Chemical Abstracts Stereochemical Nomenclature of Organic Substances in the Ninth Collective Period (1972–1976)

JAMES E. BLACKWOOD* and PATTON M. GILES, JR.

Chemical Abstracts Service, The Ohio State University, Columbus, Ohio 43210

Received October 25, 1974

This paper discusses Chemical Abstracts indexing policy revisions in the form and content of that part of the CA Index Name which depicts the stereochemistry of an organic substance. The new stereochemical descriptors are more consistent and systematic, resulting in a closer grouping of structurally related substances in the indexes.

Chemical Abstracts Service (CAS) continually develops improvements in CA indexing policies, implementing major changes in subject- and substance-indexing policies at fiveyear intervals to correspond with the beginning of a Collective Index (CI) period. Substantial changes in name selection policies have been made for the Volume 76 indexes, 1,2 the first volume of the Ninth Collective Index (9CI) period (1972-1976). These changes relate to the increased use of fully systematic names for most CA index substances. The stereochemical descriptor portion of the CA Index Name has been revised to provide more consistency and to allow a closer grouping of structurally related substances in the indexes.3

Beginning with Volume 76, chemical substances which differ only in stereochemistry appear consecutively in the Chemical Substance Index, since the stereochemical descriptor has been separated from the rest of the name and is now cited as the last element of the name. For example, previous to the 9CI revisions, 5α -Androstan-1-one and 5β -Androstan-1-one were separated by several columns of printed index entries. Now they appear at the same index heading, Androstan-1-one, with the stereochemical descriptor indicating the only point of structural difference.

> Androstan-1-one, (5α) -Androstan-1-one, (5β) -

Systematically named chemical substances are assigned a stereochemical descriptor which relates all of the asymmetric centers contained therein, together with a descriptor which specifies the absolute configuration of one center or the optical rotation of the entire molecule. For example, the carboxylic acid shown below is assigned the CA Index Name: Cyclopentanecarboxylic acid, 3-methyl-, (1R-cis)-.

Example 1

The absolute configuration of one center, as determined by the Sequence Rule,4 is described by 1R, and the relationship of the two centers is expressed by cis-. Each part of

the systematic stereochemical descriptor possesses a mutually exclusive vocabulary, thus assuring that absolute and relative stereochemistry can be distinguished. Using the Eighth Collective Index (8CI) stereochemical descriptors, this differentiation between absolute and relative was not always possible. For example, the naphthalene derivative shown below was named utilizing the stereochemical descriptors α or β to denote whether the substituent is respectively behind or in front of the plane of the ring. Although the 8CI name was intended to indicate only relative or only absolute stereochemistry, it was impossible to determine which meaning was intended without referring to the primary document.

Example 2

8CI Name

Naphthalene, 6β -ethyl- $1,2,3,4,4a\beta,5,6,7,8,8a\alpha$ decahydro-2β-isopropyl- 1β ,6-dimethyl-1-propyl-(this name was used to indicate either absolute or relative stereochemistry)

9CI Names (a) Absolute stereochemis-

- Naphthalene, 6-ethyldecahydro-1,6-dimethyl-2-(1-methylethyl)-1-propyl-,
 - $[1R-(1\alpha,2\beta,4a\beta,6\beta,$ $8a\alpha$)]-
- (b) Relative stereochemistrv Naphthalene, 6-ethyldecahydro-1,6-dimethyl-2-(1-methyl-

ethyl)-1-propyl-, $(1\alpha,2\beta,4a\beta,6\beta,8a\alpha)$ -

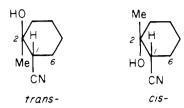
The presence of the absolute descriptor 1R in example (a) indicates that absolute stereochemistry is intended. The terms R and S are used for absolute stereochemistry and such terms as α , β , cis, trans, endo, exo, R^* , S^* , E, and Z are used as relative descriptors. Definitions of these terms and rules for their use appear later in this paper.

Revised definitions of those terms which are used in the relative portion of the stereochemical descriptor allow the

^{*} To whom correspondence should be addressed.

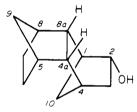
unambiguous specification of the stereochemistry of substances which could not be fully described in the 8CI. For example, cis and trans are now defined as referring to the relationship between the higher ranking substituents at each of two asymmetric centers. Thus, the isomers of 2hydroxy-2-methylcyclohexanecarbonitrile, which were described previously only by the term "stereoisomer," can now be differentiated.

Example 3



Additionally, the use of certain terms has been limited. resulting in fewer ambiguous stereochemical descriptors. Not only did the widespread use of such terms as endo and exo result in much confusion in certain areas of nomenclature, but many compounds to which these descriptors were applied can better be described utilizing the α,β system as defined in the following section. The use of endo, exo, syn, and anti has been restricted to those bicyclic systems wherein there can be no doubt as to their application. In example 4, the compound was previously described only as (endo, exo)-, with no specification of which group was "endo" to which other group. For the 9CI name, the methano bridges (9 and 10) are treated as substituents at positions 1, 4, 5, and 8.

Example 4



1,4:5,8-Dimethanonaphthalen-2-ol, decahydro-, $(1\alpha, 2\alpha, 4\alpha, 4a\beta, 5\beta, 8\beta, 8a\beta)$ -

In addition to these considerable benefits to the user, the new stereochemical descriptors require less professional analysis by CAS staff, and the controlled vocabulary will allow a computer edit of the descriptor content. The shift to descriptors which are more readily translatable into structures is an important step toward the development of computer-based procedures for verifying and generating chemical substance names.5

DEFINITIONS AND RULES

Depending on the class of compound being named and its complexity, two basic types of stereochemical descriptors are used: the systematic stereochemical descriptor and the stereoparent stereochemical descriptor. More common is the systematic stereochemical descriptor which is used with systematically named compounds.

Natural products, especially those with several asymmetric centers, are conveniently indexed at trivial names, e.g., pregnane and morphinan, which are called stereoparents since their names imply certain stereochemistry. When stereoparent derivatives have additional asymmetric centers, or when the stereochemistry which is implied by the stereoparent name has changed, this information is placed in a separate stereochemical descriptor as with systematically named substances.

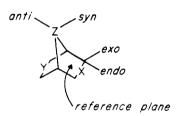
Systematic Stereochemical Descriptors. In the longestablished system for carbohydrate derivatives, two terms are used, e.g., D-gluco-, in which one center is defined absolutely by "D" and three other centers are related to it by "gluco" to describe the set of four centers absolutely.6 Analogously, among systematic stereochemical descriptors the absolute terms R and S are employed for a single "reference" center, and the following relative terms are cited under various conditions as defined below: cis. trans. exo. endo, syn, anti, α , β , R^* , S^* , E, and Z.

An absolute descriptor R or S is assigned by application of the Sequence Rule. Assignment of the absolute terms R and S depends on the ranking of atoms and groups attached to the central atom under consideration. This is based first on atomic number; e.g., for the common halogens the descending order is I, Br, Cl, F. For carbon compounds it is necessary to proceed one atom at a time away from the center along each chain until a decision is reached.4

Relative descriptors of various kinds are used as follows:

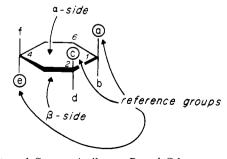
- (a) cis and trans are restricted to cyclic structures with two stereogenic atoms (an atom is stereogenic when interchange of two of the atoms or groups attached to it produces a nonidentical compound). The cis isomer has the senior groups (by the Sequence Rule) on the same side of the reference plane.
- (b) endo and exo are used only for ring positions on the X and Y bridges of bicyclo[X.Y.Z]anes in which $X \ge Y > Z$ > 0. The exo isomer has the configuration in which the senior substituent is on the same side of the reference plane as the Z bridge (see diagram below).
- (c) syn and anti are restricted to ring positions on the Z bridge of bicyclo[X.Y.Z]anes. The syn isomer has the senior substituent oriented towards the X bridge.

Example 5



(d) α and β are employed for ring positions of cyclic compounds. In the diagram below, the senior groups at the three stereogenic centers are a, c, and e. The numbering of the ring is determined by the usual rules of substitutive nomenclature. The α -side of the reference plane is that side on which the preferred substituent lies at the lowest numbered stereogenic position (i.e., position 1, in this example); c lies on the same side as a, so both are given α descriptors; e lies on the opposite side of the reference plane and is related by a β descriptor; hence: 1α , 2α , 4β .

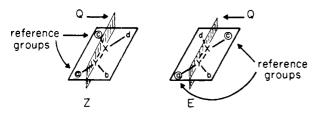
Example 6



(e) R^* and S^* are similar to R and S but are relative in that the lowest numbered (or first cited) stereogenic atom is arbitrarily assigned an R* descriptor.4

(f) E and Z are geometrical stereodescriptors for substances having achiral elements resulting from double bonds.^{4,7} For that configuration in which the two groups of higher priority are on the same side of the reference plane Q, the stereochemical descriptor Z is used; the configuration in which these groups are on opposite sides is depicted by the descriptor E.

Example 7



(g) The optical rotation descriptors (+) and (-) indicate the sign of rotation of plane-polarized light at 589 m μ (sodium D line). Racemates are indicated by a (±) descriptor.

The following rules describe the use of these descriptors (absolute, relative, and optical) to express the total stereochemical information for a chemical substance.

1. Absolute configuration

- (a) When only one chiral element is present, the absolute descriptor R or S is cited in parentheses and followed by a hyphen, thus: (R).
- (b) When two or more chiral elements are present, the reference center is given an R or S descriptor. This reference center, for cases described in 2 (a, b, e, f, and h) below, is the lowest numbered chiral center.
- (c) For 2.c and 2.g it is the first cited α position.
- (d) For 2.d it is the center which has the highest ranking substituent according to the Sequence Rule.⁸
- Relative configuration (for substances containing two or more stereogenic atoms)
 - (a) When only two stereogenic atoms are present in an eight-membered or smaller ring or a ring system consisting of such rings, and the configurations of both are known, cis or trans is cited as the relative descriptor. R* or S* is employed for spiro centers named as ortho fusions and for rings of more than eight members (see rule 2.i for cases involving chiral and achiral elements).
 - (b) endo, exo, syn, and anti are used for bicyclo[X.Y.-Z]ane compounds, in which X ≥ Y > Z > 0, with no other known centers. (See rule 2.i or cases involving chiral and achiral elements.)
 - (c) For two or more stereogenic centers in a ring system to which neither rule 2.a nor 2.b applies, α and β are cited. R* or S* is employed for spiro centers named as ortho fusions, for bridging atoms, and for rings of more than eight members. (See rule 2.i for cases involving chiral and achiral elements.)
 - (d) When two chiral centers are not in the same ring system, R* and S* are employed without locants. Two centers of like chirality are indicated by the relative descriptor R*, R* while two centers of unlike chirality are described by R*, S*.
 - (e) For substances with three or more chiral elements to which neither rule 2.b nor 2.c applies, R* and S* are used with appropriate locants. The lowest numbered chiral center is assigned R* and the other centers are related to it.
 - (f) The descriptors E and Z, for configurations about double bonds, are cited in descending order of seniority when used alone or with R* and S* as in rule 2.d.⁸ When other descriptors are present, locants are cited with E and Z.
 - (g) When both cyclic and acyclic chiral elements are present, rules 2(c-f) are invoked. When rule 2.c is

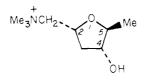
- applicable, the ring is the parent from which the descriptor is derived; the first cited α -center is the reference center and is implicitly R^* .
- (h) When chiral elements are present in two ring systems, that system containing the greater number of such elements is the parent from which the descriptor is derived (by rules 2(c-f)).
- (i) When achiral elements present in a ring can be described by rules 2(a-c), any chiral element outside the ring is described by rule 1.a.

3. Optical rotation

- (a) When the absolute stereochemistry is described, the sign of rotation is omitted.
- (b) When the substance has only one chiral element, and this is not defined by the original document, the sign of rotation is cited to distinguish enantiomers.
- (c) When the substance has more than one chiral element, the sign of rotation is cited after the relative descriptor.

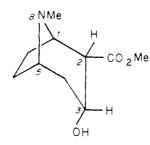
In the following examples, the absolute descriptor is cited first, followed by the relative descriptors in parentheses. The entire descriptor set is then bracketed and cited as the last element in the name.

Example 8

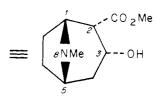


2-Furanmethanaminium, tetrahydro-4-hydroxy-N,N,N,5-tetramethyl-, $[2R-(2\alpha,4\alpha,5\beta)]$ - [The first cited chiral center is given the absolute descriptor R by rule l.c. The relationship of the three stereogenic centers is expressed according to rule 2.c.]

Example 9

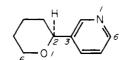


8-Azabicyclo[3.2.1]octane-2-carboxylic acid, 3-hydroxy-8-methyl-, methyl ester, [1R-(endo,endo)]-



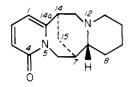
[The lowest numbered chiral center uses the absolute descriptor *R* by rule 1.b. The relationship of the two stereogenic centers on the X bridge to the Z bridge is expressed according to rule 2.b. No locants are needed for the two *endo* descriptors since there is no ambiguity without them.]

Example 10



Pyridine, 3-(tetrahydro-2H-pyran-2-yl)-, (S)- [The chirality of the single atom is expressed according to rule 1.a.]

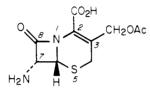
Example 11



7,14-Methano-4H,6H-dipyrido[1,2-a:1',2'-e]-[1,5]diazocin-4-one, 7,7a,8,9,10,11,13,14-octahydro-, $(7\alpha, 7a\beta, 14\alpha)$ -(+)-

The relationship of the three stereogenic centers is expressed according to rule 2.c. At positions 7 and 14 the methano bridge is treated as if it were a substituent. At position 7a, the hydrogen atom is the substituent. The optical rotation is used to distinguish this substance from its enantiomer (rule 3.c.)]

Example 12



5-Thia-1-azabicyclo-[4.2.0]oct-2-ene-2-carboxylic acid, 3-[(acetyloxy)methyl]-7-amino-8oxo-, (6R-trans)-

The reference center is expressed as before (rule 1.b). The relationship of the hydrogen at position 6 to the amino group at position 7 is indicated by trans as defined by rule 2.a.

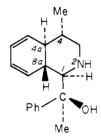
No absolute stereochemistry is possible. The two achiral stereogenic centers are related by rule 2.a.]

Example 13



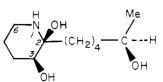
1,4-Cyclohexanediol, trans-

Example 14



1-Isoquinolinemethanol, 1,2,3,4,4a,8a-hexahydro- α ,4-dimethyl- α phenyl-, $[1R-[1\alpha(S^*),$ 4β , $4a\beta$, $8a\alpha$]]-

Example 15



2,3-Piperidinediol, 2-(5hydroxyhexyl)-, [2S- $[2\alpha, 2(S^*), 3\alpha]$ -

The reference center is expressed as before (rule 1.c). Rule 2.g is used to relate the various stereogenic centers. The chiral acyclic atom is assigned S* since it possesses S* chirality when the reference center is R^* . The S^* is placed in curves following the locant and stereochemistry of the point of attachment.]

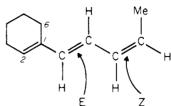
The reference center is described as before (rule 1.c). The three stereogenic centers are related according to rule 2.g.]

Example 16



2-Butanol, 3-chloro-. $[S-(R^*,R^*)]$ -

Example 17



Cyclohexene, 1-(1,3pentadienyl)-, (E,Z)- The reference center is described as before (rule 1.d). The two stereogenic centers are related according to rule 2.d.]

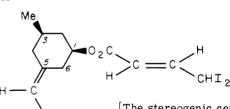
The relationship between the two acyclic double bonds is described as in rule 2.f.

Example 18

2-Furanmethanol, tetrahydro- α -methyl-, $[S-(R^*,S^*)]-$

The reference center is depicted as before (rule 1.d). The two stereogenic centers are related according to rule 2.d.]

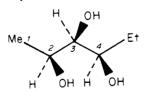
Example 19



2-Butenoic acid, 4,4-diiodo-, 3-methyl-5-propylidenecyclohexyl ester, $[1\alpha(E), 3\alpha, 5Z]$ -(a relative descriptor)

[The stereogenic centers are related according to rule 2.c. The double bonds are related according to rule 2.f.

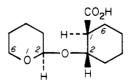
Example 20



2,3,4-Hexanetriol, $[2S-(2R^*,3S^*,4S^*)]$ -

The reference center is expressed as before (rule 1.b). The three stereogenic centers are related by rule 2.e.]

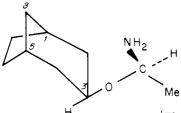
Example 21



Cyclohexanecarboxylic acid, 2-[(tetrahydro-2Hpyran-2-yl)oxy]-, $[1S-[1\alpha,2\beta(S^*)]-$

[The ring which contains the greatest number of stereogenic centers is the parent from which the descriptor is derived (rule 2.h).]

Example 22



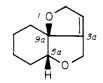
Ethanamine, 1-(bicyclo-[3.2.1]oct-3-yloxy)-, [3(R)-exo]-

The acyclic center is the only chiral stereogenic center in the molecule. The absolute descriptor is cited with a locant specifying the point of attachment of the side-chain (rule 2.i).

Example 23

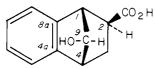
2H-Pyran-2-ol, tetrahydro-, (-)-

Example 24



7H-Furo[3,2-c]benzofuran, 2,4,5a,6,8,9-hexahydro-, [5aR - (5aR * , 9aS *)]-

Example 25



1,4-Methanonaphthalene-2-carboxylic acid, 1,2,3,4-tetrahydro-9hydroxy-, $[1S_{-}(1\alpha, 2\alpha,$ $4\alpha,9R^*)$]-

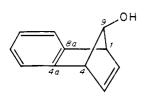
[The stereochemistry of a molecule containing a single chiral element of unknown configuration is described according to rule 3.b.

Since the spiro center, 9a, is contained in a ring assigned an ortho fusion name, the relative stereochemistry is indicated by R^*/S^* rather than cis/trans. See rule 2.a.

[The stereochemistry at position 9 is indicated by the descriptor term R^* . This descriptor term means that the bridging atom, position 9, has the R* chirality when the reference center, position 1, has the R^* chirality. See also rule 2.c.]

When an unambiguous stereochemical descriptor cannot be assigned by utilizing these rules, the substance is assigned the descriptor "stereoisomer." This implies that stereochemistry is indicated in the original document.

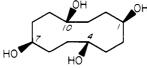
Example 26



1,4-Methanonaphthalen-9ol, 1,4-dihydro-, stereoisomer

[The stereochemistry of the hydroxy substituent on the bridge cannot be described by the R^*/S^* system because of the symmetry of the molecule. In addition, the syn/anti system (as defined above) cannot be applied since this is not a bicyclic system. Thus, the presence of unspecified stereochemistry is indicated by the descriptor "stereoisomer."

Example 27

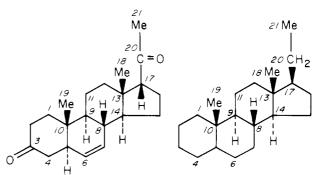


1,4,7,10-Cyclododecanetetrol, stereoisomer

This system could have been described by the α/β system were it not for the size of the ring. The symmetry of the system prohibits the use of R^*/S^* descriptors.

Stereoparent Stereochemical Descriptors. The use of stereoparents dispenses with some lengthy stereochemical descriptors. Cross-references and synonyms relating stereoparents to the corresponding systematic names appear (as stereospecifically as possible) in the CA Index Guide.9 Diagrams appear in the CA Chemical Substance Index¹⁰ to aid the user in visualizing the stereoparents. To express differences between a stereoparent diagram and the structure of the reported substance, a stereochemical descriptor is cited at the stereoparent heading. The stereochemistry of stereoparents and derivatives is assumed to be absolute unless otherwise specified. If relative or racemic substances are to be indicated, the stereochemical descriptor is suffixed by rel- or (\pm) -, respectively. In the diagrams the configurations of ring system substituents are indicated by dotted lines for α and "wedges" for β . Variations for specific derivatives are indicated by citing appropriate α and β terms with the locants of the modified centers. Stereochemistry for acyclic and bridge centers is indicated by E and Z or by R and S as described for systematic stereochemistry. Unless a variant is reported, the normal stereochemistry is considered to be retained. When the modified stereochemistry is unknown, the descriptor ξ (xi) is assigned to that position. Descriptors are also cited for added configurations when the stereoparent diagram shows no stereochemistry at a center. When the added stereochemistry occurs in a stereoparent portion of the substance which is not cited in the index heading parent but rather at a secondary index level, the descriptor is cited as a prefix, e.g., [$(5\alpha, 16E)$ -androstan-16-ylidene]. When additional stereochemistry is present in a systematically named modification of the index heading parent, the descriptor is placed in the usual position (after all other structural information) where it is identified by a locant relating to the stereoparent, e.g., Pregnane-3,20-dione, 5-hydroxy-6-(3-hydroxy-1propenyl)-, $[5\alpha,6(Z)]$ -. In the following examples, the stereoparent diagram is compared with an original document diagram for a derivative.

Example 28

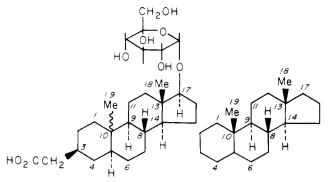


Pregn-6-ene-3,20-dione, $(5\alpha, 17\alpha)$ -

Pregnane (stereoparent)

The configurations at positions 8, 9, 10, 13 and 14 correspond; position 5 has added stereochemistry; position 17 has modified stereochemistry. The stereochemical descriptor at Pregn-6-ene-3,20-dione is $(5\alpha,17\alpha)$ -.

Example 29

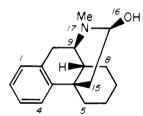


Androstane-3-acetic acid, 17-(α-p-glucopyranosyloxy)-, $(3\beta, 5\alpha, 10\xi, 17\beta)$ -

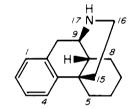
Androstane (stereoparent)

Positions 8, 9, 13, and 14 are normal; position 10 has unknown modified stereochemistry; positions 3, 5, and 17 have added stereochemistry and the substituent at 17 is a stereoparent (α -D-glucopyranosyl radical). The stereochemical descriptor is $(3\beta, 5\alpha, 10\xi, 17\beta)$. The stereochemistry within the stereoparent is expressed by the substituent radical name.

Example 30



Morphinan-16-ol, 17methyl-, (16S)-



Morphinan (stereoparent)

The stereochemistry at position 16 is cited as (16S)-. since positions 15, 16, and 17 are on a bridge.

CONCLUSION

The techniques and stereochemical notation presently used by CAS to describe the organic substances included in the Ninth Collective Indexes (1972-1976) to Chemical Abstracts have been presented with illustrative examples. Extension of this system to cover more fully the stereochemistry of chemical substances will be the objective of ongoing developmental work at CAS.

ACKNOWLEDGMENTS

The helpful comments of Dr. Hans Hirschmann of Case Western Reserve University are gratefully acknowledged. CAS gratefully acknowledges the partial support of the National Science Foundation for portions of the research efforts that have led to the development of these nomenclature techniques.

LITERATURE CITED

- (1) Donaldson, N., Powell, W. H., Rowlett, R. J., Jr., White, R. W., and Yorka, K. V., "Chemical Abstracts Index Names for Chemical Substances in the Ninth Collective Period (1972-1976)," J. Chem. Doc., 14. 3-15 (1974)
- (2) Tate, F. A., and Rowlett, R. J., Jr., "A Computer-Based System for Handling Chemical Nomenclature and Structural Representations," J. Chem. Doc., 12, 125-128 (1972).
- (3) For a discussion of the stereochemistry of inorganic substances, see Brown, M. F., Cook, B. R., and Sloan, T. E., "Stereochemical Notation in Coordination Chemistry, Mononuclear Complexes," in press.
- (4) "IUPAC Tentative Rules for the Nomenclature of Organic Chemistry. Section E. Fundamental Stereochemistry," J. Org. Chem., 35, 2849-2867 (1970). For a more detailed discussion of the Sequence Rule, see Cahn, R. S., Ingold, C. K., and Prelog, V., "Specification of Molecular Chirality," Angew. Chem., Int. Ed. Engl., 5, 385-415 (1966).
- (5) Vander Stouw, G. G., Naznitsky, I., and Rush, J. E., "Procedures for Converting Systematic Names of Organic Compounds into Atom-Bond Connection Tables," J. Chem. Doc., 7, 165-169 (1967).
- (6) "Rules of Carbohydrate Nomenclature," J. Org. Chem., 28, 281-291 (1963).
- (7) Blackwood, J. E., Gladys, C. L., Loening, K. L., Petrarca, A. E., and Rush, J. E., "Unambiguous Specification of Stereoisomerism about a Double Bond," J. Amer. Chem. Soc., 90, 509-510 (1968).
- (8) Blackwood, J. E., Gladys, C. L., Petrarca, A. E., Powell, W. H., and Rush, J. E., "Unique and Unambiguous Specification of Stereoisomerism about a Double Bond in Nomenclature and Other Notation Systems," J. Chem. Doc., 8, 30-32 (1968).
- (9) The Chemical Abstracts Index Guide is a list of entry points to the various CA Volume Indexes. It was first printed in 1972 for Volume 76 and is being supplemented annually through the Ninth Collective Index Period (1972 - 1976)
- (10) The Chemical Substance Index, which is part of the CA Volume Indexes, relates the CA Index Names of specific chemical substances to their assigned CAS Registry Numbers and corresponding CA abstract numbers.