Pure Appl. Chem., Vol. 85, No. 1, pp. 307–313, 2013.
http://dx.doi.org/10.1351/PAC-REP-12-07-12

Rules for abbreviation of protecting groups (IUPAC Technical Report)*

Margaret A. Brimble[1,‡], David StC. Black[2], Richard Hartshorn[3], Amelia P. Rauter[4], Chin-Kang Sha[5], and Leiv K. Sydnes[6]

[1]School of Chemical Sciences, The University of Auckland, New Zealand; [2]School of Chemistry, The University of New South Wales, Australia; [3]Department of Chemistry, Canterbury University, Christchurch, New Zealand; [4]Departamento de Química e Bioquímica Faculdade de Ciências da Universidade de Lisboa, Lisboa, Portugal; [5]Department of Chemistry, National Tsing Hua University, Taiwan; [6]Department of Chemistry, University of Bergen, Norway

Abstract: With the goal of presenting information in the scientific literature clearly and unambiguously, a set of rules for the abbreviation of protecting groups has been developed. It is based on principles designed to be as descriptive and systematic as possible, but also sufficiently pragmatic and flexible so as to accommodate the most important current abbreviations.

Keywords: chemical synthesis; IUPAC Organic and Biomolecular Chemistry Division; organic synthesis; protecting groups; synthesis.

INTRODUCTION

Twenty years ago the number of abbreviations and acronyms in the chemical literature was reasonable. However, in recent years the number has increased so rapidly that their introduction in text and schemes does not necessarily simplify the reading and increase the understanding of scientific papers. In order to remedy this situation, a project was initiated by the Subcommittee on Organic Synthesis, with the aim of presenting a clear and unambiguous set of rules for the abbreviation of protecting groups. Any abbreviations to be taken for granted should therefore be limited, and those abbreviations not taken for granted should be defined clearly at least once in each paper. At the outset, it was realised that the goal was difficult, given the existence of numerous well-known abbreviations based on totally different principles or no principles at all. Progress has consequently been rather tortuous, but as time went by certain fundamental points of view have become clear. First, there is no desire at all to rewrite all abbreviations so that they conform to a new set of guidelines, and a pragmatic approach has been developed. Second, the project should concentrate on providing guidelines for future abbreviations, rather than trying to change an existing culture. Third, there is no point in trying to specify a completely uniform set of rules, but it should be accepted that there will always be some exceptions. Fourth, emphasis should be on commonly used protecting groups, and the project should not be distracted by the more esoteric and complex examples. Such examples should be clarified in scientific papers as they are used.

*Sponsoring body: IUPAC Organic and Biomolecular Chemistry Division: see more details on p. 313.
‡Corresponding author

DISCUSSION

Many issues have been considered during the course of this project, and the most important are considered below, to provide background leading to the Recommendations that follow.

Consideration was given to whether or not protecting groups for all compound classes should be included. However, the conclusion was reached that any rules for the abbreviation of protecting groups should be independent of the functional group being protected.

Existing abbreviations have been constructed in some kind of rational way, seemingly on the basis of several principles. Abbreviations are used in other fields, e.g., the abbreviation of ligands in inorganic chemistry (e.g., see *Nomenclature of Inorganic Chemistry (IUPAC Recommendations 2009)*, the "Red Book" [1], and the table of abbreviations for ligands that are presented there for use in formulae), hence it is important to avoid the use of abbreviations that could be misinterpreted. The latter abbreviations are mostly only lowercase, unless they are well-established, hence abbreviations that only use lowercase should be avoided. Use of L, LH, or related things should also be avoided as they are often used in a generic way in inorganic chemistry. Use of Al and V (for allyl and vinyl, respectively, vide infra) should also be avoided to avoid confusion with symbols for elements. This is against IUPAC nomenclature principles, which state that names and abbreviations should not be ambiguous and the same abbreviation for an element and an organic group will never be accepted.

For the purposes of illustration a few selected examples are provided below with short comments to draw attention to inconsistencies that need to be dealt with.

In some cases, the same protecting group is written in several ways, e.g., with uppercase letters or a combination of lower- and uppercase letters (BOC and Boc for *tert*-butoxycarbonyl, Fmoc and FmOC for 9-fluorenylmethoxycarbonyl). It was concluded that in general, protecting group subunits which are abbreviated with only one letter, should be abbreviated with uppercase letters.

On the basis of compilations of common protecting groups, it is obvious that their abbreviations have been based on different combinations of the individual entities. Usually these entities are compiled in their order of appearance as one moves from the end of the protecting group towards the point of attachment to the group being protected. For instance, FmOC is used for 9-fluorenylmethoxycarbonyl. In some other cases this rule is not followed; thus, the abbreviation most frequently used for the benzyloxycarbonyl group is Cbz and not BzC.

Bn and Bz are the abbreviations now most commonly used for benzyl and benzoyl, respectively, although Bz was used earlier for benzyl as well as Bzl. It is therefore unfortunate that benzyloxycarbonyl is abbreviated with Bz and not Bn. Another example is BOM, which represents the benzyloxymethyl group; here benzyl is designated B instead of Bn. It seems reasonable to retain Bn and Bz as the abbreviations for benzyl and benzoyl, respectively.

Generally, the alkoxy groups represent a problem. In the BOC group, the oxy function is shown by the O. The same is the case with the MOM (methoxymethyl) group. However, in Cbz there is no indication of the presence of the oxygen atom, and the same is the case for the MEM (2-methoxyethoxymethyl) group. A reasonable solution to this dilemma would be to regard most alkoxy groups as a combination of an alkyl group, with a defined abbreviation, and one oxygen atom, which is abbreviated O (uppercase). For instance, according to this proposal, methoxymethyl could be abbreviated as MOM whereas 2-methoxyethoxymethyl, which is well established as MEM (and not MOEOM), could be treated as an exception and kept as MEM.

In BOC, the B represents *tert*-butyl, i.e., the tertiary nature of the butyl group is not indicated. However, in TBDMS this fact is visible due to the T, which means that *tert*-butyl is abbreviated TB. Due to the fact that tertiary almost without exception is indicated by either *t* or *tert*, a reasonable alternative would be tB. However, since TB is already in use, it might seem preferable to use TB as the abbreviation for *tert*-butyl. In most cases, the designator for tertiary is unnecessary, as *n*-butyl and *sec*-butyl do not appear in common protecting groups. Furthermore, T is the common abbreviation for tri.

Therefore, B seems the most reasonable abbreviation for butyl. However, in view of the widespread use of B in protecting-group abbreviations, it is considered desirable not to restrict its use to butyl, but to allow discretion to apply to its application.

Another issue is the use of S or Si for silicon. In the early stages of this project, it was proposed to use Si for silicon, but this was strongly opposed by a number of prominent chemists, mainly due to the fact that TMS has for a long time been used as the abbreviation for trimethylsilyl. The argument to switch from S to Si for silicon was that S should be used for sulfur. The TMS argument is convincing, and it is proposed that S is retained as the abbreviation for silicon. The role of sulfur in protecting groups is essentially confined to sulfonyl, which can be dealt with as a secondary unit and abbreviated as s (lowercase).

Yet another major problem is the meaning of the letter P. This is illustrated by the following names and abbreviations: *p*-nitrobenzyl = PNB, *p*-nitrophenyl = PNP, and 2-(2'-pyridyl)ethyl = Pet. From these examples it is obvious that P is used to indicate at least *para*, phenyl, and pyridyl, which is unfortunate. The situation is even more confusing when the same letter is used for propyl. The best solution seems to be the use of p for *para*, Ph for phenyl, and Py for pyridyl, but not to designate P for propyl. As in the case of butyl discussed above, the use of P is so widespread that discretion should be applied to its application.

Most of the discussion so far has very little or no relevance to the abbreviation of cyclic protecting groups, that is, protecting groups incorporating in a ring structure one or several atoms of the functional group being protected. Survey of the literature reveals that very few protecting groups of this kind are abbreviated, but are described either by their full names or as derivatives of the parent compounds (e.g., "2-alkyl-1,3-dithiane" or a "1,3-dithiane derivative"). On this basis it is proposed that such protecting groups should not be abbreviated, but should be clarified as they are introduced.

On the basis of the above discussion the following rules are recommended.

RULES

R-1: Construction of abbreviations for noncyclic protecting group

R-1.1 Subunit division

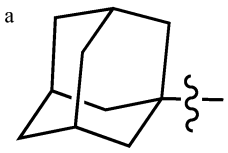
The protecting groups are divided into two types of fundamental subunits. Primary subunits are those connected to one neighbour by one bond. Secondary subunits are those connected to two neighbours by two bonds (one bond each). The example of silicon represents a quaternary subunit. Those subunits available for selection are shown in Table 1.

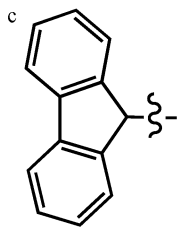
R-1.2 Subunit abbreviation

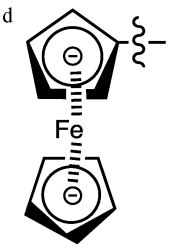
Each subunit is given an abbreviation, consisting of either one uppercase letter, or two letters, of which the first is uppercase and the second is either lower- or uppercase. If a larger subunit is available then it should be chosen so as to shorten the final abbreviation. The list is compiled in Table 1.

Table 1 Protecting-group subunits and their abbreviations.

Chemical formula	Name	Abbreviation
Primary subunits		
$CH_3CO–$	acetyl	Ac
[a]	1-adamantyl	Ad
$CH_2=CHCH_2–$	2-propenyl (allyl)	All
$PhCH_2–$	phenylmethyl (benzyl)	Bn
PhCO–	benzoyl	Bz
$CH_3CH_2–$	ethyl	E[b]
[c]	9-fluorenyl	F
[d]	ferrocenyl	Fc
H–	hydro	H
$(CH_3)_2CH–$	1-methylethyl (isopropyl)	IP
$CH_3–$	methyl	M[b]
$CH_3SO_2–$	mesyl = methanesulfonyl	Ms
$NO_2–$	nitro	N
$C_6H_5–$	phenyl	Ph
$(CH_3)_3CCO–$	2,2-dimethylpropanoyl (pivaloyl)	Piv
$C_5H_5N–$	pyridyl	Py
$CF_3SO_2–$	trifluoromethanesulfonyl	Tf
$Ph_3C–$	triphenylmethyl (trityl)	Tr
$4\text{-}CH_3C_6H_4SO_2–$	*p*-toluenesulfonyl (tosyl)	Ts
$CH_2=CH–$	ethenyl (vinyl)	Vi
Secondary subunits		
$–CH_2–$	methane-1,1-diyl (methylene)	M[b]
$–CH_2CH_2–$	ethane-1,2-diyl (ethylene)	E[b]
–O–	oxy	O
–OCO–	oxycarbonyl	oc
$–SO_2–$	sulfonyl	s
Quaternary subunits		
$–\overset{\displaystyle |}{\underset{\displaystyle |}{Si}}–$	silyl	S





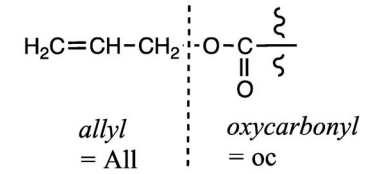
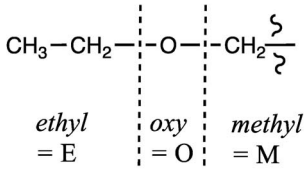


[b]Both ethyl and ethane-1,2-diyl (ethylene) are designated as E, and both methyl and methane-1,1-diyl (methylene) are designated as M.

R-1.3 Group abbreviation

The group abbreviation is generated by combining the abbreviation for each subunit as one moves from the end of the protecting group towards the group being protected. The subunit abbreviations are not separated by hyphens or spaces.

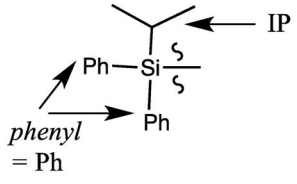
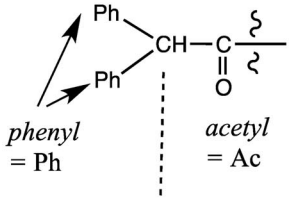
Examples:

Protecting Group	*Name*	*Abbreviation*
	allyloxycarbonyl	Alloc
	ethoxymethyl	EOM

R-1.4 Numerical prefixes

If there are several identical groups, their multiplicity is given by the numerical prefixes D for two (Di) and T for three (Tri). There is no hyphen or space between a numerical prefix and the subunit abbreviation, to which it relates. There is no need to indicate the number if there is only one group.

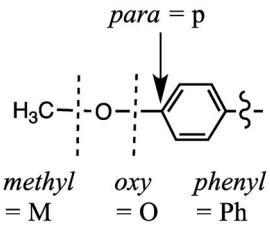
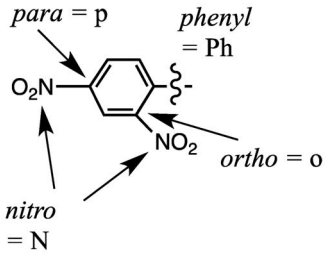
Examples:

Protecting Group	*Name*	*Abbreviation*
	diphenylisopropylsilyl	DPhIPS
	diphenylacetyl	DPhAc

R-1.5 Locants

The location of a substituent on a phenyl group is specified according to its position as *ortho*, *meta*, or *para* by o, m, or p, respectively (non-italicised). There is no hyphen or space between the locant and the related subunit abbreviation. In some cases, such as *p*-toluenesulfonyl, the locant is subsumed into the designated abbreviation.

Examples:

Protecting Group	*Name*	*Abbreviation*
	*p*-methoxyphenyl	pMOPh
	2,4-dinitrophenyl	opDNP

R-2: Abbreviation of cyclic protecting groups

Cyclic protecting groups, that is, protecting groups incorporating in a ring structure one or several atoms of the functional group being protected, are not abbreviated according to the set of rules proposed for noncyclic protecting groups (R-1).

R-3: Recommended abbreviations deviating from the rules

Certain well-established protecting-group abbreviations, which formally violate the rules above, are nevertheless accepted. Those recommended for approval are listed in Table 2.

Table 2 Accepted abbreviations for protecting groups deviating from the rules.

Name	Abbreviation
tert-Butyldimethylsilyl	TBDMS
2-(Diphenylphosphono)ethyl	DDPE
2-Methoxyethoxymethyl	MEM
Methylthiomethyl	MTM
2,2,2-Trichloroethyl	TCE
Tetrahydropyranyl	THP

R-4: Abbreviations deviating from the rules and not recommended for use

Any protecting group that deviates from the rules (R-1 and R-2) and is not included in the exceptions recommended for use (R-3) should be clearly defined in a note or section called "Abbreviations" preceding the "Experimental" part of scientific papers.

ABBREVIATIONS

The following abbreviations, deviating from the IUPAC rules, are applied:

Bus = *tert*-butylsulfonyl; CDA = cyclohexane-1,2-diacetyl; STABASE = 2,2,5,5-tetramethyl-1-aza-2,5-disilacyclopentane; TDE = (2,2,2-trifluoro-1,1-diphenyl)ethyl.

MEMBERSHIP OF SPONSORING BODIES

Membership of the IUPAC Organic and Biomolecular Chemistry Division for the period 2012–2013 is as follows:

***President*:** K. N. Ganesh (India); ***Vice President*:** M. J. Garson (New Zealand); ***Secretary*:** A. Griesbeck (Germany); ***Past President*:** G. J. Koomen (Netherlands); ***Titular Members*:** G. M. Blackburn (UK); A. Brandi (Italy); M. A. Brimble (New Zealand); T. Carell (Germany); B. Han (China); F. Nicotra (Italy); ***Associate Members*:** M. Cesa (USA); V. Dimitrov (Bulgaria); H. Jacobs (Jamaica); S. H. Kang (Korea); M. Orfanopolos (Greece); D. Sladi (Serbia); ***National Representatives*:** A. Al-Aboudi (Jordan); P. Crowley (Ireland); L. Diaz (Brazil); E. Innocent (Tanzania); K. Khalid (Pakistan); T. Krygowski (Poland); C.-C. Liao (Taiwan); D. Mizrahi (Israel); J. Streith (France); T. Vilaivan (Thailand).

Membership of the Subcommittee on Organic Synthesis is as follows:

***Chair*:** M. A. Brimble (New Zealand); ***Members*:** D. StC. Black (Australia); E. M. Carreira (Switzerland); P. Chiu (China); M. Isobe (Japan); E. Juaristi (Mexico); S. H. Kang (Korea); J. M. Khanna (India); G. J. Koomen (Netherlands); E. P. Kundig (Switzerland); S. V. Ley (UK); C.-C. Liao (Taiwan); G.-Q. Lin (China); M. Makosza (Poland); F. McDonald (USA); L. McElwee-White (USA); F. Nicotra (Italy); K. Oshima (Japan); L. Sydnes (Norway); H. Yamamoto (USA).

REFERENCE

1. IUPAC. *Nomenclature of Inorganic Chemistry, IUPAC Recommendations 2005* (the "Red Book"). Prepared for publication by N. Connelly, T. Damhus, R. M. Harshorn, RSC Publishing, Cambridge, UK (2005). ISBN 0-85404-438-8