

Nomenclature of Amino-acids

The Commission for the Reform of Biochemical Nomenclature of the International Union of Pure and Applied Chemistry has issued revised rules for the nomenclature of amino-acids and allied substances. The new rules are accepted by the editors of the *British Journal of Nutrition* and supersede those published in *Brit. J. Nutr.* 1947, 1, 109.

RULE 1. The configurational relationship of the asymmetric α -carbon atom of an amino-acid capable of optical isomerism should be indicated by a symbol prefixed to the name; however, if a specific statement or the context makes it clear which enantiomorph is under consideration, the symbol may be omitted.

Examples: leucine may be named without prefix if the preparation mentioned is stated to be, or is obviously, the enantiomorph derived from a protein source. Leucine may be named without prefix if the preparation mentioned is stated to be synthetic and not resolved and is therefore an equimolecular mixture of the enantiomorphs. Leucine may be named without prefix in a general statement that is true for either enantiomorph, or for any mixture of these.

RULE 2. (a) Distinction between the enantiomorphs of the amino-acids is made by a prefixed *small capital letter* D or L to denote the configurational family to which the α -carbon atom belongs. The D and L are to be pronounced dee and ell respectively, not *dextro* and *laevo*. An additional symbol to denote the direction of the rotation (i.e., a plus or a minus sign enclosed in parentheses) is not necessary.

Examples: L-leucine, D-valine, L-phenylalanine, L-threonine.

(b) The optically inactive mixture or racemic compound of the enantiomorphs is designated by the prefix DL in *small capital letters*.

Examples: DL-leucine, DL-valine, DL-methionine.

RULE 3. (a) The small capital letter prefixes D and L denote that the substance named is configurationally related to the corresponding enantiomorph of glyceraldehyde. Where confusion is possible between the use of the small capital letter prefix for the configuration of the α -carbon atom in amino-acid nomenclature, and for that of the highest-numbered asymmetric carbon atom in carbohydrate nomenclature, a subscript is added to the small capital letter prefix. Where the prefix is used in the amino-acid sense, the subscript s is added; where the prefix is used in the carbohydrate sense, the subscript c is added. These subscripts (lower-case roman letters) refer, respectively, to serine, the fundamental substance to which amino-acids that bear structural resemblance to the carbohydrates can be formally related, and to glyceraldehyde, the fundamental substance to which the configuration of the carbohydrates is formally related.

Examples: L_s-threonine for which the synonym in carbohydrate nomenclature is 2-amino-2:4-dideoxy-D_c-threonic acid, D_s-threonine for which the synonym is 2-amino-2:4-dideoxy-L_s-threonic acid, L_s-allothreonine for which the synonym is 2-amino-2:4-dideoxy-L_s-erythronic acid, D_s-allothreonine for which the synonym is 2-amino-2:4-dideoxy-D_s-erythronic acid.

(b) Amino-acids derived from amino-sugars should generally be named in conformity with carbohydrate nomenclature but with the use of the subscript.

Examples: D_g-glucosaminic acid for 2-amino-2-deoxy-D_g-gluconic acid, the α -carbon atom of which has the configuration of that in D-serine; D_g-mannosaminic acid for 2-amino-2-deoxy-D_g-mannonic acid, the α -carbon atom of which has the configuration of that in L-serine.

RULE 4. Where the configurational relationship of the α -carbon atom has not been definitely established, or where it is desired to emphasize the actual direction of the rotation of an enantiomorph of known configuration, the direction of the rotation in a specified solvent, preferably of the free amino-acid in water, is designated by prefixes *dextro* or *laevo* in lower case italic letters or alternatively by a plus or a minus sign enclosed in parentheses.

Examples: dextrohydroxytryptophan or (+)-hydroxytryptophan; (+)-glutamic acid or dextroglutamic acid for dextrorotatory (in water) L-glutamic acid.

RULE 5. The prefix *meso* or its abbreviation *ms* in lower-case italic letters is used to denote those amino-acids and derivatives of amino-acids that are optically inactive because of internal compensation.

Examples: mesolanthionine, ms-cystine.

RULE 6. (a) Where the amino-acid has two centres of asymmetry so constituted that internal compensation cannot occur, two diastereomeric forms are possible both of which possess the L_s-configuration at the α -carbon atom; of these forms the *first to be described* is designated the L-amino-acid, and its enantiomorph is designated the D-amino-acid. Where the name contains one or more prefixes denoting substitution, the L- or D- is, in accordance with Rule 7, placed immediately before the part of the name (usually the trivial name of the parent amino-acid) which signifies an asymmetric configuration around the α -carbon atom.

Examples: L-threonine, L-isoleucine, for the amino-acids of protein origin which have trivial names; D-threonine, D-isoleucine for their enantiomorphs. Where there is a prefix denoting substitution, the names take the form hydroxy-L-proline, hydroxy-DL-glutamic acid, hydroxy-L-lysine, or, with specification of position of substitution, 4-hydroxy-L-proline, β -hydroxy-DL-glutamic acid, δ -hydroxy-L-lysine.

(b) The other diastereomer which possesses the L_s -configuration at the α -carbon atom is distinguished by the prefix *allo* in addition to the prefix *L*. Its enantiomorph is denoted by the prefix *allo* and the prefix *D*. Where the name is wholly trivial, the *L*- or *D*- is placed before the prefix *allo* which is, in turn, attached to the parent name; but in names which contain one or more prefixes denoting substitution, the *allo* is placed before this prefix and the *L*- or *D*- is placed as stated in section (a).

Examples: *L*-allothreonine, *D*-allothreonine, *L*-alloisoleucine, *D*-alloisoleucine for the pairs of enantiomorphs of the diastereomers of the amino-acids of protein origin which have trivial names; *allohydroxy-L*-proline, *allohydroxy-D*-proline, *allohydroxy-L*-lysine, *allohydroxy-D*-lysine, and similarly for the enantiomorphs of the diastereomers named as derivatives of amino-acids which have trivial names. Where the position of the substituent group is designated, the names take the form, *allo-4*-hydroxy-*L*-proline, *allo- δ* -hydroxy-*L*-lysine. Systematic names take the form *allo- β* -hydroxy-*L- α* -aminobutyric acid (for *L*-allothreonine), *allo-4*-hydroxy-*L*-pyrrolidine-2-carboxylic acid (for *allohydroxy-L*-proline).

(c) For diastereomeric α -amino-acids which have structures not encountered in nature but which are named as derivatives of naturally occurring amino-acids with trivial names, choice between the prefixes *L* and *D* is made, (1) from the results of direct correlation with substances of known configuration, or (2), tentatively, from the results of studies of biological properties or of the change in optical rotation with change in the conditions of observations, or both (1) and (2). The assignment of the prefix *allo* to the pair of enantiomorphs of one of the diastereomers is made tentatively, if possible, in accordance with the principle in section (a). The prefixes are placed as stated in section (a).

Examples: of the four theoretically possible optically active isomers of hydroxyaspartic acid, two should yield *L*-aspartic acid on reduction and are accordingly hydroxy-*L*-aspartic acid and *allohydroxy-L*-aspartic acid, the choice of these designations being made tentatively as specified. The respective enantiomorphs are named with the prefix *D*.

RULE 7. Salts and derivatives of amino-acids including peptides are designated with the use of a *small capital letter* to denote the configurational family of the α -carbon atom or atoms, this letter being placed immediately before the name of the parent acid or its radical. The customary rules of nomenclature are otherwise observed.

Examples: *L*-histidine monohydrochloride monohydrate, copper *L*-aspartate, *D*-ornithine dihydrochloride, acetyl-*L*-tryptophan, diethyl *D*-glutamate, β -hydroxy-*DL*-glutamic acid, *L*-leucyl-*L*-valine, glycyl-*DL*-leucine. Names in which the prefixes involve amino-acid configurations are treated similarly; thus, *S*-(*D*-2-amino-2-carboxyethyl)-*D*-homocysteine for *D*-cystathionine.

Where the name contains one or more prefixes denoting substitution and where specification both of configuration and of position of substitution is required, this rule should be applied, the form taken being as in the names of the following naturally occurring substances: β -phenyl-*L*-alanine, 4-hydroxy-*L*-proline, 3:5-di-iodo-*L*-tyrosine, 3:5-dibromo-*L*-tyrosine, δ -hydroxy-*L*-lysine. However, in general biochemical writing, the names of the following four substances, the position of substitution of which is well understood, are admissible as exceptions: *L*-phenylalanine, *L*-hydroxyproline, *L*-di-iodotyrosine, and *L*-hydroxylysine.

RULE 8. Where a trivial name is applied to a compound that contains two α -amino-carboxylic acid groupings attached to dissimilar carbon chains, the configurational designation is that of the α -carbon atom of the smaller of the two chains.

Examples: *D*-cystathionine for *S*-(*D*-2-amino-carboxyethyl)-*D*-homocysteine, *L*-allocystathionine for *S*-(*L*-2-amino-2-carboxyethyl)-*D*-homocysteine.

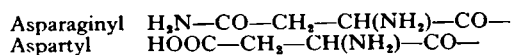
However, the introduction of new trivial names should be avoided in the absence of compelling reasons.

RULE 9. The word 'inactive' or the prefix (\pm) shall be used for inactive amino-acids and their derivatives, (a) when two or more asymmetric centres are present but the steric cause of the inactivity is unknown, and (b) for the inactive form of an amino-acid of which an active form of uncertain configurational relationship has been described.

Examples: (a) inactive β -hydroxyglutamic acid; (b) (\pm)-hydroxytryptophan.

RULE 10. The International Commission has adopted a new Rule 10 on amino-acid radicals. This rule states that amino-acid radicals shall be named in conformity with the Definitive Report of the Commission on Nomenclature in Organic Chemistry. The rule gives a full list of amino-acid radicals with their formulas. It seems unnecessary to reproduce the whole of this list. The general rule is that for an amino-acid $R \cdot CHNH_2 \cdot COOH$ (name ending in 'ine') the radical $R \cdot CHNH_2 \cdot CO-$ has the corresponding name ending in 'yl'; thus 'alanine' gives the radical 'alanyl'. 'Tryptophan' forms 'tryptophyl'.

The following may be noted specially:



Similarly, glutaminyl, glutamyl.

