Journal of Peptide Science
J. Peptide Sci. 2006; **12**: 1–12
Published online 24 October 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/psc.725

JOURNAL OF PEPTIDE SCIENCE

Editorial

Abbreviations and symbols in peptide science: a revised guide and commentary

Abstract: The abbreviations and symbols used in Peptide Science are surveyed, with comment and recommendations. Copyright © 2005 European Peptide Society and John Wiley & Sons, Ltd.

Keywords: amino acids, abbreviations for; peptide science, abbreviations in

CONTENTS

INTRODUCTION

Abbreviations, acronyms and symbolic representations are very much part of the language of peptide science – in oral communication as much as in its literature. They are not only a convenience, either – they enable the necessary but distracting complexities of long chemical names and technical terms to be pushed into the background, so the wood can be seen among the trees. Many of the abbreviations listed here are so much in use that they need no explanation. The main purpose of this Editorial, which updates and extends two previous guides [1,2], is to identify the vocabulary of the lingua franca, to outline its grammatical conventions, and to free authors from the tiresome obligation of defining commonplace abbreviations in every paper. Those in the lists that follow should be used without explanation. All other abbreviations and usage should be carefully defined in a footnote at the outset of the paper. The explanation of abbreviations as they appear (in brackets like this) interrupts the flow of the text

(just as this does); the excessive use of brackets intrudes and may distract, especially in convoluted sentences (I am sure the reader will see what I mean by now).

Previously published usage may be followed, unless it is manifestly clumsy, and even if it is, it should be followed, with due explanation, if there is a long-established convention in a series of related previous papers. Readers will see from past contributions to this Journal that much variation has been permitted and editorial insistence has only been exercised occasionally. This is because it has seemed unreasonable to delay the publication of scientifically sound papers simply to get conformity, and authors have been vigorous in defending their eccentricities. It does not mean that I approve.

If an abbreviation is employed at all, it should be used consistently. It is confusing to the naïve reader if there is alternation between e.g. dimethylformamide and DMF, bradykinin and BK, and so on. The use of two abbreviations for the same thing in the same paper, a surprisingly common defect in drafts evidently produced by the scissors and paste approach, is of course completely unacceptable.

Abbreviations should only be used in titles and abstracts if they are very familiar ones.

Where it is necessary to devise new abbreviations and symbols, the general principles behind established examples should be followed. Thus, new amino acid symbols should be of form Abc, with due thought for possible ambiguities (Dap might be obvious for diaminopropionic acid, for example, but what about diaminopimelic acid?). Further, authors poised to be inventive should ask themselves whether they really need that novel abbreviation. It is a question of frequency of appearance. If there is a term which is repeated many times, it is helpful to use a thoughtfully composed abbreviation for it. By the time the reader is a short distance into the paper, (s)he will have taken it on board, and it will not interrupt the flow of thought. But if the abbreviation only appears a few times, its appearances may slow things down.

Several dictionaries of scientific abbreviations and acronyms are available. The most comprehensive was published by our publishers in 2001 [3]: every scientific library should have a copy. Some idea of the scope for confusion can be obtained by dipping into it. ESI, for example, has 15 meanings listed.

A certain amount of common sense is called for. It is understandable, for example, that in laboratory conversation the trisyllabic TFM (*tee-eff-em*) drops off the tongue more easily than the six syllables of trifluoromethyl (*try-flew-or-oh-me-thyle*), but to use TFM or Tfm as an 'abbreviation' for CF_3 in print is an absurdity. Exchange of three characters for three is a poor rate of exchange.

Where alternatives are indicated below, the first is preferred.

AMINO ACID SYMBOLS

Proteinogenic Amino Acids

Ala	Alanine A
Arg	Arginine R
Asp	Aspartic acid D
Asx	Asn *or* Asp
Cys	Cysteine C
Gln	Glutamine Q
Glu	Glutamic acid E
Glx	Gln *or* Glu
Gly	Glycine G
His	Histidine H
Ile	Isoleucine I
Leu	Leucine L
Lys	Lysine K
Asn	Asparagine N
Met	Methionine M
Phe	Phenylalanine F
Pro	Proline P
Ser	Serine S
Thr	Threonine T
Trp	Tryptophan W
Tyr	Tyrosine Y
Val	Valine V

Other Important Amino Acids

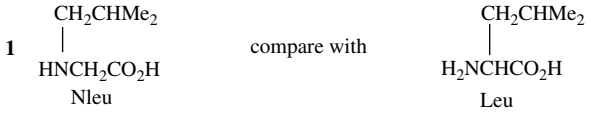
Aad	α-Aminoadipic acid
βAad	β-Aminoadipic acid
Abu	α-Aminobutyric acid
Aib	α-Aminoisobutyric acid; α-methylalanine
βAla	β-Alanine; 3-aminopropionic acid (avoid Bal)
Asu	α-Aminosuberic acid
Aze	Azetidine-2-carboxylic acid
Cha	β-Cyclohexylalanine
Cit	Citrulline; 2-amino-5-ureidovaleric acid
Gla	γ-Carboxyglutamic acid
Glp	Pyroglutamic acid; 5-oxoproline (also pGlu)
Hyl	δ-Hydroxylysine
Hyp	4-Hydroxyproline
*a*lle	*allo*-Isoleucine; 2*S*, 3*R* in the L-series
Iva	isovaline; α-ethylalanine. Care! Iva has also been used for isovaleryl; (αEt)Ala is safer.
Lan	Lanthionine; *S*-(2-amino-2-carboxyethyl)cysteine
Nle	Norleucine; α-aminocaproic acid
Orn	Ornithine; 2,5-diaminopentanoic acid
Pen	Penicillamine; β-mercaptovaline, ββ-dimethylcysteine
Phg	Phenylglycine; 2-aminophenylacetic acid
Pip	Pipecolic acid; piperidine-2-carboxylic acid
Sar	Sarcosine; *N*-methylglycine
Sec	Selenocysteine
Sta	Statine; (3*S*, 4*S*)-4-amino-3-hydroxy-6-methyl-heptanoic acid
Thi	β-Thienylalanine
Tic	1,2,3,4-Tetrahydroisoquinoline-3-carboxylic acid
*a*Thr	*allo*-Threonine; 2*S*, 3*S* in the L-series
Thz	Thiazolidine-4-carboxylic acid, thiaproline
Xaa	Unknown or unspecified (also Aaa)

Modified Proteinogenic Amino Acids

See section 'Peptide bond replacement' for reduced α-amino acid residues, α-amino thioacids and α-hydroxy acids.

***N*-alkyl-α-amino acids; α-alkyl-α-amino acids.** MeAla *N*-Methylalanine (MeVal *N*-methylvaline, etc.). This style should not be used for α-alkyl amino acids, for which either a unique symbol (such as Aib for α-methylalanine, or, with great caution, Iva for isovaline) should be used, or the position of the alkyl group should be made explicit as in (αMe)Tyr for α-methyltyrosine.

The *N*-alkylglycines analogues of proteinogenic amino acids in which side chains have been shifted sideways are a special case. The leucine analogue **1** could be rendered Bu^{i}Gly by analogy with the previous paragraph, but the style Nleu has been used.



The *N*-alkyl-α-amino acid class also includes the Hmb derivatives used in difficult sequence synthesis. They have been rendered in various ways, of which (if on-the-line format is required) the style, e.g. Fmoc-HmbGly-OH is recommended.

α-aza-α-amino acids. Replacement of the α-carbon by a nitrogen has been employed in peptide analogue design. The hydrazino acids thus derived are termed e.g. azaglycine, $H_2NNH_2CO_2H$, and can be abbreviated AzaGly, and so on. However, great care is needed here, as there are analogues of e.g. leucine and tyrosine with nitrogen-for-carbon replacements in the side chain which are known as azaleucine and azatyrosine, and which have been abbreviated

AzaLeu and AzaTyr respectively. As a further twist, azaserine is the name favored, by the Merck Index at least, for the antibiotic *O*-diazoacetyl-L-serine.

***β*-amino alcohols derived from *α*-amino acids.** Pheol Phenylalaninol, and so on. The style Phe(ol) is confusing, as indications in brackets after the three-letter symbol should refer to side chain modification; Phe(OH) is positively wrong, as it implies hydroxylation in the side chain. Lol for leucinol, Vol for valinol and so on are discouraged.

***α*-amino aldehydes derived from *α*-amino acids.** PheH Phenylalaninal, and so on. The style Phe-al is discouraged.

***αβ*-dehydro-*α*-amino acids.** Dha Dehydroalanine (also ΔAla), and so on. Strict IUPAC nomenclature requires these α-amino acids to be called didehydro-α-amino acids, but this has not generally been followed.

***Desamino-α*-amino acids.** Great care is necessary in symbolic designations of desamino (deamino) derivatives. The prefixed letter d is favored in the subfield of neurohypophyseal hormone analogues. Although well understood by those working in that subfield, this is the most unfortunate convention, because outsiders may naturally think the d, especially if spoken, has stereochemical significance. The use of e.g. dPen for desamino-penicillamine is thus discouraged; in this example, the potential for confusion is exacerbated by the fact that penicillamine and desamino-penicillamine would make equal sense in the *N*-terminal position, and, further, the usual naturally occurring penicillamine enantiomer is of the D-configuration. A more satisfactory approach is to use a unique abbreviation. Thus, desamino-cysteine is β-mercaptopropionic acid, logically abbreviated as Mpa, enabling desamino-oxytocin to be [Mpa^1]-oxytocin. It would be better still perhaps to abandon abbreviation in favor of clarity, and use the full prefix 'desamino'. Incidentally, 'desamino' is preferred to 'deamino', as the former avoids any unintended stereochemical implication.

***Glyco-α*-amino acids.** The symbolic representation of carbohydrate structures is a complex business in itself, more so when combined with the symbolic representation of peptide structures (Some IUPAC-IUB guidance may be found in *Nomenclature of glycoproteins, glycopeptides and peptidoglycans (Recommendations 1985)*, which was (9 July 2005) at http://www.chem.qmul.ac.uk/iupac/misc/glycp.html). It is unfortunate that the style recommended for side chain attachment of carbohydrates, e.g. (GlcNAc-)Asn, *N*-acetylglucosaminylasparagine, is at odds with the preferred way of indicating other kinds of side chain substitution, with the substituent in brackets after the three-letter symbol.

***Homo-α*-amino acids.** Hph Homophenylalanine (Hse homoserine, etc.) is used. Caution is necessary over the use of the prefix 'homo' in relation to amino acid nomenclature and symbolism. When the term first became current, it was applied to analogues in which a side chain CH_2 extension had been introduced. Thus, homoserine has a side chain CH_2CH_2OH, homoarginine $CH_2CH_2CH_2NHC(=NH)NH_2$, etc. In such cases, the convention is that a new three-letter symbol for the analogue is derived from the parent, by taking H for homo and combining it with the first two characters of the parental symbol. Hence Hse, Har etc. are used. Now, however, there is a considerable literature on β-amino acids which are analogues of α-amino acids in which a CH_2 group has been inserted between the α-carbon and carboxyl group. These analogues have also been called homo analogues, and there are instances e.g. not only of 'homophenylalanine', $NH_2CH(CH_2CH_2Ph)CO_2H$, abbreviated as Hph, but also of 'homophenylalanine', $NH_2CH(CH_2Ph)CH_2CO_2H$ abbreviated as Hph. Further, members of the analogue class with CH_2 interpolated between the α-carbon and the carboxyl group of the parent α-amino acid structure have been called both 'α-homo' and 'β-homo'. Clearly great care is essential, and abbreviations for 'homo' analogues ought to be fully defined on every occasion.

The prefix 'β-homo' seems preferable for backbone extension (emphasizing as it does that the residue has become a β-amino acid residue), with abbreviated symbolism as illustrated by βHph for $NH_2CH(CH_2Ph)CH_2CO_2H$.

THE USE OF THE AMINO ACID SYMBOLS

General Remarks

The three-letter symbols should generally be used in accord with the IUPAC-IUB recommendations (The recommendations have been published in many places, and were (13 June 2005) also available with other relevant documents at http://www.chem.qmw.ac.uk/iupac/. See especially *Nomenclature and symbolism for amino acids and peptides (Recommendations 1983)*, which was reproduced at http://www.chem.qmul.ac.uk/iupac/AminoAcid/).

The three-letter symbol standing alone represents the unmodified intact amino acid, of the L-configuration unless otherwise stated (but the L-configuration may be indicated if desired for emphasis: e.g. L-Ala). The same three-letter symbol, however, also stands for the corresponding amino acid *residue*, i.e. the amino acid $H_2NCHRCO_2H$ less H_2O.

$$Xaa = H_2NCHRCO_2H \textit{ or } -NHCHRCO-$$

The symbols can thus be used to represent peptides. When nothing is shown attached to either side of the three-letter symbol, it is meant to be understood that the amino group (always understood to be on the left) or carboxyl group is unmodified, but this can be emphasized, so XaaXaa = H-XaaXaa-OH. Note, however, that indicating free termini by formulating the terminal group in full is wrong: $NH_2XaaXaaCO_2H$ implies a hydrazino group at one end and an α-keto acid derivative at the other. Representation of a free terminal carboxyl group by writing H on the right is also wrong, because that implies a terminal aldehyde.

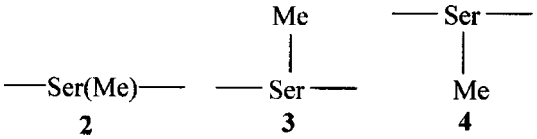
Hyphens may be used in the presentation of amino acid sequences using three-letter codes, e.g. H-Ala-Gly-Leu-OH, but are in fact superfluous in plain sequences which lack side chain or α-modifications; HAlaGlyLeuOH is sufficient, although H-AlaGlyLeu-OH may be preferred for its emphasis of terminal groups. But where there is side chain or α-modification, hyphens are helpful in making clearer what belongs to which residue, by isolating the modified residue, so e.g. HAla-Ser(Me)-LeuOH is better than HAlaSer(Me)LeuOH.

Numbered amino acid resides in a peptide chain should be referred to in discussion by the style e.g. Leu^5, not 5-Leu or Leu5, or spelt out, i.e. leucine-5.

The one-letter codes can be used in the same way as the three-letter codes, but the three-letter codes are preferred for small and medium-sized peptides. Peptides which contain a few nonproteinogenic residues embedded in long sequences of proteinogenic residues can be represented using the two codes macaronically, e.g. ..... RNDAibPQR..... The use of letters which do not appear in the regular one-letter code for proteinogenic amino acids, e.g. B, for nonproteinogenic residues is discouraged. If the one-letter codes are used in synthetic schemes, which is discouraged except for long and unwieldy molecules, remember that H can mean either a histidine residue or a hydrogen atom, and R, the conventional way of showing a generalized (especially alkyl) group, also means arginine.

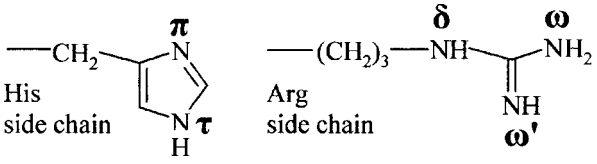
Substituted Side Chains

Side chains are understood to be unsubstituted if nothing is shown, but a substituent replacing a functional group hydrogen atom can be indicated by use of brackets or attachment by a vertical bond up or down. Thus an *O*-methylserine residue could be shown as **2**, **3**, or **4**.



Note that the oxygen atom is not shown above: it is contained in the three-letter symbol. Showing it, as in Ser(OMe), would imply that a peroxy group was present.

The position of substitution does not need to be indicated if it involves replacement of functional group hydrogen atom unless there is ambiguity, as in histidine and arginine side chains, where the Greek designators π/τ and $\delta/\omega/\omega'$, respectively, are preferred.



But if the position of substitution is not at a functional group, it does need to be indicated. This is best done with Greek characters for alkyl side chain positions and Arabic numbers for ring positions, e.g. Phe(βOH) for β-hydroxyphenylalanine alias β-phenylserine, and Phe(2Cl) for *o*-chlorophenylalanine. If no substitution position is indicated in a substituted benzene ring derivative, it will be taken as *para*, e.g. Phe(Cl) for *p*-chlorophenylalanine.

Substitutions within substituents can be indicated by the use of a hierarchy of brackets {...[...(...)...]...}; this situation is illustrated below (see section 'Substituent and protecting group symbols', Bzl) and is present in multibranched peptides (see section 'Branched peptides').

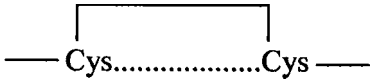
Double substitution may be indicated e.g. Arg(δ,ωZ_2) and Phe(3Me,2NO_2), but in complex cases where the side chain modification is permanent, it may be better to devise something shorter and simpler for the case in hand.

Bridged Peptides

By 'bridged' here I mean peptide derivatives in which two parts of a peptide are connected by bonds other than peptide or amide bonds; I term peptides with intramolecular connections through peptide or amide bonds 'cyclopeptides', cyclization being between α-functions unless otherwise indicated. This is an admittedly arbitrary distinction, but one that is convenient for the present purpose.

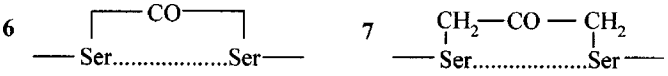
A disulfide bridge, by far the most familiar type of bridge, is indicated as in **5**.

5

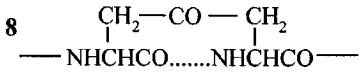


This is universally understood in the peptide and protein field, but extension of this way of indicating bridges to other types is at risk of misinterpretation, because nonspecialist organic chemists are accustomed to reading unadorned junctions of two bond lines at an angle as shorthand for carbon atoms. Thus, a carbonate bridge between two serine side chains, such as that which would be produced by cross-linking with a phosgene equivalent, might be taken as **6** if represented as in **7**, and the fuller formulation **8** is preferable.

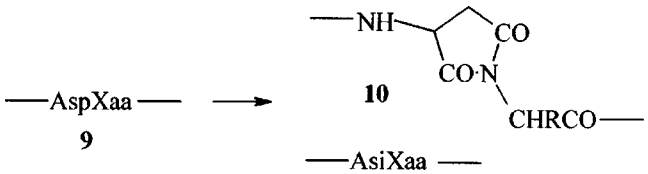
6 **7**



8



Bridges involving peptide backbone nitrogen atoms call for special care. The abbreviation Asi has been suggested [4] for the aspartimide 'residue' formed by the side chain to backbone cyclization **9** to **10**. This is a slight abuse of the way Xaa should really be used, but has the great merit of simplicity.



Branched Peptides

By the term 'branched', I mean assemblies in which amino acid residues with functional side chains which have peptide chains grown not only from the α-functions but also from the side chains. This situation can be shown in the same way as any other side chain modification, e.g. **11**, bisglycyllysine, and multiple branching involves substitutions within substituents, so use of the hierarchy of brackets {...[...(...)...]...} is appropriate. The assembly which would, in principle, be obtained by starting with an amino carrier RNH_2 and performing three cycles of coupling with BocLys(Boc)OH and Boc-removal could, thus, be presented as in **12**

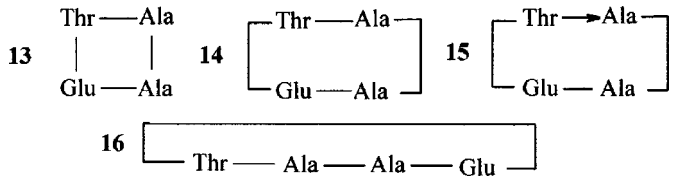
11 GlyLys(Gly) **12** Lys(Lys)Lys[Lys(Lys)]Lys{Lys[Lys(Lys)]}NHR

Cyclic Peptides

Confusions may creep in if the three-letter symbols are used thoughtlessly in representations of cyclic peptides, even simple ones. (Some guidance may be found in the 2004 IUPAC-IUB draft

prepared by GP Moss on cyclic peptide nomenclature which was (11 April 2005) available at http://www.iupac.org/reports/provisional/abstract04/moss_prs.pdf.

Consider the product of cyclizing threonylalanylalanylglutamic acid end to end. It might be thought that this compound could be economically represented as in **13**. But this is wrong, because the left-hand vertical bond implies an ester link between the two side chains, and strictly speaking, if the right-hand vertical bond means anything, it means that the two Ala α-carbons are linked by a CH_2CH_2 bridge. This objection could be circumvented by writing the structure as in **14**. But this is now ambiguous because the convention that the symbols are to be read as having the amino nitrogen to the left cannot be imposed on both lines. The direction of the peptide bond needs to be shown with an arrow pointing from CO to N, as in **15**. Actually, the simplest representation is on one line, as in **16**.



Many very complex peptidic natural products with multiple bridging and unusual features, such as vancomycin and α-amanatin, are known: some attempts to bring them within the scope of conventional abbreviation have been aired, but they are not recommended in their present stage of development.

Analogues of Bioactive Peptides

Residue-substitution analogues of bioactive peptides should be designated, e.g. the style [Gly^7]-oxytocin, or spelt out e.g. 7-glycine-oxytocin for the analogue of oxytocin in which the proline at position 7 has been replaced by glycine. Further abbreviation would be acceptable here to [Gly^7]-OT. Multiple replacements should be indicated in the style e.g. [Ser^4, Gly^7]-OT.

Extensions as in glycyloxytocin, where an additional glycine residue acylates the *N*-terminal residue of oxytocin, should be rendered Gly-OT.

Residue deletions, as in oxytocin with the leucine residue at position 8 removed so that the residue 7 in the intact hormone connects directly with the residue at position 9, should be abbreviated as des-(Leu^8)-OT. Deletion of a sequence could be given in the short form e.g. des-(7–8)-OT.

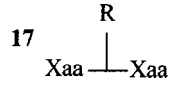
Partial but otherwise natural sequences have been alluded to in various ways; the style e.g. OT-(2–5)-peptide for the four-residue sequence which comprises positions 2 to 5 inclusive of oxytocin, is preferred. Taking this a step further, the sequence PheIleGlnAsn, which corresponds to positions 2–5 of oxytocin with phenylalanine replacing tyrosine at position 2, could be abbreviated as [Phe^2]-OT-(2–5)-peptide.

Residue interpolations, as in oxytocin with glycylglycine inserted between the proline at position 7 and leucine at position 8, could be rendered OT-(1–7)-GlyGly-(8–9)-peptide.

MODIFIED PEPTIDE BONDS

N-substitution

N-substitution with a group R can be indicated on the line as already explained in section *N*-alkyl-α-amino acids; α-alkyl-α-amino acids, but also by means of a line at 90° to the peptide bond line as shown in **17**, which is useful in the case of bridging to a peptide bond nitrogen atom.



Peptide Bond Replacement

Replacement of one or more of the C, N, or O atoms of a peptide bond between two amino acid residues Xaa-Xaa can be indicated Xaa-ψ[replacement]-Xaa, as exemplified below

$$\begin{aligned}
&\text{NHCHRCONHCHRCO} \longrightarrow \text{NHCHRCH=CHCHRCO}\\
&\qquad\qquad\qquad\qquad\qquad\quad \text{Xaa}\psi\text{[CH=CH]Xaa}\\
&\text{NHCHRCONHCHRCO} \longrightarrow \text{NHCHRCH}_2\text{CH}_2\text{ CHRCO}\\
&\qquad\qquad\qquad\qquad\qquad\quad \text{Xaa}\psi\text{[CH}_2\text{CH}_2\text{]Xaa}\\
&\text{NHCHRCONHCHRCO} \longrightarrow \text{NHCHRCH}_2\text{NHCHRCO}\\
&\qquad\qquad\qquad\qquad\qquad\quad \text{Xaa}\psi\text{[CH}_2\text{NH]Xaa}\\
&\text{NHCHRCONHCHRCO} \longrightarrow \text{NHCHRCSNHCHRCO}\\
&\qquad\qquad\qquad\qquad\qquad\quad \text{Xaa}\psi\text{[CSNH]Xaa}
\end{aligned}$$

Often, both participating amino acid residues of the dipeptide unit are modified in such operations, but there are a few important cases where only one is changed. The reduced peptide bond replacement, CH_2 NH for CO-NH, is such a case. Here, only the left-hand residue is modified. This could have been handled by devising a symbol for the reduced residue –NHCHRCH$_2$–, e.g. RedXaa, and using it e.g. –RedXaa–Xaa–; but the ψ[replacement] style has been preferred.

The case where the replacement is CS–NH for CO–NH, i.e. a thioamide link, is that of the thiopeptides. Here too, only the left-hand residue is modified, and abbreviations of the form e.g. Glyt, Leut, etc. have been used. This is acceptable in the peptide context, but note that Glyt/GlyT/glyT etc. also make multiple appearances with quite different meanings in the wider molecular biology context.

The case where the replacement is CO–O for CO–NH, i.e. an ester link, is that of the depsipeptides (peptolides). Here only the right-hand residue is modified; an α-amino acid residue is exchanged for an α-hydroxy acid residue, e.g. lactic acid for alanine, which could be simply rendered as below on the right.

$$\text{–NHCHMeCO-OCHMeCO–} = \text{–Ala}\psi\text{[COO]Ala–} = \text{–Ala-Lac–}$$

When an ester link involving a side chain is involved, there are subtle difficulties in representation [5].

SUBSTITUENT AND PROTECTING GROUP SYMBOLS

Ac	Acetyl
Acm	Acetamidomethyl
Adoc	1-Adamantyloxycarbonyl
Alloc	Allyloxycarbonyl
Boc	*t*-Butoxycarbonyl. *t*-Boc, *t*Boc and BOC are discouraged. The usages e.g. *N*(α)-BocGlyOH and N^{α}-BocGlyOH are both common in past literature, but the *N*(α)- and N^{α}- indicators are superfluous as there is nowhere else for the Boc group to be. Any group written the left of the three-letter symbol is understood to be on *N*(α). Incidentally, N^{α} is generally to be preferred to *N*(α), etc. on the grounds that no brackets are required – abbreviated formulations in peptide science already have more than enough of those.
Bom	π-Benzyloxymethyl
Bpoc	2-(4-Biphenylyl)isopropoxycarbonyl
Btm	Benzylthiomethyl
Bum	π-*t*-Butoxymethyl
Bu^{i}	*i*-Butyl
Bu^{n}	*n*-Butyl
Bu^{t}	*t*-Butyl
Bz	Benzoyl (care! confusion with benzyl is common)

Bzl	Benzyl (also Bn, discouraged); Bzl(NO_2) = 4-nitrobenzyl, etc. the substituent location only being indicated if it is not *para*, e.g. Bzl(2NO_2), *o*-nitrobenzyl. If a substituted protecting group is used on a side chain, leading to brackets within brackets, the usual hierarchy should be used, e.g. Tyr[Bzl(2NO_2)].
Cha	Cyclohexylammonium salt
Clt	2-Chlorotrityl
Dcha	Dicyclohexylammonium salt
Dde	1-(4,4-Dimethyl-2,6-dioxocyclohex-1-ylidene)ethyl
Ddz	2-(3,5-Dimethoxyphenyl)-isopropoxycarbonyl
Dnp	2,4-Dinitrophenyl
Dns	5-Dimethylaminonaphthalene-1-sulfonyl (dansyl)
Dpm	Diphenylmethyl (also Bzh, Benzhydryl)
Dpp	Diphenylphosphinyl
Et	Ethyl
Fmoc	9-Fluorenylmethoxycarbonyl
For	Formyl
Hmb	2-Hydroxyl-4-methoxybenzyl
Mbh	4, 4′-Dimethoxydiphenylmethyl, 4, 4′-Dimethoxybenzhydryl
Mbs	4-Methoxybenzenesulfonyl
Me	Methyl
Mob	4-Methoxybenzyl
Mtr	2,3,6-Trimethyl,4-methoxybenzenesulfonyl
Nps	2-Nitrophenylsulfenyl
OAll	Allyl ester
OBt	1-Benzotriazolyl ester
OcHx	Cyclohexyl ester
ONp	4-Nitrophenyl ester
OPcp	Pentachlorophenyl ester
OPfp	Pentafluorophenyl ester
OSu	Succinimido ester
OTce	2,2,2-Trichloroethyl ester
OTcp	2,4,5-Trichlorophenyl ester
Tmob	2,4,5-Trimethoxybenzyl
Mtt	4-Methyltrityl
Pac	Phenacyl, $PhCOCH_2$ (care! Pac has also been used for $PhCH_2CO$)
Ph	Phenyl
Pht	Phthaloyl
Scm	Methoxycarbonylsulfenyl
TBDMS	*t*-Butyldimethylsilyl
Tf	Trifluoromethanesulfonyl (triflyl)
TMS	Trimethylsilyl
Pmc	2,2,5,7,8-Pentamethylchroman-6-sulfonyl
Pr^i	*i*-Propyl
Pr^n	*n*-Propyl
Tfa	Trifluoroacetyl
Tos	4-Toluenesulfonyl (also Ts)
Troc	2,2,2-Trichloroethoxycarbonyl
Trt	Trityl, triphenylmethyl
Xan	9-Xanthydryl
Z	Benzyloxycarbonyl (also Cbz, discouraged). Z(2Cl) = *o*-chlorobenzyloxycarbonyl etc.: see also remarks on usage with substituted protecting groups under Bzl.

AMINO ACID HETEROCYCLIC DERIVATIVE ABBREVIATIONS

DKP	Diketopiperazine
NCA	*N*-Carboxyanhydride

PTH	Phenylthiohydantoin
UNCA	Urethane *N*-carboxyanhydride

REAGENT AND SOLVENT ABBREVIATIONS

BOP	1-Benzotriazolyloxy-tris-dimethylamino-phosphonium hexafluorophosphate
CDI	Carbonyldiimidazole
DAST	Diethylaminosulfur trifluoride
DBU	Diazabicyclo[5,4,0]-undec-7-ene
DCCI	Dicyclohexylcarbodiimide (also DCC)
DCHU	Dicyclohexylurea (also DCU)
DCM	Dichloromethane
DEAD	Diethyl azodicarboxylate (DMAD = the dimethyl analogue)
DIPCI	Diisopropylcarbodiimide (also DIC)
DIPEA	Diisopropylethylamine (also DIEA)
DMA	Dimethylacetamide
DMAP	4-Dimethylaminopyridine
DME	Dimethoxyethane
DMF	Dimethylformamide
DMS	Dimethylsulfide
DMSO	Dimethylsulfoxide
DPPA	Diphenylphosphoryl azide
EDTA	Ethylenediamine tetraacetic acid
EEDQ	2-Ethoxy-1-ethoxycarbonyl-1,2-dihydroquinoline
HATU	This is the acronym for the 'uronium' coupling reagent derived from HOAt, which was originally thought to be the ***H***exafluorophosphate salt of the *O*-(7-***A***zabenzotriazol-lyl)-***T***etramethyl ***U***ronium cation. In fact, this reagent has the isomeric *N*-oxide structure in the crystalline state, the unwieldy correct name of which does not conform logically with the acronym, but the acronym continues to be used. Similarly, the corresponding reagent derived from HOBt has the firmly attached label HBTU (the tetrafluoroborate salt is also used: TBTU), despite the fact that it is not actually a uronium salt either.
HFIP	Hexafluoroisopropanol
HMP	Hexamethylphosphoric triamide (also HMPA, HMPTA)
HOAt	1-Hydroxy-7-azabenzotriazole
HOBt	1-Hydroxybenzotriazole
HOCt	1-Hydroxy-4-ethoxycarbonyl-1,2,3-triazole
NBS	*N*-Bromosuccinimide
NMM	*N*-Methylmorpholine
PAM	Phenylacetamidomethyl resin
PEG	Polyethylene glycol
PEGA	Polyethylene glycol dimethylacrylamide copolymer
PPA	Polyphosphoric acid
PyBOP	1-Benzotriazolyloxy-tris-pyrrolidinophosphonium hexafluorophosphate
SDS	Sodium dodecylsulfate
TBAF	Tetrabutylammonium fluoride
TBTU	See remarks under HATU above
TCA	Trichloroacetic acid
TEA	Triethylamine
TFA	Trifluoroacetic acid
TFE	Trifluoroethanol
TFMSA	Trifluoromethanesulfonic acid
THF	Tetrahydrofuran
WSCI	Water-soluble carbodiimide: 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide hydrochloride (also EDC)

TECHNIQUE ABBREVIATIONS

Technique abbreviations may be combined in obvious ways, preferably with a hyphen, e.g. ESI-MS.

CD	Circular dichroism
COSY	Correlated spectroscopy
CZE	Capillary zone electrophoresis
ELISA	Enzyme-linked immunosorbent assay
EM	Electron microscopy
ESI	Electrospray ionization
ESR	Electron spin resonance
FAB	Fast atom bombardment
FT	Fourier transform
GLC	Gas liquid chromatography
HPLC	High performance liquid chromatography
IR	Infrared
MALDI	Matrix-assisted laser desorption ionization
MD	Molecular dynamics
MS	Mass spectrometry
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser effect
NOESY	Nuclear Overhauser enhanced spectroscopy
OD	Optical density
ORD	Optical rotatory dispersion
PAGE	Polyacrylamide gel electrophoresis
RIA	Radioimmunoassay
ROESY	Rotating frame nuclear Overhauser enhanced spectroscopy
RP	Reversed phase
SPPS	Solid phase peptide synthesis
TLC	Thin layer chromatography
TOCSY	Total correlation spectroscopy
TOF	Time of flight
UV	Ultraviolet

MISCELLANEOUS ABBREVIATIONS

Ab	Antibody
ACE	Angiotensin-converting enzyme
ACTH	Adrenocorticotropic hormone
Ag	Antigen
AIDS	Acquired immunodeficiency syndrome
ANP	Atrial natriuretic polypeptide
ATP	Adenosine triphosphate
AVP	Arginine vasopressin
BK	Bradykinin; Bn has also been used but should be avoided, as it has been used for bombesin too.
BSA	Bovine serum albumin
CCK	Cholecystokinin
CNS	Central nervous system
CSF	Cerebrospinal fluid
DNA	Deoxyribonucleic acid
ee	Enantiomeric excess
FSH	Follicle stimulating hormone
GH	Growth hormone
HIV	Human immunodeficiency virus
i.m.	intramuscular
i.p.	intraperitoneal
i.v.	intravenous

LHRH	Luteinizing hormone releasing hormone
MAP	Multiple antigen peptide (care! this has also been used for 'multiple-anchored protein')
Mpa	Mercaptopropionic acid, desamino-cysteine
NPY	Neuropeptide Y
MCD	Mast cell degranulating peptide
MIC	Minimum inhibitory concentration
OT	Oxytocin
PCR	Polymerase chain reaction
PNA	Peptide nucleic acid
PTH	Parathyroid hormone
QSAR	Quantitative structure activity relationship
RNA	Ribonucleic acid
RGD	ArgGluAsp (a recurrent sequence motif)
SAR	Structure activity relationship
TASP	Template-assembled synthetic protein
TRH	Thyrotropin releasing hormone
VIP	Vasoactive intestinal peptide
VP	Vasopressin

CLOSING REMARKS

In case any intending authors object that my strictures are pedantic, I say that pedantry is the proper business of editors. More importantly, it is not sufficient to say that the abbreviations and terminology used will be easily understood by those who are active and expert in the specialism concerned, so why must one make a fuss? It is surely the hope of all authors that their work will be intelligible to novices and a wider audience than their professional chums. There is certain irony in the fact that some of them who invest a lot of effort in their English also undermine the clarity of their papers by overloading them with ill-chosen abbreviations and enigmatic usage.

Complete conformity to an agreed set of rules and conventions is an editorial pipe dream. Peptide scientists come from too many scientific subcultures and lands for that. Some flexibility and allowance for taste can reluctantly be tolerated on pragmatic grounds. But the inventive and idiosyncratic approach has to be discouraged, not least because of the increasing use of electronic searching, and for that reason particular care is essential in the composition of titles and abstracts. And the nearer we can get to uniformity, the fewer the confusions and the smoother the editorial process will be.

JOHN H JONES
Editor-in-Chief

REFERENCES

1. Jones JH. *J. Pept. Sci.* 2003; **9**: 1–8.
2. Jones JH. *J. Pept. Sci.* 1999; **5**: 465–471.
3. Erb U, Keller H. *Scientific and Technical Acronyms, Symbols and Abbreviations*. John Wiley & Sons: Chichester; 2001.
4. Stathopoulos P, Papas S, Kostidis S, Tsikaris V. α- and β- aspartyl peptide ester formation *via* aspartimide ring opening. *J. Pept. Sci.* 2005; **11**: 658–664.
5. Filip SV, Cavelier F. A contribution to the nomenclature of depsipeptides. *J. Pept. Sci.* 2004; **10**: 115–118.