IUPAC Recommendations

Albert J. Dijkstra*, Karl-Heinz Hellwich, Richard M. Hartshorn, Jan Reedijk and Erik Szabó

End-of-line hyphenation of chemical names (IUPAC Provisional Recommendations)

Abstract: Chemical names and in particular systematic chemical names can be so long that, when a manuscript is printed, they have to be hyphenated/divided at the end of a line. Many systematic names already contain hyphens, but sometimes not in a suitable division position. In some cases, using these hyphens as end-of-line divisions can lead to illogical divisions in print, as can also happen when hyphens are added arbitrarily without considering the 'chemical' context. The present document provides recommendations and guidelines for authors of chemical manuscripts, their publishers and editors, on where to divide chemical names at the end of a line and instructions on how to avoid these names being divided at illogical places as often suggested by desk dictionaries. Instead, readability and chemical sense should prevail when authors insert optional hyphens. Accordingly, the software used to convert electronic manuscripts to print can now be programmed to avoid illogical end-of-line hyphenation and thereby save the author much time and annoyance when proofreading. The recommendations also allow readers of the printed article to determine which end-of-line hyphens are an integral part of the name and should not be deleted when 'undividing' the name. These recommendations may also prove useful in languages other than English.

Keywords: chemical nomenclature; dividing chemical names; end-of-line hyphenation; systematic chemical names; typesetting; word processing.

CONTENTS

*Corresponding author: Albert J. Dijkstra, Ajuinlei 13c, 9000 Gent, Belgium, e-mail: albert@dijkstra-tucker.be. https://orcid.org/0000-0001-9031-6292
Karl-Heinz Hellwich: Beilstein-Institut zur Förderung der Chemischen Wissenschaften, Trakehner Str. 7 – 9, 60487 Frankfurt, Germany. https://orcid.org/0000-0002-4811-7254
Richard M. Hartshorn: School of Physical and Chemical Sciences, University of Canterbury, Private Bag 4800, Christchurch, New Zealand. https://orcid.org/0000-0002-6737-6200
Jan Reedijk: Leiden Institute of Chemistry, PO Box 9502, 2300 RA Leiden, the Netherlands. https://orcid.org/0000-0002-6739-8514
Erik Szabó: Department of Physical and Theoretical Chemistry, Comenius University in Bratislava, Bratislava, Slovakia. https://orcid.org/0000-0002-5976-983X

Hyp-1 Introduction

Chemical compounds can be described in print in several different ways.

1. By various types of chemical names:
   a. Systematic names such as those recommended by the International Union of Pure and Applied Chemistry (IUPAC).
   b. Accepted or retained trivial names such as the names of the amino acids, several carboxylic acids, ketones such as acetone, alcohols such as ethylene glycol, and many other names.
   c. Trivial names such as soda ash (sodium carbonate).
   d. Trade names such as IMODIUM®, a registered trademark for loperamide which is the International Nonproprietary Name (INN) for 4-[4-(4-chlorophenyl)-4-hydroxypiperidinyl]-*N,N*-dimethyl--2,2-diphenylbutanamide, or Teflon® (polytetrafluoroethylene).
   e. Other kinds of names such as the INNs for pharmaceutical substances, for example trandolapril, a compound for which the systematic name reads: (2*S*,3a*R*,7a*S*)-1-[(2*S*)-2-{[(2*S*)-1-ethoxy-1-oxo-4-phenylbutan-2-yl]amino}propanoyl]octahydro-1*H*-indole-2-carboxylic acid.
2. By using abbreviations or acronyms such as DMK (dimethyl ketone; acetone; propane-2-one) or DCM (dichloromethane), but which need to be explained on their first use within a document [1].
3. By using the computer generated InChI and InChIKey.
4. By using molecular or structural notations such as:
   a. Linear line formulae, like $C_{24}H_{34}N_2O_5$ and Na_2HPO_4.

   b.   Structural formulae for linear compounds, such as acrylamide $CH_2=CH\text{-}CONH_2$ and trifluoro-methanesulfonic anhydride $(CF_3SO_2)_2O$.

   c.   Shorthand structural formulae, such as NaOMe for sodium methanolate.

   d.   Shorthand notations for proteins showing the amino acid order by using either the three-letter symbols or the one-letter symbols [2].

   e.   Shorthand notations for nucleotides that show the sequence of nucleic acid fragments in DNA by using a string of one-letter symbols, like in d(AGAGCTAGCTCT).

Chemical names, and in particular systematic chemical names, can be so long in print that they have to be broken, so that the name can be spread over more than one line of text. With respect to trivial names, dictionaries such as *Merriam-Webster's New Collegiate® Dictionary* [3], the *New Oxford Spelling Dictionary* [4], or the *Dictionary of Contemporary English* [5] indicate at which points words can be hyphenated. However, it should be noted that these dictionaries often fail to take the chemistry into account. Also, there are sometimes division differences between UK English and US English.

On the other hand, systematic names in many cases already contain hyphens, but sometimes not in a suitable position for dividing at the end of a line. This means that hyphens have to be inserted into the name between characters. Since there is a fair amount of confusion about the appropriate positions, the present document provides pertinent recommendations and guidelines. While in principle being general, it mainly focuses on the hyphenation of systematic chemical names.

When an article is about to be published, the author sends the electronic manuscript to the publisher who uses typesetting software to convert the manuscript into print. This conversion may entail dividing chemical names at the end of a line. The copyediting software used by the publisher, like the word processing software used by the authors, recognizes a hyphen as a location where a word can be divided. So a name such as '1-methyl-1,2-dihydronaphthalene' may get divided at the third hyphen in this name. However, this does not make sense chemically. In chemical terminology, the '1,2' locants are closely associated with the 'dihydro'. Just listen to yourself when you pronounce the name. You stop after '1-methyl' and again after '1,2-dihydro' and you pronounce '1,2-dihydro' as if it were a single word. There is therefore a need to prevent the software from dividing names or inserting hyphens at illogical places.

In general, word processing and typesetting software makes use of pre-loaded dictionaries that will show how to divide words at the end of a line. Accordingly, it will contain words like: 'an-a-lyse', 'anal-y-sis', 'an-a-lyst or ana-lyst',[1] and so forth. It will contain some chemical names like 'ac-et-al-de-hyde' and 'acet-amide', but '1-methyl-1,2-di-hy-dro-naph-tha-lene' is less likely to have been included. Consequently, names like that will be divided manually or by some kind of computer assumption in accordance with the typographical needs at the time, and the result of that division will then be included in the database. So, it may contain '1-methyl-1,2-dihydro-naphthalene' and/or '1-methyl-1,2-di-hydronaphthalene' and/or '1-methyl-1,2-dihydronaphthalene', etc. Consequently, there is clearly a need to provide databases with names that are divided at logical places and to prevent names that have been divided at illogical places from being included in these databases.

Authors may themselves have divided chemical names in their manuscripts for the simple reason that they were too long and didn't fit on the line. When this division is no longer necessary in print, the typesetter will have to know whether or not to maintain this end-of-line hyphen. Is it part of the name, or is it merely an end-of-line division? There is therefore a need to distinguish between these two kinds of hyphen in a manuscript.

In a later stage of publishing, the publisher sends the proofs to the author who is likely to encounter such illogical end-of-line divisions. If the author corrects these illogical divisions, these corrections may lead to subsequent illogical divisions further down the paragraph (or even on other pages), necessitating a second proofreading. Authors would therefore be well served if proofs did not contain such illogical end-of-line

---

1 These examples have been chosen to illustrate the unpredictability of where to divide normal words [3, 5]. Other dictionaries may provide even other hyphenations, e.g. [4].

divisions. These can be avoided by indicating in their manuscript where a systematic chemical name may be divided in accordance with the present recommendations and which hyphens form part of the systematic name and should not be used for end-of-line divisions. Fortunately, word processing software in use provides means to indicate this difference, but this requires that such indications are correctly interpreted by the typesetting software used by the publisher. More detailed instructions for authors, publishers and typesetters are given in the Annex (Hyp-8) to these recommendations.

Hyp-2 Definitions

Hyp-2.1 Name components

Systematic chemical names describe the structure of a compound by listing, for example, its parent name, its substituents and/or added terms, or the names of the ligands and the central atom of a complex. In the context of the present recommendations these name parts will be referred to as name components. Name components refer to atoms or groups of atoms that have a name. So, the name 'chloroethane' consists of two name components: the group of atoms forming the constituent 'ethane', a parent hydride name, and the chlorine atom as the substituent 'chloro', a substituent prefix. When the 'chloroethane' itself becomes a substituent and is then referred to as 'chloroethyl', this substituent also consists of two name components: 'chloro' and 'ethyl'.

   When more than one chlorine substituent is present as in '1,2-dichloroethane', the multiplicative prefix (di) and the locants (1 and 2) indicating the positions of the chlorine atoms form an integral part of that particular name component ('1,2-dichloro').

Hyp-2.2 Prefixes

Besides substituent prefixes, chemical names can contain various other prefixes. The multiplicative prefix indicates the number of times a name component occurs, *i.e.* 'mono', 'di', 'tri', 'tetra', 'penta', 'hexa', etc. for non-substituted name components, and 'bis', 'tris', 'tetrakis', etc. for identically substituted moieties; these prefixes can distinguish between isomers, *e.g.* 'iso', '*tert*'; they can indicate a ring structure or modification, *e.g.*, 'cyclo', 'spiro', 'homo', 'nor', 'seco', etc. Names of polymers, in most cases, have the prefix 'poly'.

   Prefixes can also specify the configuration, in which case they are called stereodescriptors (see Hyp-2.6). When used as locants in *e.g.* stereodescriptors (see Hyp-2.5), they indicate a position.

   Prefixes or part of them, depending on the kind of prefix, can be joined to the name component they are referred to by a hyphen, as in '*tert*-butyl', but not necessarily so, as in 'isopropyl'.

Hyp-2.3 Suffixes

Suffixes in chemical names always indicate a functional group, usually the principal characteristic group, or the free valence of a substituent group such as 'yl' or 'ylidene'. Suffixes can be short as the suffixes 'ol' indicating an alcohol group, 'ide' for an anionic position and 'one' indicating a ketone, but there are also longer suffixes like 'amide' or those consisting of more than one component, *e.g.* 'oic acid'.

Hyp-2.4 Endings

Endings are part of the parent name and are therefore regarded as part of the name component. The most common examples of endings are those for the kind and number of multiple bonds, *e.g.* 'ane' as in hexane,

'ene' as in propene and 'yne' as in ethyne. Other examples are 'ine' in pyridine, or in the names of the halogens, *e.g.* chlorine, or 'olidine' in pyrrolidine.

Hyp-2.5 Locants

A locant is a numeral or a letter, or a combination of both that identifies position(s) in a structure (modified from [6, 7]). Typical examples are '1', '4a', '*N*' or '*o*' (*ortho*), '*m*' (*meta*) and '*p*' (*para*). More details on the different kinds and the use of locants can be found for organic-chemical compounds in P-14.3 [8], for inorganic compounds in IR-2.14 [9], and for polymers for example in [6, 7, 10]. In the vast majority of cases, locants are separated from each other by commas and from surrounding text by hyphens. A locant can also appear in a general expression, such as '2-substituted' or 'α-hydrogen'.

Hyp-2.6 Stereodescriptors

A stereodescriptor is a prefix to specify the configuration (absolute or relative) or conformation [11] of a compound or a stereogenic unit of the compound. Typical examples are *R*, *S*; *r*, *s*; *P*, *M*; *A*, *C*; *R*$_a$, *S*$_a$; *E*, *Z*; *cis*, *trans*; D, L; *ap*, *sp*. Many, but not all, stereodescriptors are directly preceded by the corresponding locant and, together with its locant, enclosed in parentheses. For more details on the use of stereodescriptors see P-9 [8] and IR-9.3 [9].

Hyp-2.7 Hyphens and related symbols used in chemical nomenclature

In this and subsequent sections, special symbols will be described by the shortcut keys used to enter these symbols in PCs using Microsoft Word. Users of different hardware (like Mac) and/or software (like Linux or Latex), should use these descriptions as a starting point for working out equivalent procedures.

The hyphen (-, Alt + 45) is also present on the keyboard. In normal text, it is used to link words which closely belong together, or to link two parts of a word at a line break. In chemical nomenclature it is used in several contexts, most importantly to separate locants, stereodescriptors, or italic prefixes from other parts of the name. Full details on the use of hyphens can be found in sections IR-2.3.1 [9] and P-16.2.4 [8]. As in general typesetting, the hyphen is also used for dividing chemical names at a line break.

In addition to the normal or regular hyphen (cf. Hyp-3.1), two more symbols represented by a horizontal line are in use: the en dash (–, Alt + 0150) and the em dash (—, Alt + 0151). They differ in length: hyphen (-), en dash (–), and em dash (—). For details on the use of the em dash see IR-2.3.3 [9] and P-16.2.5 [8]. Examples for the use of the en dash can be found in IR-10.2.5.1 [9], and in [10]. (The use of the en dash as originally prescribed in polymer nomenclature has not been continued in the new edition of [7].)

A related symbol is the minus sign (−, 2212, Alt + X), The minus sign is very similar, but not identical to the en dash (see for example IR-2.3.2 [9]). It is used to indicate the charge on an ion or in a name, like in tetracarbonylcobaltate(1 −) and therefore never functions as a hyphen.

Word processors treat the en dash and the em dash like normal hyphens. The minus sign on the other hand is treated as a special character, just as the non-breaking hyphen, sometimes also called a hard hyphen (see Hyp-3.2).

Chemical formulae may also contain a solidus (/, Alt + 47) in polymer descriptions like 'poly[(chloromethylene)/methylene]', or to separate the Arabic numerals which indicate the proportion of individual constituents in a compound like 'boron trifluoride—water(1/2)'.

The center dot (·, also called the middle dot, Alt + 0183), is also used in compound formulae, as shown in $3Cd(SO_4) \cdot 8H_2O$ and $Cu(SO_4) \cdot 5H_2O$.

Finally, the colon (:, Alt + 58 and also present as a key) can be used in compound names, such as: di-μ-hydroxido-μ-nitrito-κ*N*:κ*O*-bis(triamminecobalt)(3 +).

Hyp-2.8 Special symbols and practices

In some documents specific symbols have been used instead of, or in addition to, a hyphen in order to indicate a division of a name at a position where a hyphen is not part of the name. In some documents the equals sign (=) was used for that purpose, for example in [12] and in the German translation of the 1990 edition of the Red Book [13].

The 2005 edition of the Red Book [9], *e.g.* on page 265, used the symbol ◠ which is a 90 degrees rotated pair of opening/closing parentheses, and similar to the symbol used in printed CAS Index Guides [14].

In the *Principles of Chemical Nomenclature* [15] the very conspicuous symbol ▸ was added when a chemical name was broken at the end of a line, even after a hyphen. Similarly, the symbol ¬ is used in *British Approved Names*, BAN [16]. The same rationale is followed by other authors who always add a hyphen – even after an existing hyphen – when a name was broken at the end of a line, like in amino acid sequences (see below in Hyp-7.5). This last approach has also been adopted in the present recommendations.

Hyp-3 Characters for correct typesetting of hyphenation

Hyp-3.1 Regular hyphen

The normal, or regular hyphen (-, see Hyp-2.7) as entered via the keyboard is shown as such in this document. If it occurs within a chemical name and the space left on the line is too short to accommodate the rest of the name, it acts as a dividing hyphen by relegating the rest of the name to the next line. This holds for the word processing software used by authors and for the typesetting software used by publishers. Office printers and typesetting software treat the regular hyphen in an identical manner.

Hyp-3.2 Non-breaking hyphen

The non-breaking or hard hyphen, *i.e.* a hyphen at which no line break is allowed, is not treated as a hyphen by word processing software. Consequently, a line never ends on a non-breaking hyphen. Using keyboard shortcut Ctrl + Shift + Hyphen (-) encodes a non-breaking hyphen. It is shown on screen as a slightly elongated hyphen (–) comparable to an en dash (see Hyp-2.7) and it is printed as a normal hyphen by office printers. The way typesetting software interprets the non-breaking hyphen encountered in manuscripts may differ from one software package to another.

In the present text, the equals sign '=' will be used to represent and distinguish the non-breaking hyphen from other hyphens.

Hyp-3.3 Optional hyphen

The optional or soft hyphen is used to indicate at what point a word can be divided if this were to be convenient. It can be entered by keyboard shortcut Ctrl + Hyphen (-). By clicking the Show/Hide (Ctrl + *) button, or clicking the '¶' symbol on the command bar in MS Word, it becomes visible on screen as a hyphen with a small vertical line descending from the right-hand side of the hyphen (¬). When the word is divided at the point indicated by the optional hyphen, a normal hyphen appears in print at the end of the line. Otherwise, office printers omit the hyphen in print. The way typesetting software interprets the optional hyphen again depends upon what software is used.

In the present text, the '~' (tilde) symbol will be used to indicate an optional hyphen.

Hyp-3.4 Non-breaking space

As will be illustrated in Hyp-4.1.1 and Hyp-4.12, there are occasions where two name components that are separated by a space should not be separated from each other by a line break. For example 'boron trifluoride—water°(1/2)' should not be divided before '(1/2)'. In such a case, a hard or non-breaking space must be used to prevent a line break at this position. Ctrl + Space can be used to encode non-breaking spaces. As illustrated by the above example, it shows on screen as a kind of 'degree' symbol: '°'

    In the present text, the '°' (degree) symbol will be used to indicate non-breaking spaces.

Hyp-3.5 No-width non break

A no-width non-break, is a non-printing character that prevents a word, number, or phrase from breaking if it falls at the end of a line. In chemical nomenclature, this is especially important when using the en dash or the em dash, since word processors treat them as normal hyphens and there are instances, like bis(pentacarbonylmanganese)(*Mn–Mn*), in which name components joined by a dash should not be divided.

    In addition to creating a shortcut key, there are three ways to introduce a no-width non break:

1. By using shortcut Alt + 65279
2. By typing FEFF and pressing Alt + X
3. By insertion: Insert → Symbol → More symbols → Special Characters → No-width non break → Insert

    In the present text, the '*' (asterisk) symbol will be used to indicate a no-width non break.

Hyp-4 Rules for dividing chemical names

Hyp-4.1 General rules for dividing chemical names at the end of a line

The rules for dividing chemical names at the end of a line consist of prohibitions and prescriptions which are imperative and are listed in the following subsections, and additional guidelines, which are mainly advisory. In fact, the only purpose of the rules is to make it easier to read a printed text and to identify hyphens that have been added when dividing the name, but that have to be deleted on 'undividing' the divided name. So **how** to divide and **where** (not) to divide will be the topics discussed in the sections below.

Hyp-4.1.1 Division at spaces

A chemical name consisting of two or more separate words, *e.g.* benzoic acid or ethyl acetate, can generally be divided at an existing space when necessary. In that case no additional symbol is used to indicate this division. For exceptions see Hyp-4.12.

Hyp-4.1.2 Division at hyphens

A chemical name containing hyphens should preferentially be divided at a hyphen that is already present in the name. However, not every hyphen in a name is suitable as the last symbol at the end of a line (see Hyp-4.4

for such situations). Moreover, and in line with the way proteins are divided at the end of a line when written as a series of amino acids, (cf. Hyp-7.5) the present recommendations prescribe that the breaking hyphen has to be repeated, so that the upper line ends with a hyphen and the lower line starts with a hyphen. This additional hyphen also greatly facilitates the undividing (cf. Hyp-5) of divided names: Just delete one hyphen when there are two.

To ensure that the lower line starts with a hyphen, the name component after which the name has to be divided should be followed by an optional hyphen that indicates that a break at that point is allowed, and a non-breaking hyphen that will move to the next line.

Hyp-4.1.3 Insertion of hyphens

A chemical name not containing a hyphen at a suitable position for dividing at the end of a line can be divided at a chemically meaningful position by adding a hyphen at the end of the line and putting the rest of the name onto the next line. Dividing a chemical name at the end of a line without adding a hyphen that indicates continuation of the name on the next line is, like in any other word, unacceptable – except at a space in the name (see Hyp-4.1.1) or in cases covered by Hyp-4.1.4. In particular, a hyphen needs to be inserted if a division is made after enclosing marks, or other symbols in a name.

For example, the compound 2-(chloromethyl)benzoic acid can be divided after 'benzoic' without an additional symbol, but if it had to be divided before 'benzoic', a hyphen would have to be inserted after the 2-(chloromethyl), *i.e.* 2-(chloromethyl)~benzoic acid, which would appear in print as a normal hyphen and line break:

2-(chloromethyl)-
benzoic acid.

The fact that the name component of the divided name on the lower line does not start with a hyphen shows that the hyphen at the end of the upper line has been inserted and should be deleted on undividing. This is in fact an example of a general rule that every hyphen at the end of a line should be deleted on undividing.

Hyp-4.1.4 Line breaks at special characters (solidus, en dash and em dash)

When a chemical name contains a solidus (oblique stroke, slash, /), an en dash or an em dash, and these symbols are used in a function similar to a hyphen, as for example in names of addition compounds, like boron trifluoride—water (1/2), they may be used to divide a chemical name at that location. Since word processors treat the en dash and the em dash as normal hyphens (cf. Hyp-2.7), a name will be automatically divided after these dashes if desirable. However, to distinguish the en dash or em dash from a normal hyphen and in line with Hyp-4.1.2, the en dash or em dash should be preceded by an optional hyphen, and insert after the hyphen a no-width non-breaking space.

If division at this point is to be prevented, no-width non breaks have to be inserted on both sides of the en dash or em dash.

The solidus on the other hand, is not treated as a normal hyphen. Consequently, an optional hyphen (~) must be inserted before the solidus to indicate that the name can be divided at that point; in case of division, this will cause the line to end with a hyphen and the next line to start with the solidus.

Hyp-4.1.5 Unacceptable insertion points for end-of-line hyphens

There are positions in a chemical name at which a division or line break is unacceptable. These are listed below:
  a.   within a locant (*e.g.* 13 must not be divided as 1-3).
  b.   within a stereodescriptor (*e.g. endo* or *RS* should not be divided, as *en-do* or *R-S*, respectively).

c. between a stereodescriptor and its locant [*e.g.* (13*R*) should not be divided as (13-*R*)].

d. within descriptors like '2λ⁵', '3δ²', or 'κ*N*¹:κ*N*², *e.g.* 1,2=di~hydro-1λ⁵=phos~phine. This should not be divided as 1,2=di~hydro-1-λ⁵=phos~phine, or 1,2=di~hydro-1λ-⁵=phos~phine; see also item i)

e. directly before a closing enclosing mark; *e.g.* ….methyl) should not be divided as ….methyl-).

f. directly after an opening enclosing mark; *e.g.* (chloro…. should not be divided as (-chloro…. .

g. in between opening or between closing marks. Accordingly, the compound with the molecular formula [Mo₆Cl₈Cl₃{(C₆H₅)₂PCH₂CH₂P(C₆H₅)₂}py]Cl should not be hyphenated according to:
[Mo₆Cl₈Cl₃{-
(C₆H₅)₂PCH₂CH₂P(C₆H₅)₂}py]Cl.

h. directly before or after a punctuation mark, such as a comma, semicolon, or colon; 1,1,1-trichloro should not be divided after a comma as 1,-1,1-trichloro,. and κ*N*:κ*O* should not be divided after the colon as κ*N*:- κ*O*. For some exceptions see Hyp-4.10.

i. directly after an em dash that represents a bond such as in bis(pentacarbonylmanganese)(*Mn–Mn*), or in the bridge descriptors, *e.g.* [B1–B1'], in names of polycyclic macromolecules [10].

j. directly before a valence or charge indication; *e.g.* not Cu-(II) or Fe-(3 +).

k. directly before a right subscript or superscript (*e.g.* not SO-₄), directly after a left subscript or super-script (*e.g.* not ³¹-P) and within subscripts or superscripts (*e.g.* not ²³-⁵U).

l. within a set of primes; *e.g.* so that N‴ should always be on the same line; when more than 3 primes are needed, a non-breaking space should be inserted to assist reading.

m. within an italicized prefix (*e.g.* *arachno* should not be divided).

n. within an element symbol (*e.g.* not N-a) or within a group of atoms (*e.g.* not N-O₃).

o. within a formula having a middle dot (*e.g.* not Cu(SO₄)·-5H₂O).

p. after a solidus within a constituent ratio (*e.g.* not boron trifluoride—water(3/-8)

q. within a common abbreviation or acronym (*e.g.* not in FeMoco, EXAFS, NADPH).

Hyp-4.2 Division between name components

A systematic chemical name written as one word and comprising two or more name components, can be divided between these components. Accordingly, 'chloroethane' can be divided according to 'chloro~ethane' where the ~ sign indicates the optional hyphen as defined above in Hyp-3.3. Similarly, 'diamminedichlorido-platinum' can be divided as 'diammine~dichlorido~platinum', or 'poly(oxyethylene)' as 'poly(oxy~ethylene)'. If a name component contains enclosing marks as in '4-(chloromethyl)pyridine' or 'dichloridobis(urea)-copper(II)', the hyphen has to be inserted after the closing enclosing mark: '4-(chloromethyl)~pyridine' and 'dichloridobis(urea)~copper(II)', respectively. Of course, a hyphen can always be inserted between two groups of parentheses, as in tricarbonyl(triethylphosphane)~(trimethylsilyl)cobalt.

Hyp-4.3 Minimum number of non-divided consecutive characters

The ACS Style Guide of the American Chemical Society [17] recommends leaving at least three characters on each line whereby a hyphen is counted as a character. Accordingly, it allows the name component 'cyclohexyl' in a name like 'cyclohexylamine' to be divided at the indicated positions: 'cy~clo~hex~yl'. In order to improve the recognizability of divided words, the present recommendations opt for a larger number of characters and suggest a number of at least six or exceptionally five consecutive characters at the end of a line (where the dividing hyphen should not be counted as a character). This count should be made from the beginning of the word. The number of consecutive characters on the next line can be smaller, but not less than two characters, for instance 'butan~one'.

Applying this rule to the above example will prevent its name component from being divided as 'cy-clohexyl' but will allow it to be divided as 'cyclo-hexyl'. On the other hand, despite the fact that dividing as

'cyclohex-ylamine' would provide sufficient characters at the end of the line, it is not recommended because the suffix '-yl' is part of the name component and should not be divided from it (see Hyp-4.2 and Hyp-4.6).

Hyp-4.4 No division after locants or descriptors

In Hyp-4.1.5 it has already been stated that a division between a stereodescriptor and its locant is unacceptable. In addition, it is recommended that a division is generally avoided at a hyphen following a locant. In the Introduction (Hyp-1), the pronunciation of '1-methyl-1,2-dihydro~naphthalene' was discussed and it was concluded that the '1,2' is closely associated with the name component 'dihydro'. In general, locants are to be considered part of the name component that follows. Accordingly, the hyphen between the locant or locant set and the name component must be considered as a non-breaking hyphen (see Hyp-3.2). This is particularly important for contracted forms of substituent groups such as '2-pyridyl'. Using the notation defined in Hyp-3 therefore leads to '1=methyl~=1,2=dihydro~naphthalene' and '2=pyridyl', respectively.

⠀⠀⠀This procedure also extends to the situation where a name component is preceded by both a stereodescriptor and a locant. Accordingly, non-breaking hyphens should be used, as shown by the example '(2*R*)=2=chloro~butane'.

⠀⠀⠀Systematic names can also contain letters that are followed by numerals or *vice versa*, such as 'indicated hydrogen' or 'added hydrogen' (see P-14.7 [8]), *e.g.* '2*H*', '2(3*H*)', lambda and delta descriptors (P-21.2.4 and P-25.7.2 [8]), *e.g.* '2λ^5', '3δ^2', and other descriptors that do not contain hyphens, *e.g.* '(^{13}C,^{2}H)' (Chapter P-8 [8]), 'κN^1:κN^2' (IR-9.2.4.2 [9]). These descriptors should not be separated from their moieties by an end-of-line break. See Hyp-4.1.5 item d.

Hyp-4.5 Division between name components and prefixes

In general, an end-of-line division between a name component and a prefix is acceptable if they are written as one word, *i.e.* without space or hyphen. Therefore, 'cyclohexane' can be divided between the prefix and the compound name according to 'cyclo~hexane'. The name '1,2-dichloroethane' can also be divided after the prefix, *i.e.* '1,2=dichloro~ethane'. In fact, it can also be divided after the prefix 'di' according to: '1,2=di~chloroethane', because the number of symbols in front of the optional hyphen equals six, but a division according to '1,2=dichloro~ethane' is far easier to read.

⠀⠀⠀In this respect, authors should use their discretion in applying dividing rules. A systematic name starting with 'tetrakis(….' can be divided before the parenthesis according to 'tetrakis ~(….' since 'tetrakis-' has eight consecutive characters before the hyphen and a chemist knows that the substituted moiety will be described after the parenthesis. Consequently, a name starting with 'bis(….' or 'tris(….' is also allowed to be divided in a similar way, despite the fact that 'bis-' and 'tris-' have fewer than six consecutive characters. Accordingly, 'tris~(ethane~=1,2=di~amine)~cobalt(III)', 'tetra~ammine~aqua~cop~per(II)' and '*cis*=di~ammine~di~chlorido~pla~ti~num(II)' are perfectly acceptable. The same rationale holds for polymers in which a division is generally allowed after the prefix 'poly', *e.g.* 'polystyrene' can be divided as 'poly~styrene'. Readability and chemical sense should prevail.

Hyp-4.6 Division before suffixes

When suffixes are preceded by a locant, the name can be divided at the hyphen before this locant (or locant set), *e.g.* 'pentan~=3=one'. If there is no locant, division before the suffix is similarly allowed, but a hyphen has to be inserted, *e.g.* 'cyclo~hexan~one', benzene~thiol, or 'heptan~oic acid'.

⠀⠀⠀However, in contracted forms such as 'phenol', 'ethyl', or '4=pyridyl' the suffixes are regarded as part of the name component, and a division before the suffix is unacceptable (just as the locant for the suffix is not

inserted in such contracted forms). Analogously, a division of 'tetrachloridoplatinate(II)' is possible at the following positions 'tetra~chlor~ido~pla~tin~ate(II)', and 'acetone' may be divided according to 'acet~one' if it is substituted as in '1,3=di~chloro~acet~one', but a division according to '1,3=di~chloro~acetone' is far easier to read.

Hyp-4.7 Division before endings

Endings are regarded as part of the name component and should therefore not be divided from the parent structure name. Contrary to the ACS Style Guide [17], the present recommendations keep name components with an ending like 'ethyne' together by shifting them to the next line. However, division before the locant or locants is permitted when the ending is preceded by a locant, *e.g.* 'hepta~=1,3=diene'. The systematic name for 'oleic acid' can therefore be divided according to '(9*Z*)=octa~dec~=9=en~oic acid' and the systematic name for another fatty acid, 'agonandoic acid', can be divided according to '(11*E*)=octa~dec~=11=en-9=yn~oic acid', where the '-ynoic' and the '-enoic', respectively, are regarded as endings with part of a suffix (see Hyp-4.6).

According to this exception '(2*Z*)=pent~=2=ene' can be divided as shown. Consequently, it is also acceptable to divide 'pentane' in the same way ('pent~ane') provided there are enough characters in front of 'pent~' to justify this division, as in '2=methyl~pent~ane'. This preferred way of division may therefore not be in accordance with the way some people pronounce it ('pen~tane') and should therefore preferably be avoided (*cf.* Hyp-4.9). Another example with similar reasoning is '*catena*-tribor~ane'.

However, dividing 'but-2=ene' at the first hyphen is not recommended since the number of letters in front of this hyphen is too small, which leads to 'but=2=ene', *i.e.* no division at all. When the number of characters is increased as in '1=chloro~but~=2=ene', division at the indicated hyphen would be possible, but a division according to '1=chloro~but=2=ene' is far easier to read.

Hyp-4.8 Generally accepted divisions

The last two examples in Hyp-4.7 also illustrate that the readability of words does not suffer when they are divided in a way that is generally accepted and readers will recognize. So 'octadec-' can be divided after 'octa~'. Similarly, 'phosphorus' can be divided according to: 'phos~phorus'. The prefix 'hydro' can be divided according to 'hy~dro' provided there are at least four characters in front as in '1=methyl~=1,2=dihy~dronaphthalene'.

Hyp-4.9 Division in accordance with pronunciation and chemical meaning

Division should not hamper pronunciation. Take the word 'anthracene'. This consists of two chemical terms: the root word 'anthrac' and the ending 'ene', but nobody says 'anthrac-ene', people say 'anthra-cene'. If the line had ended on 'anthrac-', readers might have started to pronounce this as 'anthrak'. So 'anthra~cene' is the generally accepted way of dividing this word. This way of dividing is preferred to the division according to 'anthr~acene'. This latter division is in accordance with the way 'poly~acenes' can be divided, but then endings like 'acenes' are part of a name component or name. So 'anthra~cene is preferred because that is how the name is pronounced. We note, however, that this means that the division of names may be both language and dialect dependent. Similarly, 'naphthalene' should not be divided according to 'naphthal~ene', but in accordance with its pronunciation as 'naphtha~lene', which is again preferred to 'naphth~alene'.

On the other hand, there are instances where a division may be acceptable according to the above described recommendations but would mislead the reader by implying a different chemical meaning. For example 'acetaldehyde' may be divided as 'acet-aldehyde', but not as 'acetal-dehyde' which is allowed by

some dictionaries [3, 4] and also the ACS Style Guide [17], 'formaldehyde' may be divided as' form-aldehyde', but not as 'formal-dehyde'.

Hyp-4.10 Division within a set of stereodescriptors or locants

Special attention should be given to names that include stereochemical information in the form of a string of descriptors. The systematic name of 'DHA' (docosahexaenoic acid) is: '(4*Z*,7*Z*,10*Z*,13*Z*,16*Z*,19*Z*)=do~cosa-4, 7,10,13,16,19=hexa~enoic acid'. If this name were to be divided after the 'do-', the preceding string (including the hyphen) would be twenty-seven characters long. That can lead to unsightly print. Accordingly, the stereo-descriptors should be divisible, according to:

$$\text{'}(4Z,7Z,{\sim}10Z,{\sim}13Z,{\sim}16Z,{\sim}19Z)=\text{'}$$

   In this example, no optional hyphen has been inserted between the first two stereodescriptors to ensure there are at least six characters at the end of a line. Accordingly, a name starting with two stereodescriptors such as the systematic name for 'linoleic acid', *i.e.* '(9*Z*,12*Z*)=octa~deca-9,12=di~en~oic acid', should keep these stereodescriptors together and keep them attached to the prefix ('octa'), despite the fact that this leads to a string of fourteen (including the hyphen) consecutive characters. The locant set for the double bonds in 'DHA' can also be divided according to:

$$\text{'do~cosa~}=4,7,{\sim}10,{\sim}13,{\sim}16,{\sim}19=\text{hexa-'}$$

   The same rules apply to the stereodescriptors '*R*' and '*S*'. Take, for instance, 'eudesmane'. The systematic name of this compound is: '(1*R*,4a*R*,7*R*,8a*S*)-1,4a-dimethyl-7-(propan-2-yl)decahydronaphtha-lene''. This name could be presented to the typesetter in a way that illustrates the rules outlined above: '(1*R*,4a*R*,~7*R*,~8a*S*)~=1,4a=di~methyl~=7=(propan~=2=yl)~deca~hy~dro~naphtha~lene)'.

Hyp-4.11 Division at special characters (solidus, en dash and em dash)

Chemical names can be divided at symbols which have a function comparable to a hyphen. So, names in which the components are separated by a solidus, '/', can be divided by preceding the solidus with an optional hyphen, which then becomes the last character on the line (see Hyp-4.1.4), for example 'poly[(chloromethylene)-/methylene]' can be divided according to 'poly ~[(chloro~methylene)~/methylene]'.

   poly[(chloromethylene)-
   /methylene].

   Also, if an em dash or an en dash is used in a chemical name in a function similar to a hyphen, as for example in names of addition compounds, they should be preceded by an optional hyphen and followed by a no-width non break as already explained in Hyp-4.1.4.
   Thus 'cadmium sulfate—water (3/8)' provides an example for both of the previous paragraphs in this section, in that it can be divided after 'sulfate' but not after '3/' (sulfate~—*water°(3/8)). For line breaks at dashes in formulae, see Hyp-7.

Hyp-4.12 Unacceptable division at spaces

As stated in Hyp-4.1.1, a chemical name can, in general, be divided at an existing space within the name. Exceptions to this rule are the following: A division at the space before a descriptor following a name, such as the numerical descriptor for the composition of an addition compound, or of a polymer should be avoided,

in order not to split such a descriptor from the name at a line break. Similarly, spaces within such descriptors should be considered as non-breaking spaces. For example 'boron trifluoride—water°(1/2)' should not be divided before '(1/2)', as indicated by the non-breaking space, and '(polyethylene)-*mod*-chloro(w_{Cl} = 0.32)' should not be divided after 'chloro', and a line break within the parenthetical descriptor '(w_{Cl} = 0.32)' is not acceptable either. Spaces before and after an equal sign, =, are considered as non-breaking spaces.

Hyp-5 'Undividing' divided systematic chemical names

The division of a chemical name at an existing hyphen, an en dash or an em dash by inserting an optional hyphen before the existing hyphen, en dash or em dash leads to a bis-hyphen sequence, on division this leads to the upper line ending on a hyphen and the lower line starting with a hyphen, an en dash or an em dash. Undividing this divided name is straightforward: It only requires the deletion of the optional hyphen, *i.e.* the hyphen at the end of the upper line leaving the original hyphen, the en dash, or the em dash in the undivided name. If the chemical name had been divided by inserting an optional hyphen between two characters, this optional hyphen will appear in print at the end of the upper line. On undividing, it has to be deleted.

Undividing systematic chemical names can therefore be summarized by the simple rule that all hyphens at the end of a line have been inserted for division purposes and must therefore be deleted on undividing.

Hyp-6 InChI and InChIKey

Closely related to chemical names are the InChI (IUPAC International Chemical Identifier) and the related InChIKey for a chemical structure. An InChI is produced by computer from structures drawn on-screen (or from derived 'mol' files) to represent a compound in a completely unequivocal manner. An InChIKey is a contracted form of defined length and format that is produced from the InChI by another algorithm. Whereas the length of an InChI depends on the compound, all InChIKeys have the same 27 character length, including two hyphens in fixed positions (positions 15 and 26).

Examples (InChI and InChIKey for chalcone):
  InChI: 1S/C15H12O/c16-15(14-9-5-2-6-10-14)12-11-13-7-3-1-4-8-13/h1-12H/b12-11+
  InChIKey: DQFBYFPFKXHELB-VAWYXSNFSA-N

Both InChIs and InChIKeys are electronic codes, *i.e.* machine-readable strings of symbols primarily intended to be interpreted by computers. Therefore, any addition of symbols, including spaces, is unacceptable for InChIs and InChIKeys. When they appear in print – as is increasingly the case – and do not fit into a line, which may be quite common for InChIs in particular, they may only be divided at a suitable existing symbol, *i.e.* after an hyphen that is already present, or a solidus. The first symbol on the next line should not be a comma. The InChIKey can only be divided at the first hyphen.

This method of dividing therefore does not involve the use of the bis-hyphen sequence. Accordingly, the hyphen at the end of part of the InChI or InChIKey must not be deleted on undividing, as would be the case for a systematic chemical name.

Hyp-7 Linear chemical formulae

The treatment of end-of line hyphenation would be incomplete without also considering formulae, even though formulae are not chemical names. Whenever possible, a line break within a line formula should be avoided. However, linear formulae can become quite long so in order to avoid unsightly appearance, formulae may also be divided at the end of a line. In that case, the rules described in Hyp-4.1 apply.

Hyp-7.1 The first type of formula listed in the introduction is the molecular formula that just lists how many carbon, hydrogen, oxygen and other atoms a molecule contains. Given the limited number of different atoms present in most molecules for which the molecular formulae are used, it is unlikely that molecular formulae will require end-of-line division. If it turns out to be desirable, or inevitable, molecular formulae can be divided by inserting an optional hyphen in front of an atom symbol and the presence of this hyphen will indicate that the molecular formula has been divided and that the hyphen must be removed on undividing.

Hyp-7.2 Another type of linear formula is the shorthand notation, such as NaOMe for sodium methanolate or EtOAc for ethyl acetate. Such shorthand formulae will usually be quite short so that no end-of-line division will be necessary. In the unlikely event that this division is deemed to be desirable, the guidelines of Hyp-4.1 provide an answer where to insert the optional hyphen, which has to be removed on undividing.

Hyp-7.3 Another slightly different type of formula is the structural linear formula of which $NaOCH_3$ is an example. Again, it is unlikely that such formulae will have to be divided at the end of a line, but when this happens to be the case the same principle as above should be used.

Hyp-7.4 Linear formulae can also show single, double and triple bonds like for instance in $CH_2=CH_2\text{-}CONH_2$ (acrylamide) or $CH_3\text{-}[CH_2]_8=[CH_2]_8\text{-}COOH$ (oleic acid). When such formulae have to be divided at the end of a line, they can be divided at any indicated single bond. To this end the original bond is shifted to the next line by replacing it by a non-breaking hyphen and inserting an optional hyphen in front of this non-breaking hyphen, thereby forming the bis-hyphen sequence (Hyp-4.1.2). On undividing, the inserted hyphen is deleted.

Hyp-7.5 A common type of linear formula deals with biomacromolecules such as proteins. Although two linear proteins can be linked by cystine bridges as in insulin, the linear chains themselves can be represented in print as a series of amino acid residues, like the Gly-Pro-Pro repeat unit in collagen. This example shows that hyphens are dispersed at regular intervals in these amino-acid sequences and so are well suited for an end-of-line division. The IUPAC-IUB document on amino acids and peptides [2] recommends that in such cases the hyphen is repeated at the beginning of the next line, which recommendation has also been adopted in the present recommendations as the basis of the hyphenation rules. Example:

<div align="center">

Ala-Ser-Tyr-Phe-Ser-
-Gly-Pro-Gly-Tyr-Arg

</div>

Just as in systematic chemical names, this repeated hyphen has to be removed on undividing.

Hyp-7.6 If the one-letter system for describing proteins or peptides is used, the letters representing the amino acids are presented in groups of ten that are separated by spaces (*e.g.* …..GTPQDRRLRL ECHETRPLRG RCGCGERRVP …..). There is no need to divide these groups at the end of a line. This has the additional advantage that in a text column, the groups are vertically aligned if the strings are very long.

Hyp-7.7 The last linear formula to be discussed is used for sequences of nucleotides, denoted by single capitals notations, *e.g.* by using a string of one-letter symbols, like d(AGAGCTAGCTCT). Dividing strings of this kind should be avoided wherever possible (*i.e.* they should appear on one line), but if absolutely necessary, the rules on numbers of characters on either line should be obeyed. An optional hyphen should be inserted and since it is a hyphen at the end of a line, it should be deleted on undividing,

Hyp-8 Annex

The typesetting software used by the various publishing houses varies. Some software packages may differentiate between differently encoded hyphens whereas other software may automatically remove such

encodings during import of a submitted manuscript file into the typesetting software. At worst, optional hyphens may be transformed into obligatory (regular) hyphens; otherwise, information provided by the author may simply be lost.

Both authors and publishers can contribute to an improvement of this situation. The following sections will describe the roles of both parties.

Hyp-8.1 Role of the author

Any author should realize that the end-of-line hyphenation in the manuscript will lead to problems in typesetting because in a typeset manuscript the line breaks are almost sure to be at different positions; this then results in excess hyphens in the middle of a line. Since manually inserted end-of-line hyphens cannot be easily recognized as such and can therefore not be removed automatically, authors are advised to refrain from inserting a normal hyphen at the end of a line, but rather to insert optional hyphens if a division is felt to be desirable.

For ease of reference, a table listing the various punctuation symbols used in this document summarizes their names, character codes, and shortcut keys.

Name	Character code	Shortcut keys
Regular hyphen	2010, Alt + X	Keyboard, or Alt + 45
Non-breaking hyphen	2011, Alt + X	Ctrl + Shift + _
Optional hyphen	00AC, Alt + X	Ctrl + -
Non-breaking space	00A0, Alt + X	Ctrl + Shift + Space
No-width non break	FEFF, Alt + X	Special characters → No-width non break
Minus sign	2012, Alt + X	
En-dash	2013, Alt + X	Ctrl + Num − [1]
Em-dash	2014, Alt + X	Alt + Ctrl + Num - [1]
Solidus	002F, Alt + X	Keyboard, or Alt + 47
Center dot	00B7, Alt + X	Alt + 0183
Colon	003A, Alt + X	Keyboard, or Alt + 58

[1]'Num −' indicates that the minus key on the numerical pad must be used.

If an author wishes to retain some control of precisely where names and formulae are divided during typesetting of a manuscript, then the following steps are advised. We acknowledge that many authors may be reluctant to expend this effort:

1. During writing, the author is advised to use non-breaking hyphens for any hyphen that may occur in the chemical name;
2. When preparing the manuscript for typesetting, the author could/should insert optional hyphens in front of non-breaking hyphens if a division is allowed at that point;
3. The author could/should surround any en dash or em dash that should not act as end-of-line divisions by no-width non breaks;
4. The author could/should also insert non-breaking spaces where needed.
5. In a later stage of manuscript preparation, the author may decide to look for systematic names with long stretches of non-interrupted characters, whereby the non-breaking hyphens that have been introduced before are counted as characters. The author may then introduce optional hyphens by using the division rules listed above but **when in doubt should refrain from inserting an optional hyphen**.

When inserting optional hyphens, the author should use his/her discretion in applying the recommended division rules. **Readability and chemical sense should prevail**. Therefore, when consulting a desk dictionary, *e.g.* [3–5], it should be kept in mind that in these dictionaries, chemical sense is often disregarded.

Because both typesetting and word processing software may store words in their databases and may add new words when they arise, the need to specify what type of hyphen should be used will eventually decrease.

Until that time, applying the above rules will expedite proofreading and reduce the need of proofreading for a second time.

Hyp-8.2 Role of the publisher/typesetter

The publisher should first of all incorporate the above recommendations in the Author Guidelines or Instructions for Authors and develop appropriate dictionaries of relevant examples. Also embedding the rules in the journal templates would be very helpful, although in some cases this may be challenging to implement.

Secondly, publishers should ensure that their typesetting software incorporates these recommendations and recognises the various types of hyphens used by the authors. Authors could then, in an ideal case, perhaps be advised to neither insert any manual line breaks, nor any encodings for specific hyphens. An even better solution would result if the software could also recognise and act on author intentions, as revealed by their deliberate insertion of non-breaking hyphens and any other specifications regarding hyphenation. Ideally the software should also flag situations where the author may have specified an incorrect or inappropriate hyphenation.

"This manuscript was prepared in the framework of IUPAC project 2014-003-2-800"

Membership of sponsoring bodies

Membership of the Division Committee of the IUPAC Division of Chemical Nomenclature and Structure Representation for the period 2018 – 2019 is as follows: **President**: A. T. Hutton (South Africa); **Past President**: K.-H. Hellwich (Germany); **Secretary**: R. S. Laitinen (Finland); **Titular Members**: M. A. Beckett (UK); E. C. Constable (Switzerland); T. Damhus (Denmark); R. T. Macaluso (USA); E. Nordlander (Sweden); A. P. Rauter (Portugal); M. M. Rogers (USA); **Associate Members**: E. Mansfield (USA); J. Nagy (Hungary); M. A. Strausbaugh (USA); K. T. Taylor (USA); C. A. Tovee (UK); J. Vohlídal (Czech Republic); **National Representatives**: F. Aricò (Italy); N. Burford (Canada); A. M. da Costa Ferreira (Brazil); S. Erdem (Turkey); S. Koo (Korea); R. Kruszyński (Poland); L. Meesuk (Thailand); M. A. Petrova (Bulgaria); E. Szabó (Slovakia); A. Yerin (Russia); **Ex Officio**: R. M. Hartshorn (New Zealand); G. P. Moss (UK).

Membership of the Division Committee of the IUPAC Division of Inorganic Chemistry for the period 2018 – 2019 is as follows: **President**: L. Öhrström (Sweden); **Vice-President:** J. García-Martínez; **Past President**: J. Reedijk (the Netherlands); **Secretary**: M. Leskelä (Finland); **Titular Members**: L. Armelao (Italy), X. Zhu (China); P. Karen (Norway), R. Macaluso (USA); M. Hasegawa (Japan), M. Drabik (Slovakia); **Associate Members:** A. Powell (Germany), L. Meesuk, (Thailand), F. Abdul Aziz, (Malaysia) L. Colón, (Puerto Rico); T. Walczyk (Switzerland), D. Rabinovich, (USA); **National Representatives**: J. Galamba Correia (Portugal). S. Kalmykov (Russian Federation), P. Knaut, (France), K. Yoon (South Korea), M. Diop (Senegal), G.J. Leigh (UK), K. Sakai (Japan), N. Trendafilova (Bulgaria), V. Stilinović (Croatia), O. Metin (Turkey), M.H. Lim (South Korea), P. Kulesza (Poland).

References

[1]　J. He, J. Chen, K.-H. Hellwich, M. Hess, K. Horie, R. G. Jones, J. Kahovec, T. Kitayama, P. Kratochvíl, S. V. Meille, I. Mita, C. dos Santos, M. Vert, J. Vohlídal. *Pure Appl. Chem.* **86**, 1003 (2014).
[2]　IUPAC, IUB. *Pure Appl. Chem*. **56**, 595 (1984).
[3]　*Merriam-Webster's New Collegiate® Dictionary*, Merriam-Webster Inc., Springfield, MA, US (2014).
[4]　*New Oxford Spelling Dictionary*, M. Waite (Ed.), xii + 596 pp., Oxford University Press, Oxford, UK (2014).

[5] *Dictionary of Contemporary English*, P. Procter (Ed.), Longman Group Ltd., UK (1978).

[6] R. G. Jones, T. Kitayama, K.-H. Hellwich, M. Hess, A. D. Jenkins, J. Kahovec, P. Kratochvíl, I. Mita, W. Mormann, C. K. Ober, S. Penczek, R. F. T. Stepto, K. Thurlow, J. Vohlídal, E. S. Wilks. *Pure Appl. Chem.* **88**, 1073 (2016).

[7] IUPAC. *Compendium of Polymer Terminology and Nomenclature (IUPAC Recommendations 2008)* (the 'Purple Book'), R. G. Jones, J. Kahovec, R. Stepto, E. S. Wilks, M. Hess, T. Kitayama, W. V. Metanomski (Eds.), xix + 443 pp., RSC Publishing, Cambridge, UK (2008); PDF on-line version: https://iupac.org/project/2002-048-1-400.

[8] IUPAC, *Nomenclature of Organic Chemistry, IUPAC Recommendations and Preferred Names 2013* (the 'Blue Book'), H. A. Favre, W. H. Powell (Eds.), xliii + 1568 pp., RSC Publishing, Cambridge, UK (2014); a list of errata can be found on the web at the address https://www.qmul.ac.uk/sbcs/iupac/bibliog/BBerrors.html.

[9] IUPAC *Nomenclature of Inorganic Chemistry (IUPAC Recommendations 2005)* (the 'Red Book'), N. G. Connelly, T. Damhus, R. M. Hartshorn, A. T. Hutton (Eds.), xiiI + 366 pp., RSC Publishing, Cambridge, UK (2005).

[10] M. Barón, R. B. Fox, A. Fradet, K.-H. Hellwich, P. Hodge, K. Horie, J. Kahovec, P. Kubisa, E. Maréchal, W. Mormann, R. F. T. Stepto, J. L. Schultz, E. S. Wilks. *Pure Appl. Chem.* **80**, 201 (2008).

[11] G. P. Moss. *Pure Appl. Chem.* **68**, 2193 (1996).

[12] W. Mormann, K.-H. Hellwich, J. Chen, E. S. Wilks. *Pure Appl. Chem.* **89**, 1695 (2017).

[13] *Nomenklatur der Anorganischen Chemie, Deutsche Fassung der Empfehlungen 1990*, W. Liebscher (Ed.), VCH, Weinheim (1994).

[14] Naming and Indexing of Chemical Substances for Chemical Abstracts™, Chemical Abstracts Services, ACS, Columbus, USA (2007). https://www.cas.org/sites/default/files/documents/indexguideapp.pdf.

[15] IUPAC, *Principles of Chemical Nomenclature, A guide to IUPAC Recommendations, 2011 Edition*, G. J. Leigh (Ed.), xii + 257 pp. RSC Publishing, Cambridge, UK (2011).

[16] British Approved Names; BAN Book, ISBN: 9780113230334 British Pharmacopoeia, Norwich, UK (2017).

[17] *The ACS Style Guide*, A. M. Coghill, L. R. Garson (Eds.), American Chemical Society, Washington, DC, USA (2006), ISBN13: 9780841239999.