

Changes in the Indexing, Registration, and Naming of Chemical Substances for the 14th Collective Period (1997-2001)

Kappa in Coordination Nomenclature

The Kappa system replaces the donor atom system for describing the position of ligand attachment to a coordination center.

In the Kappa system, when it is necessary to indicate the attachment of a ligand to a coordination center, the ligating atom(s) are indicated by the italic element symbol of the atom(s) preceded by a Greek Kappa (κ). This combination of italic element symbol and Greek Kappa is known as the Coordination Bond Site Identifier (CBSI). The Coordination Bond Site Identifier is placed after that portion of the *Chemical Abstracts (CA)* index name to which it directly applies; i.e., the part that contains the coordinating function, atom, or ring. For example: (2-aminoethanolato- κO), [*N*-[2-(amino- κN)ethyl]-*N'*-(2-aminoethyl)-1,2-ethanediamine- $\kappa N, \kappa N'$], [μ -(pyrimidine- $\kappa N^1: \kappa N^3$)]. For conjunctively named parents, the CBSI is placed after the portion of the name denoting the function: (3-pyridinemethanol- κN^1). In multiplicative nomenclature, the CBSIs are placed with the hetero atoms that are part of the multiplying segment of the name and at the end of the parent portion of the name: [[2,2'-[1,2-ethanediy]bis[(nitriolo- κN)methylidyne]]bis[phenolato- κO]](2-).

The use of enclosing marks follows established conventions except that extra enclosing marks are used around that portion of the *CA* name that contains the coordinating function, atom, or ring if they are not already in use. Extra enclosing marks are also used if a locant is required to designate the point of attachment of the coordinating atom, function or ring to the ligand. For example: [2-(amino- κN)ethyl], (methylamino- κN). The standard multiplicative prefixes (di-, tri-, bis-, tris-, etc.) operate on the CBSI.

Revision of Stereochemical Practices

To provide more accurate descriptions of and improved access to substances whose stereochemistry has not been completely defined, CAS now registers and names substances with partially defined stereochemistry. Previously, partial stereochemistry was generally ignored.

Existing descriptor formats are retained for cases where all centers are known. The presence of unknown chiral centers is indicated by the addition of the term “[*partial*]-” to the end of the normal stereochemical descriptor. Unknown double bond geometry does not in itself require the “[*partial*]-” designation. When the reference ring or chain has incompletely defined chiral atoms/bonds, the format cites the stereo using *R* and *S* terms with their nomenclature locants for all known centers. If this method is used to describe a substance for which only relative stereochemistry is known, “*rel*” is added to the stereochemical descriptor. Any stereochemical descriptor marked as “*rel*” always cites the first center as *R* and relate all other centers back to this first center using an *R* if the chirality is identical and an *S* if the chirality is the opposite.

Racemic mixtures of substances having single chiral centers are now indexed, registered, and named as non-stereospecific substances. The decision to discontinue the racemate distinction

was based on the observation that most of the chemical substances encountered in the literature are presented without any explicit indications of their racemic nature but which in the majority of the cases are probably racemates. The distinction that CAS was making between the racemic and non-stereospecific (or relative) forms was judged to cause scattering in indexing and registration, and complications in database searching.

Thus, DL-malic acid and malic acid (stereospecificity unspecified) now receive the same CAS Registry Number® and CA index name: 6915-15-7, Butanedioic acid, hydroxy-. The Registry Number for the DL-form (617-48-1) will be cross-referred to the Registry Number of the nonstereospecific form (6915-15-7).

Racemates having more than one chiral center are indexed, registered, and named as having only relative stereochemistry. Thus, DL-threitol and threitol (absolute stereospecificity unspecified but having two chiral centers with the same relative configuration) now receive the same CAS Registry Number and index name: 7493-90-5, 1,2,3,4-Butanetetrol, (*R*,R**)-. Again, the Registry Number for the DL-form (6968-16-7) will be cross-referred to the Registry Number of the relative-only stereospecific form (7493-90-5).

Stereoparent Nomenclature Simplification

In order to reduce the number of trivial and semi-systematic terms that appear in CA index names, some infrequently used terpene, steroid, alkaloid, and antibiotic stereoparent terms have been replaced by systematic names. For example, Nemuarine and 15-Thialanostane are no longer used in index names. CAS has maintained the frequently occurring stereoparents, such as Pregnane, Cholane, Cholestane, 9,10-Secocholestane, Morphinan, Retinoic acid, and Erythromycin. Cross-references from the previously used stereoparent names will guide users to the corresponding systematic index names.

For the 14th Collective, several revisions in peptide nomenclature have been made. Highlights include the following:

- No structure is assigned a peptide name unless it contains at least two standard amino acid residues. A standard amino acid residue is any amino acid that can stand alone as a stereoparent (e.g., glycine or tryptophan), plus α -asparagine and α -glutamine. (2-Aminobutanoic acid and 2,4-diaminobutanoic acid are now considered nonstandard amino acids.)
- Amino acid sequence names are now assigned to most systematically named linear peptides. In an amino acid sequence name, the C-terminal residue is the index heading parent, and the other residues are cited in the substituent, beginning with the N-terminal residue and continuing from left to right in the sequence; e.g., L-Lysine, D-alanyl-glycyl-L-leucyl-. This simpler nomenclature replaces the many locants and enclosing marks that have been used in the name in the past, e.g., L-Lysine, N^2 -[N-(N-D-alanyl-glycyl)-L-leucyl]-.
- Whether terminal or nonterminal, α -hydroxy acids and nonstandard α -amino acids are treated as residues and assigned systematic acyl or acid names. Other unusual residues are permitted, but only in nonterminal positions. The stereochemistry of all these residues is described using systematic *R/S* stereodescriptors rather than *D/L*, *cis/trans*, etc.

- Psi (ψ) nomenclature is now used to describe certain modifications of the peptide bond. The Greek letter ψ conveys the fact that a peptide bond has been replaced by a pseudopeptide bond. In an amino acid sequence name, the format of the ψ term is ...-A- ψ (X-X')-B-..., where A is the amino acyl radical whose carbonyl group has been modified to X and B the amino acyl radical whose α -amino group has been modified to X'. X and X' are shown as strings of element symbols, separated by a bond; e.g., ...-L-valyl- ψ (CH₂-NH)-L-tyrosyl-. In a peptide stereoparent derivative name, the format is similar, except that residue-number locants precede and follow the ψ term; e.g., Bradykinin, ⁷ ψ ⁸(CH=CH,*E*)-8-L-tyrosine-.