24 placed as a superscript on the left of the symbol of the element concerned. Deuterium and  
25 tritium are written <sup>2</sup>H and <sup>3</sup>H, respectively [1].

26 The carbon atom in a ring or a simplified angular-chain representation is explicitly  
27 designated when its mass number is shown [1].



35 Figure A-5-1. – *Flumazenil (N-[<sup>11</sup>C]methyl) injection (1917)*

36 To indicate isotopic labelling (partial replacement of the atom by the nuclide shown),  
37 atomic symbols in formulae should be in square brackets [1].



44 Figure A-5-2. – *Fludeoxyglucose (<sup>18</sup>F) injection (1325)*

45 When atomic symbols and formulae are drawn without square brackets, the compounds  
46 are assumed to be isotopically substituted, i.e., the atom concerned is completely replaced  
47 by the nuclide shown [1]. This is not usually the case in the monographs.

## A-6. CO-ORDINATION COMPOUNDS

## I. NON-CYCLIC LINEAR STRUCTURES

According to current usage [2], for example, *Sodium nitroprusside (0565)*:

$\text{Na}_2[\text{Fe}(\text{CN})_5(\text{NO})], 2\text{H}_2\text{O}$ , a noncyclic structure is constructed in the following order:

- the symbol of the central atom placed on the left,
- ionic ligands with cations first and then anions,
- neutral ligands.

Polyatomic ligands are placed in parentheses, with the atom linked to the central atom on the left, for example,  $\text{Na}_2[\text{Fe}(\text{CN})_5(\text{NO})], 2\text{H}_2\text{O}$ . If several identical ligands are attached to the central atom, their number is indicated as a subscript to the right. In each class of ligand, the symbol of the linking atom is shown, followed by the other atoms in alphabetical order. The complete formula of the co-ordination entity (neutral group or complex ion) is placed in square brackets [1]. No spaces should be left between representations of ionic species within the formula of a co-ordination compound.

If the charge of the co-ordination entity needs to be specified, it is placed outside the square bracket as a right superscript [1]. The individual charges usually carried by the central atoms and the ligands are not normally shown; they may, however, be shown in structural formulae when it is difficult to show all the co-ordination links [1].

## II. CYCLIC STRUCTURES

The rings follow the conventions for cyclic compounds (see section A-3). Where possible, the metal atom is placed in the centre of the group. Square brackets are placed round every co-ordination entity containing one or more rings, even if the charge is zero [1].

‘Sandwich’ structures are shown with the rings connected to the central atom by a line starting from inside the cycle and passing through one side [1]. Benzene rings and condensed benzene systems in ‘sandwich’ compounds are drawn with alternating single and double bonds. Pentagonal and heptagonal rings are shown with a circle inside [1].

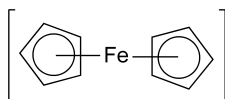
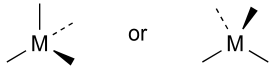







Figure A-6-1

## III. STEREOCHEMISTRY

The stereochemistry of mononuclear complexes is expressed by means of special descriptors formed from an abbreviation for the central atom geometry and the co-ordination number [1].

Table A-6-1

Polyhedral symbol	Representation
<i>T-4</i>	
<i>SP-4</i>	
<i>TBPY-5</i>	
<i>SPY-5</i>	
<i>OC-6</i>	
<i>PBPY-7</i>	

Co-ordination bonds of the shared-electron-pair type are shown as arrows directed towards the central atom (see Figure A-6-2); these arrows do not project in front of or behind the paper and must be drawn in the plane of the paper. The other co-ordination entities are orientated accordingly.

### 1. *T-4*: tetrahedral complexes

They are described by the chirality symbols (*R*) and (*S*); they are shown in the same way as stereogenic carbon atoms, following the same convention [1].

### 2. *SP-4*: square planar complex

The 4 co-ordination links are shown in the plane of the paper [1].

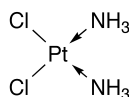


Figure A-6-2. – *Cisplatin (0599)*

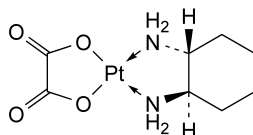


Figure A-6-3. – *Oxaliplatin (2017)*

### 3. *TBPY-5*: trigonal bipyramidal complex

The reference axis is shown in the plane of the paper; of the 3 other, equatorial ligands, 1 is assumed to be also in the plane of the paper, 1 in front of it and the other behind it [1].

### 4. *SPY-5*: square pyramidal complex

The reference axis with its lone co-ordinating atom is shown in the plane of the paper and 4 co-ordination links are assumed to be in a plane perpendicular to the reference axis, 2 in front of and 2 behind the plane of the paper [1].

1 **5. OC-6: octahedral complex**

2 Two co-ordination links are shown as the axis in the plane of the paper and 4 are assumed  
3 to be in a plane perpendicular to the reference axis, 2 in front of and 2 behind the plane of  
4 the paper [1].

5 **6. PBPY-7: pentagonal bipyramidal complex**

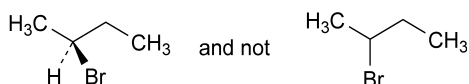
6 Two co-ordination links are shown as the axis in the plane of the paper and the 5 other  
7 co-ordination links are shown as their projection onto the plane perpendicular to this axis:  
8 1 in the plane of the paper, 2 in front of it and 2 behind it [1].

9 **A-7. STEREOCHEMISTRY**

10 **I. GENERAL ASPECTS**

11 The rules are applied when permitted by the spatial constraints of the molecule.

12 Stereochemistry is the branch of chemistry concerned with the 3-dimensional arrangement  
13 of atoms in molecules, and stereoisomers are isomers with no differences in connectivity  
14 or bond multiplicity, but whose atomic spacial arrangements differ. As already mentioned,  
15 a broken line denotes a bond projecting behind the plane of the paper and a filled wedge  
16 denotes one projecting in front of that plane. A line of normal thickness denotes a bond  
17 lying in the plane of the paper [1]. Hydrogen is represented by its symbol 'H' whenever a  
18 configuration has to be shown [1].



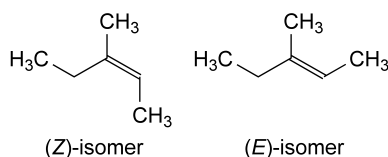
22  
23  
24 Figure A-7-1

25 **II. GEOMETRIC ISOMERISM**

26 **1. (EZ) isomerism**

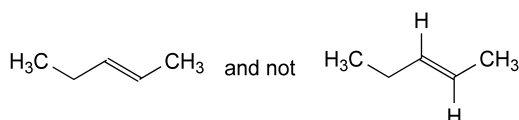
27 *General rules for representation*

28 Compounds containing carbon-carbon double bonds are represented as lines at an angle to  
29 each other; the representations of (*EZ*) isomers show that the (*Z*) configuration produces  
30 a bend in the chain whereas the (*E*) configuration does not.



33  
34  
35  
36  
37 Figure A-7-2

38 It should be noted that the hydrogens attached to the 2 carbon atoms forming the double  
39 bond are omitted [1] in the case of (*EZ*) isomerism.



42  
43  
44 Figure A-7-3

45 *(EZ) isomerism of imines/oximes*

46 The same convention is used for the isomers of imines/oximes and compounds containing  
47 several double bonds [1].

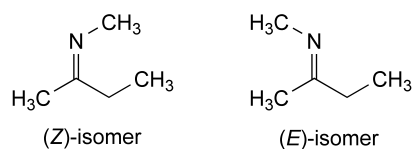
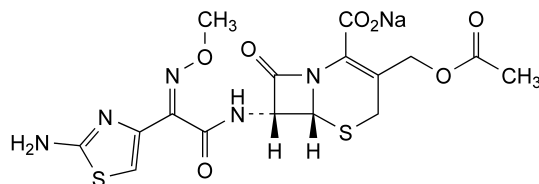
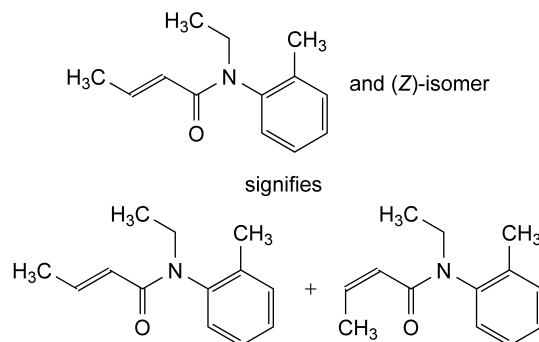


Figure A-7-4

Figure A-7-5. – *Cefotaxime sodium (0989)**How to represent a mixture of (E) and (Z) isomers of a compound*

12  
13  
14  
15  
16  
17

To describe a mixture of (*Z*) and (*E*) isomers of a compound, the (*E*) isomer is shown and the legend ‘and (*Z*)-isomer’ is added to the right of the structure. This type of representation does not necessarily mean that the 2 isomers are present in the same quantities.

Figure A-7-6. – *Crotamiton (1194)**How to represent the stereochemistry of a double bond when it is undefined*

27  
28  
29  
30  
31  
32  
33

In certain cases, the stereochemistry of a double bond may be unknown, i.e., the group of experts does not know whether the impurity consists of the (*E*) isomer or the (*Z*) isomer or both. This lack of knowledge is represented by convention by drawing the double bond so that it is aligned with 1 of the 2 C-C bonds next to it (see Figure A-7-7).

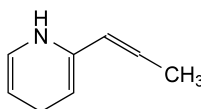


Figure A-7-7

37  
38  
39  
40

The information conveyed by this representation is considerably different from that conveyed by the representation of a mixture of both isomers (Figure A-7-8).

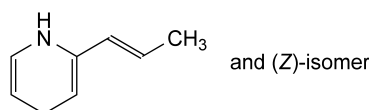


Figure A-7-8

45  
46

Comments have already been made on this subject (see section E).

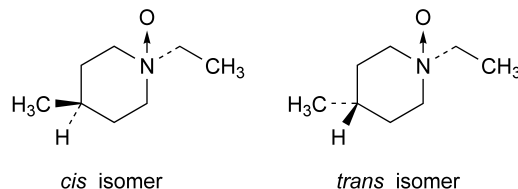
47  
48

*Summary of modes of representation for (E/Z) isomerism*

1 See Table A-7-1.

2 **2. *cis-trans* Ring isomerism**

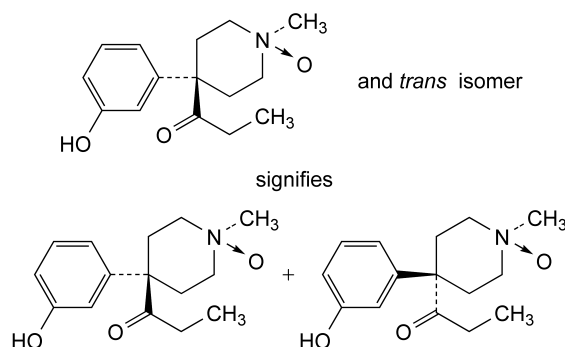
3 *cis-trans* Ring chirality is defined for a molecule with a plane of symmetry as the existence  
4 of *cis* and *trans* isomerism according to the arrangement of substituents on a ring with  
5 respect to the plane of this ring. *N*-oxides of cyclic amines, present as oxidation impurities  
6 in the monographs on certain cyclic amines, display this type of isomerism.  
7



13 Figure A-7-9

14 *How to represent a mixture of cis-trans ring isomers of a compound*

15 To describe a mixture of *cis-trans* ring isomers of a compound, the *cis* isomer is shown  
16 and the legend 'and *trans* isomer' is added to the right of the structure. This type of  
17 representation does not necessarily mean that the 2 isomers are present in the same  
18 quantities.  
19



30 Figure A-7-10. – Impurity A of *Cetobemidone hydrochloride (1746)*

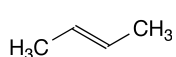
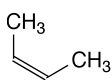
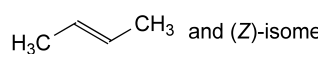
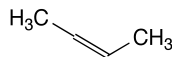
31 Comments have already been made on this subject (see section E).

32 **III. COMPOUNDS WITH ONE CENTRE OF ASYMMETRY**

33 *Rules for representation of enantiomers*

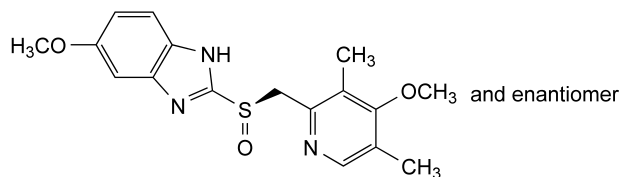
34 An enantiomer is one of a pair of molecular entities that are mirror images of each other  
35 and non-superposable [4]. Diastereoisomerism is defined as stereoisomerism other than  
36 enantiomerism. Diastereoisomers are stereoisomers not related as mirror images [4].  
37

38 Table A-7-1. – Representations of (*EZ*) isomerism

Representation	Isomer present	Type of stereochemistry
	( <i>E</i> )	defined
	( <i>Z</i> )	defined
 and ( <i>Z</i> )-isomer	( <i>E</i> ) + ( <i>Z</i> )	defined
	( <i>E</i> ) or ( <i>Z</i> ) or ( <i>E</i> ) + ( <i>Z</i> )	undefined

1 *Enantiomerism of sulphoxides*

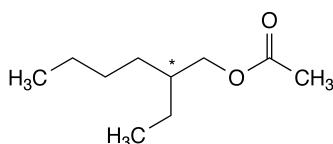
2 Sulphoxides show enantiomerism. Questions are frequently asked about this (see  
3 section E).



10 Figure A-7-12. – *Omeprazole (0942)*

11 *Undefined chirality*

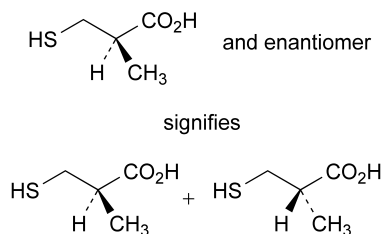
12 If the chirality of the centre of asymmetry is unknown or not specified, the bonds adjoining  
13 atoms or groups to the chiral atom are shown as lines of 'normal' thickness and an asterisk  
14 identifies this chiral centre [1]. Monosaccharides and macrolides are exceptions for which  
15 a bond with an unknown orientation is symbolised by a wavy line.



21 Figure A-7-13.

22 *How to represent a mixture of enantiomers of a compound*

23 To describe a mixture of 2 enantiomers of a compound, the (*R*)-isomer is shown and the  
24 phrase 'and enantiomer' is added to the right of the structure. This type of representation  
25 does not necessarily mean that the 2 isomers are present in the same quantities. If they  
26 are present in the same quantities, the term used is racemate or racemic mixture.



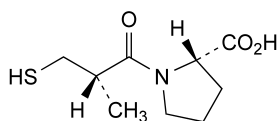
35 Figure A-7-14

36 *Summary of modes of representation of (RS) isomerism*

37 See Table A-7-2.

38 *IV. COMPOUNDS WITH SEVERAL CENTRES OF ASYMMETRY*

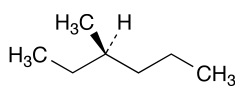
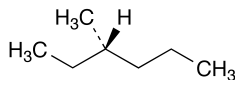
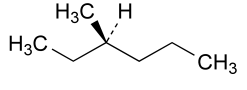
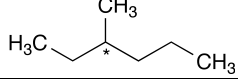
39 In compounds containing several centres of asymmetry, the same conventions apply to  
40 each of these centres [1].



46 Figure A-7-15. – *Captopril (1079)*

47 This is also the case for the representation of *cis-trans* isomerism relative to a ring.

Table A-7-2. – Representations of (*RS*) isomerism

Representation	Isomer present	Type of stereochemistry
	( <i>R</i> )	defined
	( <i>S</i> )	defined
 and enantiomer	( <i>R</i> ) + ( <i>S</i> )	defined
	( <i>R</i> ) or ( <i>S</i> ) or ( <i>R</i> ) + ( <i>S</i> )	undefined

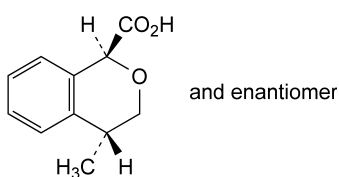


Figure A-7-16

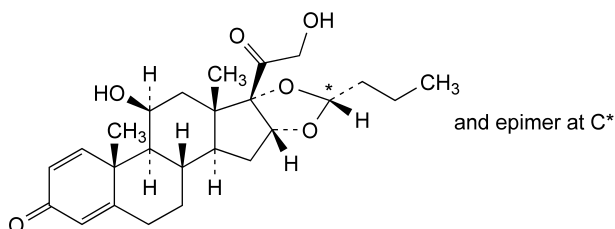
## 1. Epimers

### *General rules of presentation*

Epimers are diastereoisomers that have the opposite configuration at only 1 of 2 or more tetrahedral stereogenic centres present in the respective molecular entities [4]. If a compound has 2 centres of asymmetry, there are 4 different stereoisomers. Two epimers of this compound have the same configuration at one centre of asymmetry and a different configuration at the other.

### *Mixtures of 2 epimers*

To define a mixture of 2 epimers of a compound containing at least 2 asymmetrical carbons, draw the (*R*) isomer with respect to the centre of asymmetry of interest marked with an asterisk and add the legend 'and epimer at C\*' to the right of the structure [1].

Figure A-7-17. – *Budesonide* (1075)

If the 2 centres of asymmetry are, for example, a carbon atom and a nitrogen atom, draw the (*R*) isomer with respect to the centre of asymmetry of interest, without an asterisk, and add the legend 'and epimer at C' or 'and epimer at N' to the right of the structure.

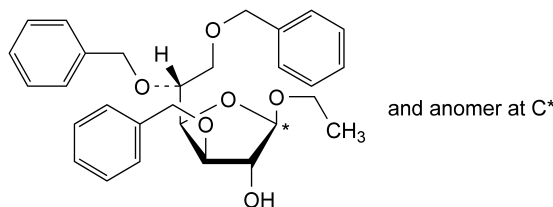
## 2. Anomers

### *General rules of presentation*

1 A special case of epimerism is that of anomerism of sugars. Anomers are defined as  
2 diastereoisomers of glycosides, hemiacetals or related cyclic forms of sugars, or related  
3 molecules differing in configuration only at C-1 of an aldose, C-2 of a 2-ketose, etc [4].

#### 4 *Mixtures of 2 anomers*

5 To define a mixture of 2 anomers of a compound containing at least 2 centres of  
6 asymmetry, draw the (*R*) isomer with respect to the centre of asymmetry of interest  
7 marked with an asterisk and add the legend 'and anomer at C\*' to the right of the structure [1].  
8



15 Figure A-7-18. – *Tribenoside (1740)*

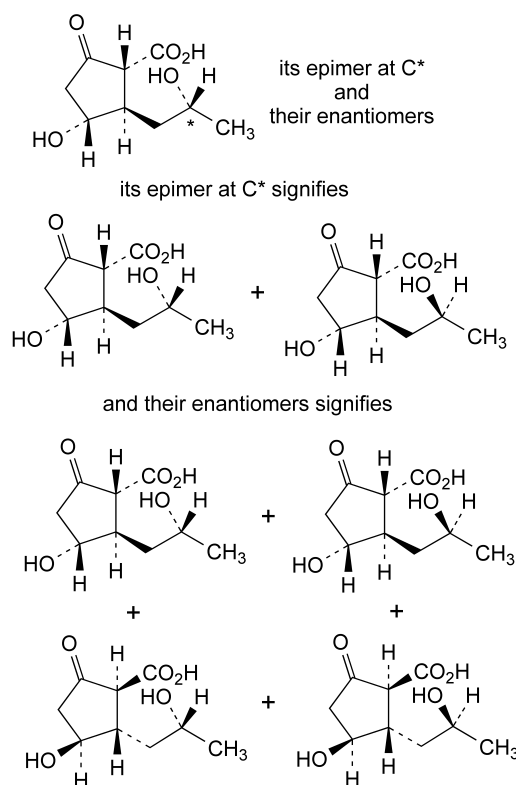
### 16 **3. Mixtures of more than 2 stereoisomers**

17 To define a mixture of more than 2 isomers of a compound containing at least 2 centres of  
18 asymmetry, proceed in 2 steps:

- 19
- 20 – first, identify the centres of asymmetry that permute together and isolate 2 groups of  
21 enantiomers;
  - 22 – second, isolate the marginal centre of asymmetry (C\*), which is the centre used to  
23 establish the link between the 2 groups of enantiomers.
- 24

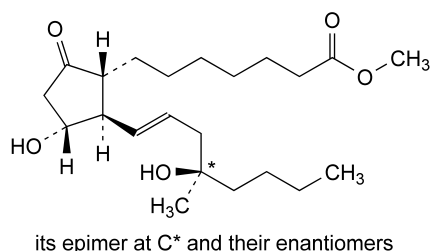
25 Next, show the (*R*) isomer at C\* and add the legend 'and epimer at C\*', which defines both  
26 of the above mentioned groups. The legend 'and their enantiomers' is also added, which  
27 means that for the 2 defined epimers there are 2 isomers, giving a total of 4 isomers. This  
28 construction can be broken down for greater clarity (see Figure A-7-19).

29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47

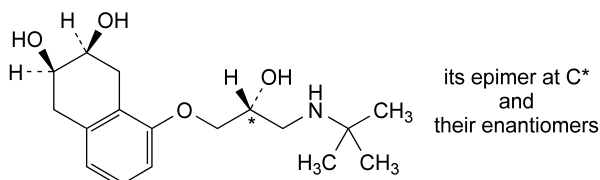


22 Figure A-7-19

23 The Ph. Eur. includes monographs on such mixtures (Figures A-7-20 and A-7-21).



33 Figure A-7-20. – *Misoprostol (1731)*



42 Figure A-7-21. – *Nadolol (1789)*

43 **V. STEREODESCRIPTORS IN SYSTEMATIC NOMENCLATURE**

44 **(R) and (S)**

45  
46 These symbols (proposed by Cahn, Ingold and Prelog) indicate the absolute configuration  
47 around 4-co-ordinate (quadriligant) and 6-co-ordinate (sexiligant) stereogenic centres.

**(r) and (s)**

The stereodescriptors *r* and *s* (lower-case italics) are used instead of *R* and *S* to designate the stereochemistry of a carbon atom that is stereogenic because it has 4 different substituents (2 of which differ only in terms of stereochemistry), but which is part of a structure that is non-chiral due to a plane of symmetry passing through this carbon atom.

**SECTION B - Graphic rules specific to the Ph. Eur.****INTRODUCTORY NOTE**

The general rules described in section A of this guide are not always strictly followed and sometimes have been adapted to the scientific and editorial requirements of the Ph. Eur. In addition, unlike the WHO, the Ph. Eur. describes impurities (Impurities section of monographs, usually called the 'transparency list'). As the structures are not available in the INN database, they are drawn and named entirely by the Ph. Eur. for its users.

- Large structures in particular must be drawn so that they fit into one of the 2 columns on a page (i.e., maximum 8.5 cm wide).
- For reasons of clarity, impurities are not drawn independently of the parent substance that is the subject of the monograph. The impurities are in fact related to this substance since most of the time they are its degradation products. The drawings of the impurities are derived from the structure of the parent substance, modified as necessary.
- For editorial reasons (to save space), the structures of impurities are grouped together wherever possible. For reasons of clarity, a special effort should be made to determine generic structures (consisting of the parts in common) and to define the simplest possible substituent groups (variable parts).

This section describes the specific rules to be applied to achieve these 3 objectives.

**B-1. GROUPING OF STRUCTURES*****I. GENERAL PRINCIPLES***

The Ph. Eur. groups structures together in the 2 following cases:

- the parent substance is defined as a mixture of compounds with similar structure;
- the impurities in the transparency list show clear structural analogy ('specified impurities' will be grouped separately from 'other detectable impurities').

Grouping of structures involves:

- identifying the constant part among the molecules, so as to define the largest common denominator, and drawing this generic structure;
- identifying the parts that vary between structures and defining the relevant substituent groups to be specified in the name.

Section A of this guide describes the general rules for drawing generic structures; an additional objective is to show the structures of the impurities in a manner that makes it clear that they are related. The section below describes the recommendations to be followed for substituent groups.

***II. THE VARIOUS SUBSTITUENT GROUPS***

The Ph. Eur. groups structures together by means of substituent groups, whose nature depends on the group to be described and on the bond in particular:

- 1 – *monovalent groups* are attached to the generic structure by a single bond;  
 2  
 3 – *divalent groups* are attached to the generic structure by 2 single bonds or 1 double  
 4 bond;  
 5 – *repetition units* describe a repeating unit (in the carbon chain of the generic structure).

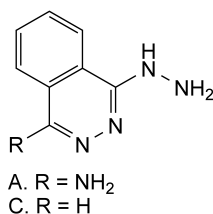
6 It is possible to describe 1 or more substituent groups of a given type for the same  
 7 structure. The editorial recommendations to be followed are described below for each  
 8 type. Substituent groups must be as small as possible and generic structures must be as  
 9 large as possible. For certain structures shown together, the various types of substituent  
 10 groups may be combined on the same structure. These should therefore be considered  
 11 independently and the rules described below for each type should be followed.

### 13 1. Monovalent groups

14 A monovalent group is an atom or group of atoms linked to a generic structure by only 1  
 15 single bond. Several cases can be described depending on whether 1, 2 or more groups  
 16 have to be defined; there are specific editorial rules for each case.

#### 18 Case No. 1 – 1 monovalent group

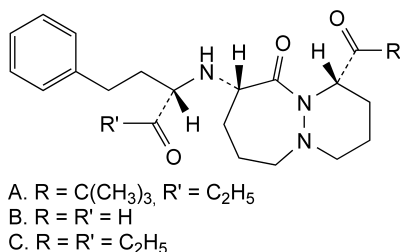
19 When there is a single monovalent group it is designated R. The group is specified just  
 20 before the name of each entity.



29 Figure B-1-1

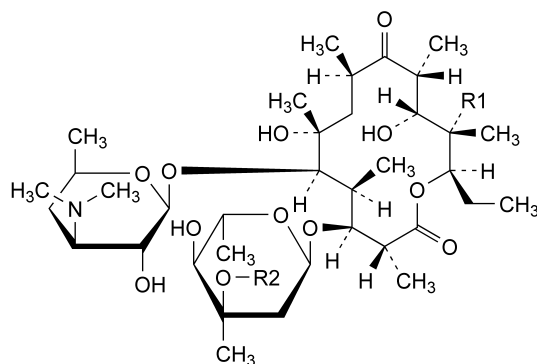
#### 31 Case No. 2 – 2 monovalent groups

32 When there are exactly 2 monovalent groups, they are usually designated R and R'.



42 Figure B-1-2

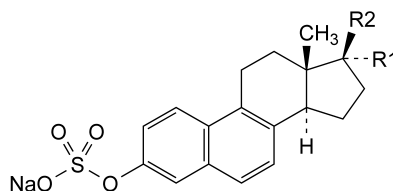
43 They may, however, be designated R1 and R2, in particular in crowded structures where  
 44 the prime symbol might not be noticed. It should be noted that the number given to  
 45 the substituent group should not be written as a subscript to avoid confusion with the  
 46 subscripts in chemical groups (indicating the number of times that the atom is present  
 47 in a polyatomic group).



Erythromycin	Mol. Formula	M <sub>r</sub>	R1	R2
A	C <sub>37</sub> H <sub>67</sub> NO <sub>13</sub>	734	OH	CH <sub>3</sub>
B	C <sub>37</sub> H <sub>67</sub> NO <sub>12</sub>	718	H	CH <sub>3</sub>
C	C <sub>36</sub> H <sub>65</sub> NO <sub>13</sub>	720	OH	H

Figure B-1-3

The case of  $R + R' = O$  or  $R1 + R2 = O$  is a special case that makes it possible to show a ketone or an imine with its homologous alkane/amine/alcohol, as illustrated below.



E: R1 = OH, R2 = H  
 F: R1 = H, R2 = OH  
 H: R1 + R2 = O

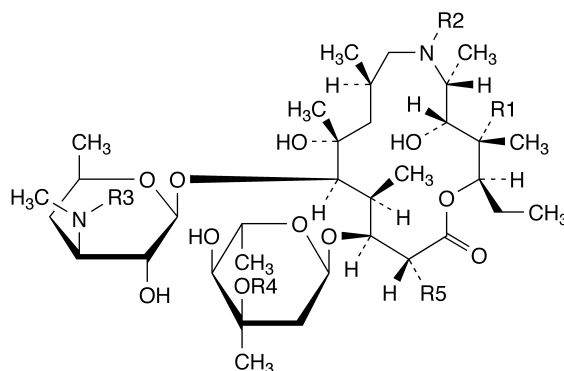
Figure B-1-4

It is also possible to do this using divalent groups (see case No. 4).

Comments have already been received on this subject (see section E).

### Case No. 3 – more than 2 monovalent groups

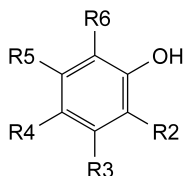
When more than 2 monovalent groups are identified, they are designated R1, R2, R3... R<sub>n</sub>.



A: R1 = OH, R2 = H, R3 = R4 = R5 = CH<sub>3</sub>  
 B: R1 = H, R2 = R3 = R4 = R5 = CH<sub>3</sub>  
 C: R1 = OH, R2 = R3 = R5 = CH<sub>3</sub>, R4 = H  
 D: R1 = OH, R2 = R3 = R4 = CH<sub>3</sub>, R5 = CH<sub>2</sub>OH  
 F: R1 = OH, R2 = R4 = R5 = CH<sub>3</sub>, R3 = CHO  
 G: R1 = OH, R2 = R4 = R5 = CH<sub>3</sub>, R3 = SO<sub>2</sub>-C<sub>6</sub>H<sub>4</sub>-CH<sub>3</sub>  
 I: R1 = OH, R2 = R4 = R5 = CH<sub>3</sub>, R3 = H

Figure B-1-5

In principle, the number assigned to a substituent group is independent of its position in the structure, except in the simple case where the groups are attached to an aromatic ring, and where systematic numbering applies. It should be noted that R1 is almost always missing (the principal group is in position 1). Other intermediary substituent groups may also be missing (R3 may be missing without affecting the numbering of R2, R4, R5 and R6).



A : R2 = R3 = R4 = R5 = R6 = H  
 B : R2 = CH<sub>3</sub>, R3 = R4 = R5 = R6 = H  
 C : R2 = R3 = R5 = R6 = H, R4 = CH<sub>3</sub>  
 D : R2 = R6 = CH<sub>3</sub>, R3 = R4 = R5 = H

Figure B-1-6

Comments have already been received on this subject (see section E).

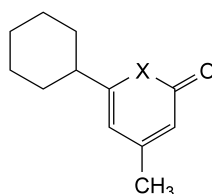
## 2. Divalent groups

A divalent group is an atom or group of atoms able to form 2 single bonds or 1 double bond with the generic structure.

### Case No. 4 – 1 divalent group

When a single divalent group is identified, it is designated X. X may form:

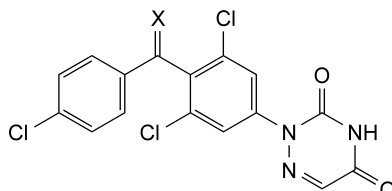
– 2 single bonds to 2 different atoms:



B : X = O  
 C : X = NH

Figure B-1-7

– 1 double bond to a single atom:



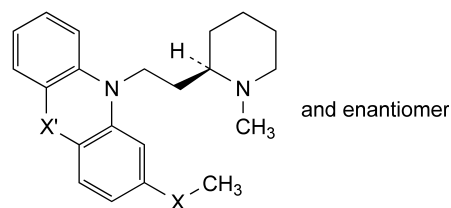
D : X = O  
 F : X = H<sub>2</sub>

Figure B-1-8

X = H<sub>2</sub> is a special case which enables a ketone or an imine to be grouped with its homologous alkane, as shown above.

### Case No. 5 – 2 divalent groups

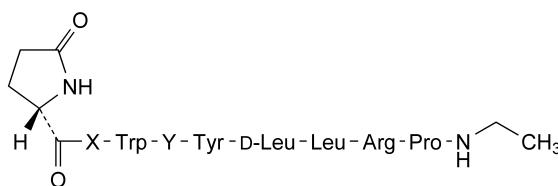
In chemical compounds, when 2 divalent groups are identified, they are designated X and X'.



- 7  
8  
9  
10  
11
- A : X = X' = SO<sub>2</sub>  
 B : X = SO, X' = S  
 C : X = S, X' = SO  
 D : X = SO, X' = SO  
 E : X = SO<sub>2</sub>, X' = S

Figure B-1-9

12  
13 In peptides, when 2 divalent groups are identified, they are designated X and Y.

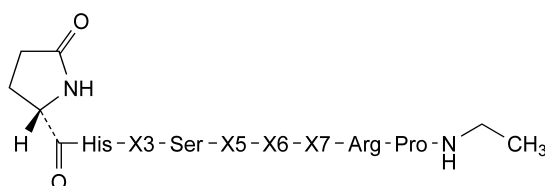


- 20  
21  
22
- A. X = L-His, Y = D-Ser  
 B. X = D-His, Y = L-Ser  
 F. X = D-His, Y = D-Ser

Figure B-1-10

23  
24  
25 *Case No. 6 – More than 2 divalent groups*

26 When more than 2 divalent groups are identified, they are designated X1, X2, X3... Xn.  
 27 This is very rare for chemical products but is very frequent for peptides.  
 28



- 35  
36  
37  
38
- C. X3 = L-Trp, X5 = L-Tyr, X6 = X7 = L-Leu  
 E. X3 = D-Trp, X5 = L-Tyr, X6 = D-Leu, X7 = L-Leu  
 G. X3 = L-Trp, X5 = D-Tyr, X6 = D-Leu, X7 = L-Leu  
 H. X3 = L-Trp, X5 = L-Tyr, X6 = X7 = D-Leu

Figure B-1-11

39  
40 For peptides/proteins, the number assigned to the substituent group corresponds to the  
 41 number assigned to the amino acid that it replaces in the current structure. This makes it  
 42 easier to check the name.

### 43 3. Repetition units

44  
45 As is the case for polymers, repeating units can be shown in square brackets, where *n* is  
 46 the number of times the unit is repeated.

47 *Case No. 7 – 1 repetition unit*

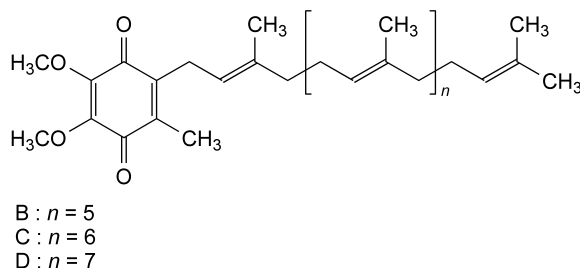


Figure B-1-12

It is not absolutely necessary to do this and a substituent group R that is itself defined as a polymer can also be introduced.

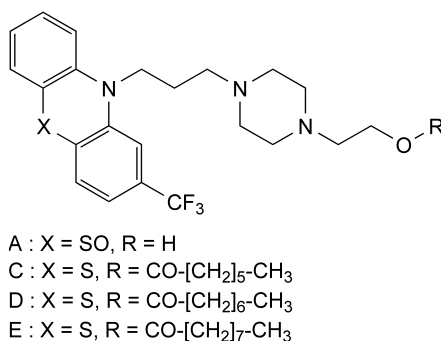


Figure B-1-13

### Case No. 8 – 2 repetition units

When 2 repetition units are defined, they are designated  $n$  and  $m$ .

### Case No. 9 – More than 2 repetition units

When more than 2 repetition units are defined, they are designated  $x$ ,  $y$ ,  $z$ ,  $n$ ,  $m$ .

### Summary of the various substituent groups

See Table B-1-1.

### III. SPECIAL RECOMMENDATIONS FOR PEPTIDES / PROTEINS

For peptides, when several substituent groups have to be described for an impurity and some are from the L series and others are from the D series, the relevant series should be specified clearly.

Table B-1-1. – Representations of substituent groups

Representation	Type of group	Area of application
R	monovalent	1 monovalent group defined
R and R' or R1 and R2	monovalent	2 monovalent groups defined
R1, R2, R3... Rn	monovalent	more than 2 monovalent groups defined
X	divalent	1 divalent group defined
X and X' or X and Y	divalent	2 divalent groups defined
X1, X2, X3... Xn	divalent	more than 2 divalent groups defined
$n$	repetition	1 repetition group defined
$n$ and $m$	repetition	2 repetition groups defined
$x$ , $y$ , $z$ , $n$ , $m$	repetition	more than 2 repetition groups defined

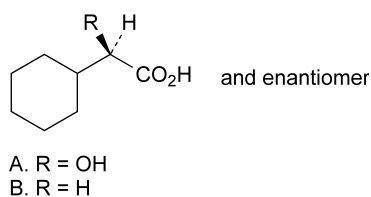
1 When an amino acid is replaced, there are 2 equivalent nomenclature systems that can be  
 2 used. For example, if the 2<sup>nd</sup> amino acid of leuprorelin is replaced by D-leucine, the impurity  
 3 is called [D-Leu<sup>2</sup>]leuprorelin or [2-D-leucine]leuprorelin, with the first name being preferred.

#### 4 IV. SPECIAL RECOMMENDATIONS FOR TERTIARY AND QUATERNARY AMMONIUM 5 SALTS

6 A tertiary ammonium salt and its quaternary homologue are never grouped by means of a  
 7 substituent group R linked to the nitrogen atom (which would take the value R = H for the  
 8 tertiary amine). It is not the amine salts that are described but rather the corresponding  
 9 bases.

#### 10 V. GROUPING AND ISOMERISM

11 A chiral impurity and a non-chiral impurity can be grouped by defining substituent groups  
 12 attached to the centre of asymmetry.



20 Figure B-1-14

21 Comments are often made on this subject (see section E).

22 In the case of a grouped structure defined as a mixture of (*E*) and (*Z*) isomers, see  
 23 impurities B and C in the monograph on *Halothane (0393)*.

#### 24 B-2. MIXTURES OF COMPOUNDS

##### 25 I. GENERAL PRINCIPLES

26 Some parent substances or impurities are defined as mixtures of 2 or more compounds  
 27 with different structures or relatively similar structures. For more than 50 years, WHO  
 28 policy on INNs [9] has been not to assign an INN to mixtures of substances except in the  
 29 case of natural products (in particular, antibiotics produced by fermentation) for which  
 30 the substances meet the 3 following criteria:  
 31 the substances meet the 3 following criteria:

- 32 – they have very similar structures;
- 33 – they have comparable activities;
- 34 – they cannot be separated during their isolation.

35 Other substances for pharmaceutical use are described as mixtures; this is the case for  
 36 most excipients (polymers in particular). Sometimes these have been assigned an INN but  
 37 usually it is not used, with the trivial name or chemical name being favoured [9]. Finally,  
 38 there are rare cases of associations between several compounds in fixed proportions that  
 39 have been assigned an INN [9].

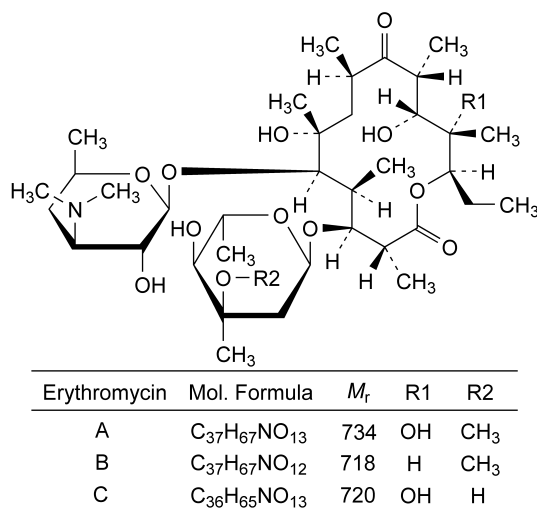
40 There are various editorial rules according to the case:

- 41 – mixtures of structures that can be grouped using radicals;
- 42 – mixtures of structures that cannot be grouped using radicals;
- 43 – mixtures of 2 structural groups;
- 44 – mixtures of structures that are not shown in the Ph. Eur.

1 *II. MIXTURES OF STRUCTURES THAT CAN BE GROUPED*

2 When the parent substance is a mixture of grouped structures, the generic structure is  
 3 shown along with a table that provides information on the various compounds making up  
 4 the mixture. These compounds are usually classified in order of decreasing content in the  
 5 mixture; the main compound is therefore mentioned first.

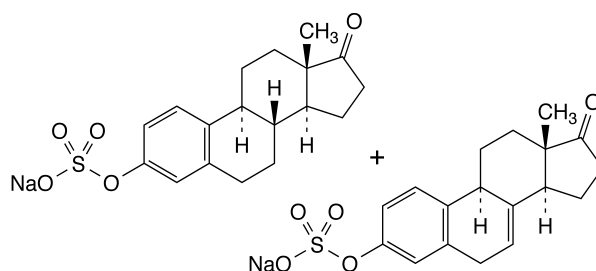
6  
 7 Macrolide antibiotics that are produced by isolation from fermentation media of  
 8 *Streptomyces* strains are usually described as mixtures of compounds with structures  
 9 that can be grouped using radicals.



25 Figure B-2.1. – *Erythromycin (0179)*

26  
 27 *III. MIXTURES OF STRUCTURES THAT CANNOT BE GROUPED*

28 In this case, the 2 structures are shown with a + sign between them. When the molecular  
 29 ratio between the 2 entities is not defined, see the content of each compound given under  
 30 Definition.



41 Figure B-2.2. – *Conjugated estrogens (1512)*

42 *VI. MIXTURES OF STRUCTURES THAT ARE NOT SHOWN IN THE PH. EUR.*

43 In certain cases, a monograph may describe a mixture of structures (in particular for  
 44 heparins and for oils). In this case, a decision is taken with the administrator responsible  
 45 for the monograph not to show the structures since information that is relevant to the user  
 46 cannot be presented because of the complexity of the mixture. The mixture is described  
 47 under Definition as a mixture of compound A and compound B, and the approximate  
 proportions are indicated.

## 1 B-3. ORGANIC AND INORGANIC SALTS

## 2 I. GENERAL PROVISIONS

## 3 1. Salts of a parent substance

4 *Representation of salts*

5 Both organic and inorganic salts are represented in the Ph. Eur. with a comma separating  
6 the parent structure from the acid or base with which it forms the salt.

7 *Molecular ratio*

8 The molecular ratio is usually shown just before the acid or base used to form the salt. If  
9 no molecular ratio is shown, the ratio is 1:1 by default. If this ratio is unknown or variable,  
10 an 'x' is indicated before the molecular formula of the acid or base used to form the salt;  
11 this is done especially for products of fermentation and peptides/proteins.

## 12 2. Salts of an impurity

13 Impurities are not shown in the Ph. Eur. as salts, even if the subject of the monograph is a  
14 parent substance in the form of a salt. Primary, secondary and tertiary ammonium salts  
15 are represented as the corresponding uncharged base. For quaternary ammonium salts,  
16 the cation is shown but not the counter-ion.

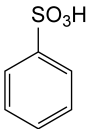
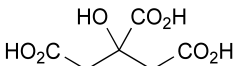
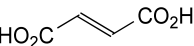
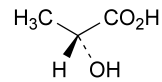
## 17 II. INORGANIC SALTS

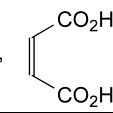
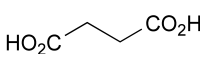
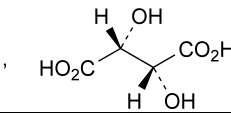
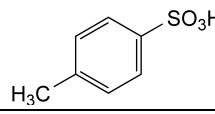
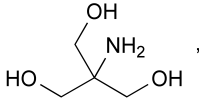
18 The salts of inorganic acids are shown by placing the molecular formula of the acid after a  
19 comma. Various types of inorganic acid salts are described in the Ph. Eur. (see Table B-3-1).

## 20 III. ORGANIC SALTS

21 For reasons of consistency in the Ph. Eur., the salts of organic acids or bases are always  
22 drawn in the same manner. However, rotation of the organic acid or base in question is  
23 tolerated to save space.

24 Table B-3-1. – Principal salts of the Ph. Eur.

Salt	Representation
Acetate	, H <sub>3</sub> C – CO <sub>2</sub> H
Besilate	, 
Hydrobromide	, HBr
Hydrochloride	, HCl
Citrate	, 
Fumarate	, 
Lactate	,  and enantiomer

Salt	Representation
Maleate	
Mesilate	$\text{H}_3\text{C}-\text{SO}_3\text{H}$
Nitrate	$\text{HNO}_3$
Oxalate	$\text{HO}_2\text{C}-\text{CO}_2\text{H}$
Phosphate	$\text{H}_3\text{PO}_4$
Succinate	
Sulphate	$\text{H}_2\text{SO}_4$
Tartrate	
Tosilate	
Trometamol	

#### IV. ORDER OF PRESENTATION OF THE COMPONENTS OF THE SALT

For organic salts, the general principle for representing ions is followed: the structures are presented in the same order as when the salt has been formed so that the cation is placed first followed by the anion, which give the 2 following cases. When an organic acid is used to form the salt, it is placed last:

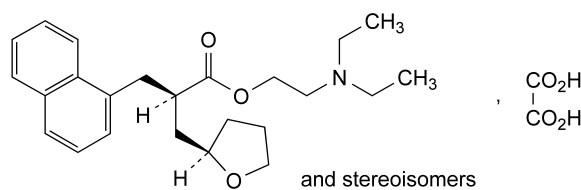


Figure B-3-7. – *Naftidrofuryl hydrogen oxalate (1594)*

When an organic base is used to form the salt, it is placed first:

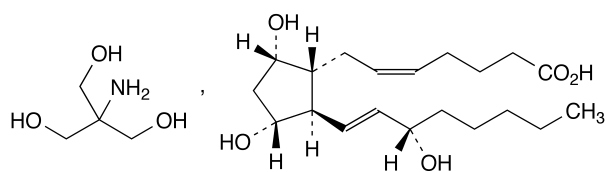


Figure B-3-8. – *Dinoprost trometamol (1312)*

---

## B-4. TITLES OF MONOGRAPHS

### I. INTERNATIONAL NONPROPRIETARY NAMES (INN)

INNs are assigned by the WHO, and when they are recommended they must be used as the titles of monographs. If objections to an INN have been formulated, its use is more delicate and a decision must be taken with the producer on a case-by-case basis.

#### 1. Gender of INNs in French

A rule has been established to determine the gender of INNs in French: names ending in 'one' or 'ine' are feminine and all the others are masculine [7].

#### 2. Salts and esters

When an INN is assigned to a particular salt or ester (for example, levothyroxine sodium), the name of the acid or the base, or that of another other salt or ester, may be chosen as a modified INN (INNM) (for example, levothyroxine) derived from the recommended INN [7].

#### 3. INNs for substituents and groups

INNs have been devised for groups whose names are too long or too awkward to use (for example, mesilate for methanesulphonate) [9]. These INNs make it possible to avoid using stereodescriptors or italics in the titles, thus simplifying the labelling of pharmaceutical substances (for example, derbumine salt is preferable to *tert*-butylamine salt).

#### 4. Substances not covered by INNs

In principle, INNs are not assigned to:

- mixtures of substances (with certain exceptions);
- substances that are not completely characterised (with certain exceptions);
- herbal substances;
- homoeopathic preparations;
- substances that have been in medical use for a long time under a well established name (for example, alkaloids) or under a trivial chemical name.

Exceptions to the above are the following:

- mixtures isolated from biological sources (mainly antibiotics);
- products obtained via a chemical reaction yielding a mixture of homologous compounds (in particular synthetic polymers).

### II. MODIFIED INNS

WHO guidance now exists on the elaboration of modified INNs (namely salts and esters, which are of particular interest to the Ph. Eur.) from recommended INNs [21].

### III. EXCIPIENTS

The derivation of Ph. Eur. monograph titles for excipients is much less homogeneous than for active substances since it may be based on the chemical name, the trivial name or (less frequently) the INN, where it exists. Usage is the main selection criterion [7].

## SECTION C - Main structural classes

The Ph. Eur. has monographs on many substances belonging to the main structural classes. Structures of substances belonging to the same class are drawn in a manner that shows their structural analogy. This makes it easier to compare visually the structures of

1 substances in the same class and also to identify the common parts as well as the specific  
2 parts of the structures. In addition, this approach has the editorial advantage of ensuring  
3 that structures are depicted uniformly and consistently in all the monographs.

4 Despite the existence of general rules described in section A of this guide, it may be  
5 difficult to apply them to certain structural classes, in particular those for which there are  
6 special rules for systematic numbering, for stereochemistry or for nomenclature.

7 Some of the main structural classes for substances in the Ph. Eur. are listed below, for  
8 information:

- 9  
10 – amino acids;  
11 – monosaccharides and their derivatives;  
12 – derivatives of purine and pyrimidine bases;  
13 – steroids;  
14 – terpenoids;  
15 – prostanoids;  
16 – alkaloids;  
17 – antibiotics;  
18 – peptides and proteins;  
19 – polymers.

## 24 SECTION D - Nomenclature and application of IUPAC rules

### 26 I - IUPAC RULES APPLIED

27  
28 The chemical names used in the Ph. Eur. are based on the rules of the International Union  
29 of Pure and Applied Chemistry (IUPAC) published in 'A Guide to IUPAC Nomenclature  
30 of Organic Compounds (Recommendations, 1993)' [14] and on the nomenclature rules  
31 for carbohydrates elaborated by a IUPAC and IUBMB joint commission 'Nomenclature of  
32 carbohydrates' [15].

### 34 II - CHANGES TO IUPAC RULES

35 Proposed changes to IUPAC rules on specific topics have been published in *Pure &*  
36 *Appl. Chem.* These can be downloaded free of charge from the official IUPAC website [19].  
37 Caution should be exercised when consulting websites that display these proposed  
38 changes mixed imprudently with the formally adopted rules (especially site [20]). The  
39 Ph. Eur. does not follow these recommendations adopted after 1993.

### 41 III - PREFERRED IUPAC NAME

42 For a given compound, there often are several names that are acceptable by IUPAC rules.  
43 A guide to selecting a 'preferred name' is currently being prepared by IUPAC and it is  
44 expected to be available for enquiry in 2006. Nevertheless, the Ph. Eur. will not follow  
45 these recommendations until the document is formally adopted. For now, there are no  
46 plans to use this document for a systematic revision of the names in already published  
47 monographs.

1 IV - NAME SOFTWARE BY ACD/Labs

2 The Ph. Eur. has been using Name software produced by ACD/Labs (Advanced Chemistry  
3 Development [17]) to elaborate the names in its texts since Pharmeuropa 18.1. This  
4 software uses the same IUPAC nomenclature rules that are used by the Ph. Eur. [14,15].  
5 The software can also be set up to operate using IUPAC 1979 rules [13]. However, IUPAC  
6 recommendations published after 1993 are not applied in the software.  
7

8  
9 **SECTION E - Frequently Asked Questions (FAQ)**

10 Questions on structures and names are often received by the Ph. Eur. after monographs  
11 are published for enquiry in Pharmeuropa. The most frequently asked questions are  
12 shown below. For more detailed answers see the corresponding section of this guide.  
13

Your question/comment	Answer in paragraph
When are S/N corrections introduced into monographs?	I
How to report an error in a structure	II
How to report an error in a name	III
How to report a discrepancy between a structure and its name	IV
Why does the name indicate <i>r/s</i> instead of <i>R/S</i> ?	V
What does $R + R' = O$ mean?	VI
Where is the R3 radical?	VII
Why is a nonchiral impurity grouped with a chiral impurity?	VII
The bond angle is wrong	II-1
The bond length is wrong	II-1
The counter-ion of an impurity is missing	II-1
The salt of an impurity is missing	II-1
The <i>N</i> -oxide is not represented properly	II-1
You wish to propose an equivalent IUPAC name	III-1
'Sulphate' is misspelt in English	III-1

30 I - WHEN ARE S/N CORRECTIONS INTRODUCED INTO MONOGRAPHS?

31  
32 The structure and nomenclature expert sends his corrections between the publication of  
33 a monograph in Pharmeuropa (ANP) and its adoption by the European Pharmacopoeia  
34 Commission (COM). Usually these corrections have already been introduced into the draft  
35 submitted to the Commission for adoption. Delays nevertheless may occur and in this case  
36 it is not necessary to re-send a comment that had already been made at the Pharmeuropa  
37 stage. The corrections will be introduced before the text is published in a supplement.

38 II - HOW TO REPORT AN ERROR IN A STRUCTURE

39 *1 - EDITORIAL COMMENTS*

40  
41 It is not necessary to point out editorial errors in the representation of a structure such as  
42 bond angles, bond lengths, the absence of the counter-ion of an impurity, the absence of  
43 the salt of an impurity or the manner of representing *N*-oxides.

44 *2 - COMMENT ON THE CONTENT*

45 All errors in content concerning the structure must be reported. Requests for revision  
46 or comments will be processed much more easily if they are supported by licensing data  
47 or structural determination data.

## III - HOW TO REPORT AN ERROR IN A NAME

## 1 - EDITORIAL COMMENTS

It is not necessary to submit editorial comments on the name of a structure such as proposing equivalent IUPAC names, correcting the spelling of sulphate derivatives in English, etc.

## 2 - COMMENT ON THE CONTENT

Any non-editorial error concerning nomenclature must be reported. Corrections are easier to implement if the requests for revision or the comments follow IUPAC numbering rules for atoms (trivial numbering should not be used).

As the Ph. Eur. does not follow IUPAC recommendations published after 1993, it is not necessary to point out that they have not been applied.

In general, any comment on nomenclature must mention the IUPAC rule that it is based on, as well as the document in which the rule was published.

## IV - HOW TO REPORT A DISCREPANCY BETWEEN A STRUCTURE AND ITS NAME

## 1 - SUPPLEMENT

Any relevant discrepancies between a structure and its name must be pointed out for a supplement.

## 2 - PHARMEUROPA

Discrepancies between a structure and its name may be pointed out at the Pharmeuropa stage. It should be noted that since the expert's corrections are introduced only after the Pharmeuropa stage, it is likely that the published text will include errors.

V - WHY DOES THE NAME INDICATE *r/s* INSTEAD OF *R/S*?

The stereodescriptors *r* and *s* (lower-case italics) are used instead of *R* and *S* to designate the stereochemistry of a carbon atom that is stereogenic because it has 4 different substituents (2 of which differ only in terms of stereochemistry), but which is part of a structure that is non-chiral due to a plane of symmetry passing through this carbon atom. This case applies in particular to derivatives of atropine.

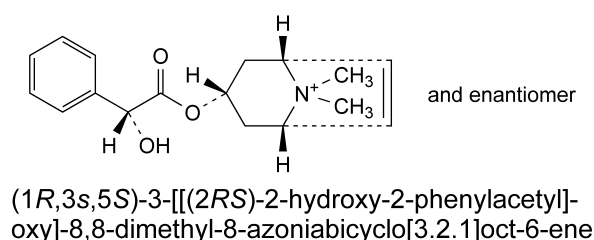
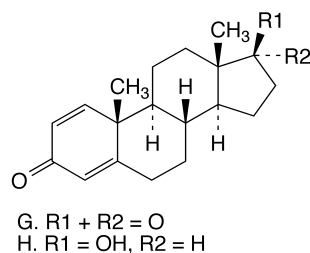


Figure FAQ-5. – Impurity A of *Homatropine methylbromide* (0720)

VI - WHAT DOES  $R + R' = O$  MEAN?

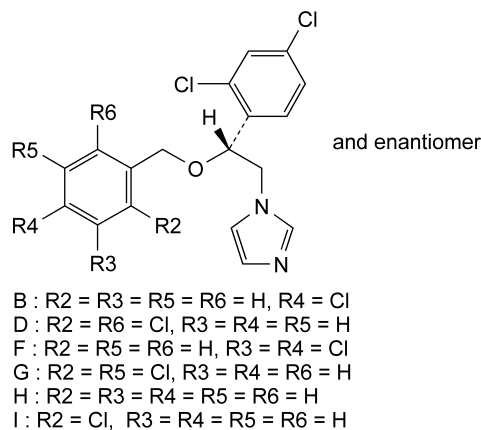
$R + R' = O$  or  $R_1 + R_2 = O$  makes it possible to show a ketone or an imine with its homologous alkane/amine/alcohol. Obviously the bond angle is not accurately shown in this case but such a grouped presentation saves space and also shows the structural analogy between the impurities to make the transparency list easier to understand.



8  
9  
Figure FAQ-6. – Impurities G to H of *Testosterone* (1373)

10 VII - WHERE IS THE R3 SUBSTITUENT GROUP?

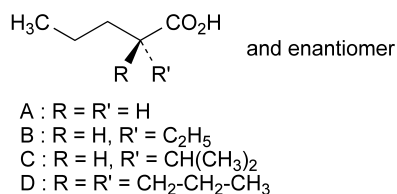
11 If the substituent groups are attached to an aromatic ring, they are numbered in  
12 accordance with systematic numbering. It should be noted that in this context R1 is  
13 almost always absent (the main functional group is in position 1). Other intermediary  
14 substituent groups may also be missing (R3 may for example be absent without modifying  
15 the numbering for R2, R4, R5 and R6).



28  
29  
Figure FAQ-7. – Impurities B, D, F, G, H and I of *Miconazole nitrate* (0513)

30 VIII - WHY IS A NONCHIRAL IMPURITY GROUPED WITH A CHIRAL IMPURITY?

31 In the Ph. Eur. a chiral impurity and a nonchiral impurity may be grouped by defining  
32 radicals attached to the centre of asymmetry. The graphic representation with the legend  
33 'and enantiomer' is not really correct for the nonchiral impurity but the main objective of  
34 this type of grouping is to save space and also to show the structural analogy between the  
35 impurities to make the transparency list easier to understand.



42  
43  
44  
Figure FAQ-8. – Impurities A to D of *Valproic acid* (1378)

45  
46  
47  
REFERENCES

[1] The graphic representation of chemical formulae in the publications of international nonproprietary names (INN) for pharmaceutical substances. WHO/PHARM/95.579.

- 1 [2] International Union of Pure and Applied Chemistry, Organic Chemistry Division,  
2 Commission on the Nomenclature of Organic Chemistry, Nomenclature of Organic  
3 Chemistry, sections A, B, C, D, E, F and H, 4<sup>th</sup> ed. Oxford, Pergamon, 1979.
- 4 [3] Leigh GJ, ed. Nomenclature of Inorganic Chemistry: Recommendations 1990. Oxford,  
5 Blackwell Scientific, 1990.
- 6 [4] Compendium of Chemical Terminology, IUPAC Recommendations, 1999.
- 7 [5] World Health Organisation. Guidelines on the use of International Nonproprietary  
8 Names (INNs) for Pharmaceutical Substances. WHO/PHARM S/NOM 1570 (1997).
- 9 [6] World Health Organisation. International Nonproprietary Names (INN). *WHO Drug*  
10 *Information* 2002;**16**(4):293-307.
- 11 [7] World Health Organisation. International Nonproprietary Names (INN) for  
12 pharmaceutical substances, Cumulative List No. 10 (2002).
- 13 [8] WHO Expert Committee on Specifications for Pharmaceutical Preparations. WHO,  
14 Technical Report Series No. 863.
- 15 [9] International Nonproprietary Names (INN) for pharmaceutical use – Comprehensive  
16 list of names for radicals and groups. WHO/EDM/QSM/2003.1.
- 17 [10] The use of common stems in the selection of International Nonproprietary  
18 Names (INN) for pharmaceutical use. WHO/EDM/QSM/2003.2.
- 19 [11] INNs for stereoisomers. World Health Organisation. PHARM S/NOM 1554, Rev2.
- 20 [12] WHO Mednet database, access is free of charge and requires membership:  
21 <http://mednet.who.int>
- 22 [13] Nomenclature of Organic Chemistry, Sections A, B, C, D, E, F, and H, Pergamon  
23 Press, Oxford, 1979. Copyright 1979 IUPAC.
- 24 [14] A Guide to IUPAC Nomenclature of Organic Compounds (Recommendations 1993),  
25 1993, Blackwell Scientific publications, Copyright 1993 IUPAC.
- 26 [15] Nomenclature of carbohydrates (Recommendations 1996). *Pure & Appl. Chem.*,  
27 Vol. 68, No. 10, pp. 1919-2008, 1996.
- 28 [16] CambridgeSoft website: [www.cambridgesoft.com](http://www.cambridgesoft.com).
- 29 [17] ACDLabs website: [www.acdlabs.com](http://www.acdlabs.com).
- 30 [18] IUPAC website: [www.iupac.org](http://www.iupac.org).
- 31 [19] *Pure & Appl. Chem* free access through internet: [www.iupac.org/publications/pac/index.html](http://www.iupac.org/publications/pac/index.html).
- 32 [20] Dr. Moss's IUPAC website: [www.chem.qmul.ac.uk/iupac](http://www.chem.qmul.ac.uk/iupac).
- 33 [21] World Health Organisation. International Nonproprietary Names Modified. INN  
34 Working Document 05.167/3; 2006.
- 35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47