

IR-4 Formulae (Draft March 2004)

CONTENTS

- IR-4.1 Introduction
- IR-4.2 Definitions of types of formulae
 - IR-4.2.1 Empirical formulae
 - IR-4.2.2 Molecular formulae
 - IR-4.2.3 Structural formulae and the use of enclosing marks in formulae
 - IR-4.2.4 Formulae of addition compounds
 - IR-4.2.5 Solid state structural information
- IR-4.3 Indication of ionic charge
- IR-4.4 Sequence of citation of symbols in formulae
 - IR-4.4.1 Introduction
 - IR-4.4.2 Ordering principles
 - IR-4.4.2.1 Electronegativity
 - IR-4.4.2.2 Alphabetical order
 - IR-4.4.3 Formulae for specific classes of compounds
 - IR-4.4.3.1 Binary compounds
 - IR-4.4.3.2 (Formal) treatment as coordination compounds
 - IR-4.4.3.3 Chain compounds
 - IR-4.4.3.4 Generalised salt formulae
 - IR-4.4.3.5 Addition compounds, multiple salts, and solvates
 - IR-4.4.4 Use of ligand abbreviations
- IR-4.5 Isotopically modified compounds
 - IR-4.5.1 General formalism
 - IR-4.5.2 Isotopically substituted compounds
 - IR-4.5.3 Isotopically labelled compounds
 - IR-4.5.3.1 Types of labelling
 - IR-4.5.3.2 Specifically labelled compounds
 - IR-4.5.3.3 Selectively labelled compounds
- IR-4.6 Optional modifiers of formulae
 - IR-4.6.1 Indication of oxidation state
 - IR-4.6.2 Formulae of radicals
 - IR-4.6.3 Formulae of optically active compounds
 - IR-4.6.4 Indication of excited states
 - IR-4.6.5 Structural descriptors
- IR-4.7 References

IR-4.1 INTRODUCTION

Formulae (empirical, molecular, and structural formulae as described below) provide a simple and clear method of designating compounds. They are of particular importance in chemical equations and in descriptions of chemical procedures. In order to avoid ambiguity and for many other purposes, *e.g.* in databases, indexing, *etc.*, standardisation is recommended.

IR-4.2 DEFINITIONS OF TYPES OF FORMULA

IR-4.2.1 Empirical formulae

The empirical formula of a compound is formed by juxtaposition of the atomic symbols with appropriate (integer) subscripts to give the simplest possible formula expressing the composition. For the order of citation of symbols in formulae, see Section IR-4.4, but, *in the absence of any other ordering criterion* (for example, if little structural information is available), the alphabetical order of atomic symbols should be used in an empirical formula, except that in carbon-containing compounds, C and H are usually cited first and second, respectively.¹

Examples:

1. BrClH₃N₂NaO₂Pt
2. C₁₀H₁₀ClFe

IR-4.2.2 Molecular formulae

For compounds consisting of discrete molecules, the *molecular formula*, as opposed to the empirical formula, may be used to indicate the actual composition of the molecules. For the order of citation of symbols in molecular formulae, see Section IR-4.4.

The choice of formula depends on the chemical context. In some cases, the empirical formula may also correspond to a molecular composition, in which case the only possible difference between the two formulae is the ordering of the atomic symbols. If it is not desirable or possible to specify the composition, *e.g.* in the case of polymers, a letter subscript such as *n* may be used.

Examples:

	<i>Molecular formula</i>	<i>Empirical formula</i>
1.	S ₈	S
2.	S _{<i>n</i>}	S
3.	SF ₆	F ₆ S
3.	S ₂ Cl ₂	ClS

4.	$\text{H}_4\text{P}_2\text{O}_6$	$\text{H}_2\text{O}_3\text{P}$
5.	Hg_2Cl_2	ClHg
6.	N_2O_4	NO_2

IR-4.2.3 Structural formulae and the use of enclosing marks in formulae

A structural formula gives partial or complete information about the way in which the atoms in a molecule are connected and arranged in space. In simple cases, a line formula that is just a sequence of atomic symbols gives structural information provided the reader knows that the formula represents the order of the atoms in the linear structure.

Examples:

1.	HOCN	(empirical formula CHNO)
2.	HNCO	(empirical formula also CHNO)
3.	HOOH	(empirical formula HO)

As soon as the compound has even a slightly more complex structure, it becomes necessary to use enclosing marks in line formulae to separate subgroups of atoms. Different enclosing marks must be used for repeating units and sidechains in order to avoid ambiguity.

The basic rules for applying enclosing marks in structural formula are as follows:

- (i) Repeating units in chain compounds are enclosed in square brackets.
- (ii) Side groups to a main chain and groups (ligands) attached to a central atom are enclosed in parentheses (except single atoms when there is no ambiguity regarding their attachment in the structure, *e.g.* hydrogen in hydrides with a chain structure).
- (iii) A formula or part of a formula which represents a molecular entity may be placed in enclosing marks. If an entire formula is enclosed, square brackets must be used, except if rule (v) applies.
- (iv) A part of a formula which is to be multiplied by a subscript may also be enclosed in parentheses or braces, except in the case of repeating units in chain compounds, *cf.* rule (i).
- (v) In the case of polymers, if the bonds between repeating units are to be shown, the repeating unit is enclosed in strike-through parentheses, with the dash superimposed on the parentheses representing the bond. (If this is typographically inconvenient, dashes can be placed before and after the parentheses).

(vi) Inside square brackets, enclosing marks are nested as follows:

$()$, $\{()\}$, $\{\{()\}\}$, $\{\{\{()\}\}\}$, *etc.*

(vii) Atoms or groups of atoms which are represented together with a prefixed symbol, *e.g.* a structural modifier such as ' μ ', are placed within enclosing marks, using the nesting order $()$, $\{()\}$, $\{\{()\}\}$, $\{\{\{()\}\}\}$, *etc.*

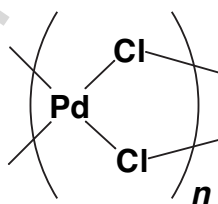
The use of enclosing marks for the specification of isotopic substitution is described in Section IR-4.5.

Compared to line formulae, displayed formulae (Example 13 below) give more (or full) information about the structure.

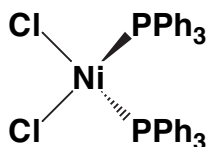
(The rules needed for ordering the symbols in some of the example formulae below are given in Section IR-4.4.3.)

Examples:

- | | | |
|-----|--|----------------------------|
| 4. | $\text{SiH}_3[\text{SiH}_2]_8\text{SiH}_3$ | [rule (i)] |
| 5. | $\text{SiH}_3[\text{SiH}_2]_6\text{SiH}(\text{SiH}_3)\text{SiH}_3$ | [rules (i) and (ii)] |
| 6. | $\text{Ca}_3(\text{PO}_4)_2$ | [rule (iv)] |
| 7. | $[\text{Co}(\text{NH}_3)_6]_2(\text{SO}_4)_3$ | [rules (iii), (iv), (vi)] |
| 8. | $[\{\text{Rh}(\mu\text{-Cl})(\text{CO})_2\}_2]$ | [rules (iii), (vi), (vii)] |
| 9. | $\text{K}[\text{Os}(\text{N})_3]$ | [rules (ii), (iii)] |
| 10. | $\{\text{S}\}_n$ | [rule (v)] |
| 11. | $(\text{HBO}_2)_n$, or $\{\text{B}(\text{OH})\text{O}\}_n$ | [rules (ii) and (v)] |
| 12. | $([\text{PdCl}_2])_n$, or | |



13.



- | | |
|-----|-----------------|
| 14. | NaCl |
| 15. | $[\text{NaCl}]$ |

The first formula in each of Examples 11 and 12 may be considered to be molecular formulae (Section IR-4.2.2) with no implications about the structure of the polymers in question.

In Examples 14 and 15, the formula [NaCl] may be used to distinguish the molecular compound consisting of one sodium atom and one chlorine atom from the solid with the composition NaCl.

IR-4.2.4 **Formulae of addition compounds**

In the formulae of addition compounds, including multiple salts and solvates (particularly hydrates), a special format is used. The proportions of constituents are indicated by arabic numerals preceding the formulae of the constituents, and the formulae of the constituents are separated by a centre dot. The rules for ordering the constituent formulae are described in Section IR-4.4.3.5.

Examples:

1. $\text{Na}_2\text{CO}_3 \cdot 10\text{H}_2\text{O}$
2. $8\text{H}_2\text{S} \cdot 46\text{H}_2\text{O}$
3. $\text{BMe}_3 \cdot \text{NH}_3$

IR-4.2.5 **Solid state structural information**

Structural information can also be given by indicating structural type as a qualification of a molecular formula (see Section IR-11.7.2). For example, polymorphs may be indicated by adding in parentheses an abbreviated expression for the crystal system. Structures may also be designated by adding the name of a type-compound in italics in parentheses, but such usage may not be unambiguous. There are at least ten varieties of $\text{ZnS}(h)$. Where several polymorphs crystallise in the same crystal system they may be differentiated by the Pearson symbol (see Sections IR-3.5.3 and IR-11.5.2). Greek letters are frequently employed to designate polymorphs, but their use is often confused and contradictory and is not recommended.

Examples:

1. $\text{TiO}_2(t)$ (*anatase* type)
2. $\text{TiO}_2(r)$ (*brookite* type)
3. $\text{AuCd}(c)$, or AuCd (*CsCl* type)

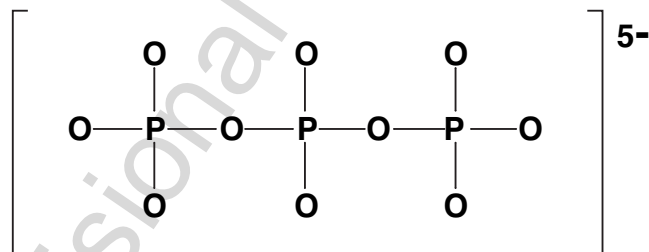
For the formulae of solid solutions and non-stoichiometric phases, see Chapter IR-11.

IR-4.3 INDICATION OF IONIC CHARGE

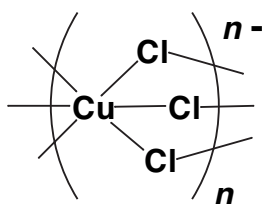
Ionic charge is indicated by means of a right upper index, as in A^{n+} or A^{n-} (not A^{+n} or A^{-n}). If the formula is placed in enclosing marks, the right upper index is placed outside the enclosing marks. For polymeric ions, the charge of a single repeating unit should be placed inside the parentheses that comprise the polymeric structure or the total charge of the polymeric species should be placed outside the polymer parentheses. (The rules needed for ordering the symbols in some of the example formulae below are given in Section IR-4.4.3.)

Examples:

- | | |
|--|--|
| 1. Cu^+ | 7. As^{3-} |
| 2. Cu^{2+} | 8. HF_2^- |
| 3. NO^+ | 9. CN^- |
| 4. $[\text{Al}(\text{H}_2\text{O})_6]^{3+}$ | 10. $\text{S}_2\text{O}_7^{2-}$ |
| 5. H_2NO_3^+ | 11. $[\text{Fe}(\text{CN})_6]^{4-}$ |
| 6. $[\text{PCl}_4]^+$ | 12. $[\text{PW}_{12}\text{O}_{40}]^{3-}$ |
| 13. $[\text{P}_3\text{O}_{10}]^{5-}$, or $[\text{O}_3\text{POP}(\text{O})_2\text{OPO}_3]^{5-}$, or | |



14. $([\text{CuCl}_3]^-)_n$, or



IR-4.4 SEQUENCE OF CITATION OF SYMBOLS IN FORMULAE

IR-4.4.1 Introduction

Atomic symbols in formulae may be ordered in various ways. Section IR-4.4.3 describes the conventions usually adopted for some important classes of compounds. As a prerequisite,

Section IR-4.4.2 explains what is meant by the two ordering principles 'electronegativity' and 'alphabetical ordering'.

IR-4.4.2 **Ordering principles**

IR-4.4.2.1 *Electronegativity*

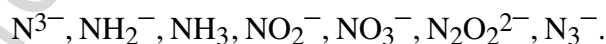
If electronegativity is taken as the ordering principle in a formula or a part of a formula, the atomic symbols are cited according to *relative* electronegativities, the least electronegative element being cited first. For this purpose, Table VI* is used as a guide, except that oxygen is placed between chlorine and fluorine.

IR-4.4.2.2 *Alphabetical order*

A single letter symbol always precedes a two-letter symbol with the same initial letter, *e.g.* B before Be, O before OH. The group NH_4 is treated as a single symbol and so is listed after Na, for example.

Where the entities to be arranged in a formula are polyatomic, the order of citation is decided by selecting a particular atomic symbol to characterise the entity. The first symbol in the formula of a polyatomic group, as written according to the appropriate rule in Section IR-4.4.3, determines the alphabetical order. For example, C_5H_5 , SCN, UO_2 , NO_3 , OH, and $[\text{Zn}(\text{H}_2\text{O})_6]^{2+}$ are ordered under C, S, U, N, O, and Zn, respectively. If the first symbols are the same, the symbol with the lesser right index is cited first, *e.g.* NO_2 precedes N_2O_2 . If this still does not discriminate, the subsequent symbols are used alphabetically and numerically to define the order, *e.g.* NH_2 precedes NO_2 which precedes NO_3 .

To summarise and exemplify, the order of citation of some nitrogen-containing compounds is:



IR-4.4.3. **Formulae for specific classes of compounds**

IR-4.4.3.1 *Binary compounds*

In accordance with established practice, the electronegativity criterion (Section IR-4.4.2.1) is generally used in binary compounds.²

Examples:

* Tables numbered with a Roman numeral are collected together at the end of this book.

1. NH₃
2. H₂S
3. OF₂
4. Cl₂O
5. ClO⁻
6. PH₄⁺
7. P₂O₇⁴⁻
8. [SiAs₄]⁸⁻
9. RbBr
10. [Re₂Cl₉]⁻
11. HO⁻ or OH⁻

Note that the formula for the hydroxide ion should be HO⁻ to be consistent with the above convention.

12. Rb₁₅Hg₁₆
13. Cu₅Zn₈ and Cu₅Cd₈

Ordering by electronegativity could, in principle, be applied to ternary, quaternary, *etc.* compounds. For most compounds consisting of more than two elements, however, other criteria for ordering the element symbols in the formula are more often used (see Sections IR-4.4.3.2 to IR-4.4.3.4).

IR-4.4.3.2 (Formal) treatment as coordination compounds

The nomenclature of coordination compounds is described in detail in Chapter IR-9. A brief summary of the construction of *formulae* of coordination compounds is given here. Many polyatomic compounds may conveniently be treated as coordination compounds for the purpose of constructing a formula.

In the formula of a coordination entity, the symbol of the central atom(s) is/are placed first, followed by the symbols or formulae of the ligands, unless additional structural information can be presented by changing the order (see, for example, Section IR-4.4.3.3).

The order of citation of central atoms is based on electronegativity as described in Section IR-4.4.2.1. Ligands are cited alphabetically (Section IR-4.4.2.2) according to the first symbol of the ligand formula or ligand abbreviation (see Section IR-4.4.4) *as written*. Where possible, the ligand formula should be written in such a way that a/the donor atom is closest to the central atom to which it is attached.

Square brackets may be used to enclose the whole coordination entity whether charged or not. Established practice is always to use square brackets for coordination entities with a transition metal as the central atom.

Examples:

1. PBrCl_2
2. SbCl_2F or $[\text{SbCl}_2\text{F}]$
3. $[\text{Mo}_6\text{O}_{18}]^{2-}$
4. $[\text{CuSb}_2]^{5-}$
5. $[\text{UO}_2]^{2+}$
6. $[\text{SiW}_{12}\text{O}_{40}]^{4-}$
7. $[\text{BH}_4]^-$
8. $[\text{ICl}_4]^-$
9. $[\text{PtCl}_2\{\text{P}(\text{OEt})_3\}_2]$
10. $[\text{Al}(\text{OH})(\text{OH}_2)_5]^{2+}$
11. $[\text{PtBrCl}(\text{NH}_3)(\text{NO}_2)]^-$
12. $[\text{PtCl}_2(\text{NH}_3)(\text{py})]$
13. $[\text{Co}(\text{en})\text{F}_2(\text{NH}_3)_2]^+$, but $[\text{CoF}_2(\text{NH}_2\text{CH}_2\text{CH}_2\text{NH}_2)(\text{NH}_3)_2]^+$
14. $[\text{Co}(\text{NH}_3)_5(\text{N}_3)]^{2-}$

In a few cases, a moiety which comprises different atoms and which occurs in a series of compounds is considered as an entity that acts as a central atom and is cited as such, even if this violates the alphabetical order of ligands. For example, PO and UO_2 are regarded as single entities in Examples 15 and 16.

Examples:

15. POBr_3 (alphabetically, PBr_3O)
16. $[\text{UO}_2\text{Cl}_2]$ (alphabetically, $[\text{UCl}_2\text{O}_2]$)

For derivatives of parent hydrides, the alphabetical order of ligands is traditionally disobeyed in that remaining hydrogen atoms are listed first among the ligands in the formula.

Examples:

17. GeH_2F_2
18. SiH_2BrCl
19. $\text{B}_2\text{H}_5\text{Cl}$

For carbaboranes, there has previously been some uncertainty over the order of B and C.³ The order 'B before C' recommended here conforms to both electronegativity and alphabetical order (*i.e.* it is an exception to the Hill order¹ in Section IR-4.2.1). In addition, carbon atoms that replace skeletal boron atoms are cited immediately after boron, regardless of what other elements are present. (See also Section IR-6.2.4.4).

Examples:

20. $B_3C_2H_5$ (recommended)
21. $B_3C_2H_4Br$ (recommended)

For inorganic oxoacids, there is a traditional ordering of formulae in which the 'acid' or 'replaceable' hydrogen atoms (hydrogen atoms bound to oxygen) are listed first, followed by the central atom, then 'non-replaceable' hydrogen atoms (hydrogen atoms bound directly to the central atom), and finally oxygen. This format is an alternative to writing the formulae as coordination compound formulae (see Section IR-8.3).

Examples:

22. HNO_3 (traditional) or $[NO_2(OH)]$ (coordination)
23. H_2PHO_3 (traditional) or $[PHO(OH)_2]$ (coordination)
24. $H_2PO_4^-$ (traditional) or $[PO_2(OH)_2]^-$ (coordination)
25. $H_5P_3O_{10}$ (traditional) or
 $[(HO)_2P(O)OP(O)(OH)OP(O)(OH)_2]$ (coordination)
26. $(HBO_2)_n$ (traditional) or $(B(OH)O)_n$ (coordination)

IR-4.4.3.3 Chain compounds

For chain compounds containing three or more different elements, the sequence of atomic symbols should generally be in accord with the order in which the atoms are bound in the molecule or ion, rather than using alphabetical order or order based on electronegativity. However, if one wishes to view a compound formally as a coordination compound, *e.g.* in connection with a discussion of additive naming of the compound, one may use a coordination-compound type of formula, as in Example 1 below.

Examples:

1. SCN^- (not CNS^-) = $[C(N)S]^-$, nitridosulfidocarbonate(1-)
2. $BrSCN$ (not $BrCNS$)
3. $HOCN$ (cyanic acid)
4. $HONC$ (fulminic acid)
5. $[O_3POSO_3]^{3-}$

IR-4.4.3.4 *Generalised salt formulae*

If the formula of a compound containing three or more elements is not naturally assigned using the preceding two sections, the compound can be treated as a generalised salt. This term is taken to mean any compound in which it is possible to identify at least one constituent which is a positive ion or can be classified as electropositive or more electropositive than the other constituents, and at least one constituent which is a negative ion or can be classified as electronegative or more electronegative than the rest of the constituents. The ordering principle is then:

- (i) all electropositive constituents precede all electronegative constituents;
- (ii) within each of the two groups of constituents, alphabetical order is used.

Examples:

1. KMgF_3
2. $\text{MgCl}(\text{OH})$
3. $\text{FeO}(\text{OH})$
4. $\text{NaTl}(\text{NO}_3)_2$
5. $\text{Li}[\text{H}_2\text{PO}_4]$
6. $\text{NaNH}_4[\text{HPO}_4]$
7. $\text{Na}[\text{HPO}_3]$
8. CuK_5Sb_2 or K_5CuSb_2
9. $\text{K}_5[\text{CuSb}_2]$
10. $\text{H}[\text{AuCl}_4]$
11. $\text{Na}(\text{UO}_2)_3[\text{Zn}(\text{H}_2\text{O})_6](\text{O}_2\text{CMe})_9$

The first formula in Example 8 was arrived at by considering K and Cu to be electropositive constituents and Sb to be electronegative, the second by considering K to be electropositive and Cu and Sb to be electronegative. No structural information is conveyed by these formulae. The formula in Example 9, on the other hand, implies the presence of a molecular entity or coordination entity $[\text{CuSb}_2]^{5-}$.

Deviation from alphabetical order of constituents in the same class is allowed to emphasise similarities between compounds.

Example:

12. CaTiO_3 and ZnTiO_3 (rather than TiZnO_3)

Some generalised salts may also be treated as addition compounds, see Section IR-4.4.3.5.

IR-4.4.3.5 *Addition compounds, multiple salts and solvates*

In the formulae of addition compounds, multiple salts and solvates, the component molecules or entities are cited in order of increasing number; if they occur in equal numbers, they are cited in alphabetical order of the first symbols. In addition compounds containing water, the water remains conventionally cited last. However, boron compounds are no longer treated as exceptions.

Examples:

1. $3\text{CdSO}_4 \cdot 8\text{H}_2\text{O}$
2. $\text{Na}_2\text{CO}_3 \cdot 10\text{H}_2\text{O}$
3. $\text{Al}_2(\text{SO}_4)_3 \cdot \text{K}_2\text{SO}_4 \cdot 24\text{H}_2\text{O}$
4. $\text{AlCl}_3 \cdot 4\text{EtOH}$
5. $8\text{H}_2\text{S} \cdot 46\text{H}_2\text{O}$
6. $\text{C}_6\text{H}_6 \cdot \text{NH}_3 \cdot \text{Ni}(\text{CN})_2$
7. $\text{BF}_3 \cdot 2\text{H}_2\text{O}$
8. $\text{BF}_3 \cdot 2\text{MeOH}$

IR-4.4.4 *Ligand abbreviations*

Since abbreviations are widely used in the chemical literature, agreement on their use and meaning is desirable. This Section provides guidelines for the selection of ligand abbreviations for application in the formulae of coordination compounds (Section IR-9.2.3.4). Some commonly used ligand abbreviations are listed in Table VII with diagrams of most of the ligands shown in Table VIII.

An abbreviation for an organic ligand should be derived from a name consistent with the current rules for the systematic nomenclature of organic compounds.⁴ (For some ligands a non-systematic name is included in Table VII if it was the source of the abbreviation and if that abbreviation is still commonly used). However, new abbreviations should be constructed according to the following recommendations:

- (i) Ligand abbreviations should be formulated so as to avoid confusion and misunderstanding. Since a reader may not be familiar with an abbreviation, it should be explained at the beginning of the text or when it first appears in a publication.
- (ii) New meanings should not be suggested for abbreviations or acronyms that have generally accepted meanings, *e.g.* DNA, NMR, ESR, HPLC, Me (for methyl), Et (for ethyl), *etc.*

- (iii) An abbreviation should readily suggest the ligand name, *e.g.* ida for iminodiacetato.
- (iv) Abbreviations should be as short as practicable, but should contain more than one letter or symbol.
- (v) The use of non-systematic names (and their abbreviations) is discouraged.
- (vi) Abbreviations should normally use only lower-case letters, with several well established exceptions:
 - (a) abbreviations for alkyl, aryl and similar groups, which have the first letter capitalised with the remaining letters in lower case, *e.g.* Me (for methyl), Ac (for acetyl), Cp (for cyclopentadienyl), *etc.*;
 - (b) abbreviations containing atomic symbols, *e.g.* [12]aneS₄;
 - (c) abbreviations containing Roman numerals, *e.g.* H₂ppIX for protoporphyrin IX;
 - (d) abbreviations for ligands containing readily removable hydrons (see vii).

(N.B. Abbreviations for solvents that behave as ligands should also be in lower case letters [*e.g.* dmsO for dimethyl sulfoxide{(methylsulfinyl)methane}, thf for tetrahydrofuran]; the practice of capitalising the abbreviation of a solvent when it does not behave as a ligand is strongly discouraged as an unnecessary distinction).

- (vii) Hydration of anionic ligands, *e.g.* ida, leads to acids which may be abbreviated by the addition of H, *e.g.* Hida, H₂ida.
- (viii) Ligands which are normally neutral, but which continue to behave as ligands on losing one or more hydrons, are abbreviated by adding -1H, -2H, *etc.* as subscripts (including the numeral 1) after the usual abbreviation of the ligand. For example, if Ph₂PCH₂PPh₂ (dppm) loses one hydron to give [Ph₂PCHPPh₂]⁻ its abbreviation is dppm_{-1H}; if it loses two hydrons, its abbreviation is dppm_{-2H}, *etc.*

IR-4.5 ISOTOPICALLY MODIFIED COMPOUNDS.⁵

IR-4.5.1 General formalism

The mass number of any specific nuclide can be indicated in the usual way with a left superscript preceding the appropriate atomic symbol (see Section IR-3.2).

When it is necessary to cite different nuclides at the same position in a formula, the nuclide symbols are written in alphabetical order; when their atomic symbols are identical the order is that of increasing mass number. Isotopically modified compounds may be classified as *isotopically substituted* compounds and *isotopically labelled* compounds.

IR-4.5.2 **Isotopically substituted compounds**

An isotopically substituted compound has a composition such that all the molecules of the compound have only the indicated nuclide(s) at each designated position. The substituted nuclides are indicated by insertion of the mass numbers as left superscripts preceding the appropriate atom symbols in the normal formula.

Examples:

- | | | | |
|----|-----------------------------------|----|--|
| 1. | H^3HO | 5. | $^{32}\text{PCl}_3$ |
| 2. | H^{36}Cl | 6. | $\text{K}[^{32}\text{PF}_6]$ |
| 3. | $^{235}\text{UF}_6$ | 7. | $\text{K}_3^{42}\text{K}[\text{Fe}(^{14}\text{CN})_6]$ |
| 4. | $^{42}\text{KNa}^{14}\text{CO}_3$ | | |

IR-4.5.3 **Isotopically labelled compounds**

IR-4.5.3.1 *Types of labelling*

An isotopically labelled compound may be considered formally as a mixture of an isotopically unmodified compound and one or more analogous isotopically substituted compounds. They may be divided into several different types. Specifically labelled compounds and selectively labelled compounds are treated here.

IR-4.5.3.2 *Specifically labelled compounds*

An isotopically labelled compound is called a specifically labelled compound when a unique isotopically substituted compound is added formally to the analogous isotopically unmodified compound. A specifically labelled compound is indicated by enclosing the appropriate nuclide symbol(s) and multiplying subscript (if any) in square brackets.

Examples:

- | | | | |
|----|---|----|--|
| 1. | $\text{H}[^{36}\text{Cl}]$ | 4. | $[^{13}\text{C}]\text{O}[^{17}\text{O}]$ |
| 2. | $[^{32}\text{P}]\text{Cl}_3$ | 5. | $[^{32}\text{P}]\text{O}[^{18}\text{F}_3]$ |
| 3. | $[^{15}\text{N}]\text{H}_2[{}^2\text{H}]$ | 6. | $\text{Ge}[{}^2\text{H}_2]\text{F}_2$ |

IR-4.5.3.3 *Selectively labelled compounds*

A selectively labelled compound may be considered as a mixture of specifically labelled compounds. It is indicated by prefixing the formula by the nuclide symbol(s) preceded by any necessary locant(s) (but without multiplying subscripts) enclosed in square brackets.

Examples:

1. $[^{36}\text{Cl}]\text{SOCl}_2$
2. $[^2\text{H}]\text{PH}_3$
3. $[^{10}\text{B}]\text{B}_2\text{H}_5\text{Cl}$

The number of possible labels for a given position may be indicated by subscripts separated by semicolons added to the atomic symbol(s) in the isotopic descriptor.

Example:

4. $[1\text{-}^2\text{H}_{1;2}]\text{SiH}_3\text{OSiH}_2\text{OSiH}_3$

IR-4.6 OPTIONAL MODIFIERS OF FORMULAE

IR-4.6.1 Oxidation state

The oxidation state of an element in a formula may be indicated by an oxidation number written as a right upper index (superscript) in Roman numerals. Oxidation state zero may be represented by the numeral 0 but is not usually shown. If an element occurs with more than one oxidation state in the same formula, the element symbol is repeated, each symbol being assigned a number cited in sequence in increasing magnitude and from negative to positive.

Examples:

- | | |
|--|---|
| 1. $[\text{P}^{\text{V}}_2\text{Mo}_{18}\text{O}_{62}]^{6-}$ | 4. $\text{Pb}^{\text{II}}_2\text{Pb}^{\text{IV}}\text{O}_4$ |
| 2. $\text{K}[\text{Os}^{\text{VIII}}(\text{N})\text{O}_3]$ | 5. $\text{Na}_2\text{O}^{-\text{I}}_2$ |
| 3. $[\text{Mo}^{\text{V}}_2\text{Mo}^{\text{VI}}_4\text{O}_{18}]^{2-}$ | 6. $[\text{Os}^0(\text{CO})_5]$ |

Where it is not feasible or reasonable to define an oxidation state for each individual member of a group (or cluster), the overall oxidation level of the group should be defined by a formal ionic charge, indicated as in Section IR-4.3. This avoids the use of fractional oxidation states.

Examples:

- | | |
|-------------------|---------------------------------|
| 7. O_2^- | 8. $\text{Fe}_4\text{S}_4^{3+}$ |
|-------------------|---------------------------------|

IR-4.6.2 Formulae of radicals

A radical is an atom or molecule with one or more unpaired electrons. It may have positive, negative or zero charge. An unpaired electron may be indicated in a formula by a superscript dot. The dot is placed as a right upper index to the chemical symbol, so as not to interfere with indications of mass number, atomic number or composition. In the case of diradicals, *etc.*, the superscript dot is preceded by the appropriate superscript multiplier. The radical dot with its multiplier, if any, precedes any charge. To avoid confusion, the multiplier and the radical dot can be placed within parentheses.

Metals and their ions or complexes often possess unpaired electrons but, by convention, they are not considered to be radicals, and radical dots are not used in their formulae. However, there may be occasions when a radical ligand is bound to a metal or metal ion where it is desirable to use a radical dot.

Examples:

- | | | | |
|----|------------------------------|----|--------------------------------|
| 1. | H [•] | 5. | O ₂ ^{•-} |
| 2. | HO [•] | 6. | BH ₃ ^{•+} |
| 3. | NO ₂ [•] | 7. | PO ₃ ^{•2-} |
| 4. | O ₂ ^{2•} | 8. | NO ^{(2•)-} |

IR-4.6.3 **Formulae of optically active compounds**

The sign of optical rotation is placed in parentheses, the wavelength (in nm) being indicated as a right subscript. The whole symbol is placed before the formula and refers to the sodium D-line unless otherwise stated.

Examples:

- (+)₅₈₉[Co(en)₃]Cl₃
- (-)₅₈₉[Co{(-)NH₂CH(CH₃)CH₂NH₂}₃]Cl₃

IR-4.6.4 **Indication of excited states**

Excited electronic states may be indicated by an asterisk as right superscript. This practice does not differentiate between different excited states.

Examples:

- He*
- NO*

IR-4.6.5 **Structural descriptors**

Structural descriptors such as *cis-*, *trans-*, *etc.*, are listed in Table V. Usually such descriptors are used as italicised prefixes and are connected to the formula by a hyphen.

Examples:

1. *cis*-[PtCl₂(NH₃)₂]
2. *trans*-[PtCl₄(NH₃)₂]

The descriptor μ designates an atom or group bridging coordination centres.

Example:

3. [(H₃N)₅Cr(μ -OH)Cr(NH₃)₅]⁵⁺

IR-4.8

REFERENCES

1. This is the so-called Hill order. See, E.A. Hill, *J. Amer. Chem. Soc.*, **22**, 479 (1900).
2. For intermetallic compounds, earlier recommendations prescribed alphabetical ordering rather than by electronegativity (see Section I-4.6.6 of *Nomenclature of Inorganic Chemistry, Recommendations 1990*, Blackwell Scientific Publications, Oxford, 1990).
3. For example, the ordering of B and C in formulae was inconsistent in *Nomenclature of Inorganic Chemistry, Recommendations 1990*, Blackwell Scientific Publications, Oxford, 1990.
4. **Title of New Blue Book needed to replace:** *Nomenclature of Organic Chemistry*, 1979 edition, p. 480, Rule E-4.4.
5. Chapter II-2 of *Nomenclature of Inorganic Chemistry II, Recommendations 2000*, Royal Society of Chemistry, 2001.