The Chemical Abstracts Service Chemical Registry System. III. Stereochemistry†

JAMES E. BLACKWOOD,* PATRICIA M. ELLIOTT, ROBERT E. STOBAUGH, and CHARLES E. WATSON

Chemical Abstracts Service, Columbus, Ohio 43210

Received December 28, 1976

The Chemical Abstracts Service (CAS) Chemical Registry System is a computer-based system that uniquely identifies chemical substances on the basis of composition and structure. The unique and unambiguous representation of stereoisomers has been an integral part of the CAS Registry System since its inception as Registry I in 1964. Various techniques for recording stereoisomers and refinements made to the stereochemical portion of the computer record are described. The stereochemical terms used in the current system, Registry III, are emphasized, and input and validation techniques for these terms are discussed.

INTRODUCTION

The Chemical Abstracts Service Chemical Registry System is a computer-based system that uniquely identifies chemical substances on the basis of their molecular structure. The overall design of the system has been described by Dittmar, Stobaugh, and Watson.[1]

Stereochemical representation has been an integral part of the CAS Chemical Registry System since its inception as Registry I late in 1964.[2] Stereoisomers are considered to be different, individual substances, and each is recognized as unique. The unique identification of stereoisomers permits the storage and retrieval of information collected from scientific literature about a specific isomer. In addition, this feature permits some degree of generic searching by computer for structures which have specified stereochemistry.

The representations of stereoisomers have evolved in scope and complexity from relatively simple beginnings to the sophisticated, highly systematized treatment in the present CAS Registry System. A brief account of developments in the stereochemical field is presented as background information for the major part of this paper, which deals with the present handling of stereoisomer representations.

STEREOCHEMICAL REPRESENTATIONS IN REGISTRY I AND II

Parallel to the development of the specifications, design, and early operation of the Registry System, research was carried out to develop a method for generating unique machine representations of stereoisomers in which the steric environment of each stereogenic atom was described.[3,4] Although the resulting method did meet the stated objective, it was judged unsuitable for the entire CAS Registry.

Attention was then turned to the steric description of a substance as a whole, as provided in the original literature. This became the basis of stereochemical representation in the Registry System. To accommodate descriptive terms for stereochemistry, a separate segment was added to the atom-bond description, or connection table, in the computer structure record.

Initially these descriptive terms consisted of (1) conventional prefixes commonly used in chemistry, illustrated by the examples in Figure 1; (2) names of substances that imply the stereochemistry of the substance, such as those shown in Figure 2; and (3) the single asterisk (*) to denote that stereochemistry was shown in the structural diagram in the original article but was not described in the text.

In the early system, machine registration of stereoisomers was based on matching the unique connection table[5] and stereochemical descriptor for a given structure against the Registry Master Structure File. Registration was completed entirely by machine if (1) there was an exact match on file for both the unique connection table and stereochemical descriptor (with the exception of the asterisk) or (2) the unique table was different from any other on file. When there was an exact match on the unique table but not on the stereochemical descriptor, registration was completed only after review by a chemist who determined whether or not the input substance was new. The latter case resulted when the stereochemical descriptors were synonymous, e.g., *d* and (+); D and *R*.

The first stereochemical program enabled CAS to provide an operable registry system that both accommodated the substances registered at that time and could be expanded as the scope of the system expanded. However, the disadvantages included an uncontrolled vocabulary, total lack of machine editing and validation, and machine registration of stereoisomers only with manual intervention.

Building a standardized, controlled descriptor vocabulary began soon after installation of the original Registry System. The natural product areas were considered first, since most substances of this type are stereospecific. For each class of steroids, alkaloids, terpenes, and carbohydrates, word roots were established which implied specific stereochemistry at several locations in a basic unit. Examples of these terms and their structural equivalents are shown in Figure 3. In addition to the sets of descriptors for these four classes, common stereochemical terms were included in the control file for descriptors. These included such terms as ANTI, CIS, E, ENDO, EXO, (R), (S), SYN, TRANS, Z, (+), (−), and (±).

For many natural products it was necessary to use prefixes with the terms to indicate the presence of additional or modified stereochemistry. For ring systems, Ring Index[6] numbering was used to generate various alphanumeric prefixes which were used with these descriptors, for example, 3α-PREGN and 3β,7β-OLEANANE. However, since the prefix terms were not a controlled vocabulary, their use required that the complete descriptor be considered potentially ambiguous and subject to chemist review.

With the establishment of a list of valid stereochemical descriptors, a machine-editing program became possible and indeed essential. Such a program was installed in 1968 as a part of the second version of the Registry System, called Registry II. An important feature of this program was that once the descriptor had been machine validated, complete machine registration of structure representations containing such valid descriptors was accomplished without a manual review step.

The experience gained in the development and use of stereochemical representations in Registry I and Registry II indicated the need for a more extensive, more exactly defined

† Presented at the 170th National Meeting of the American Chemical Society, Chicago, Ill., Aug 25, 1975.

* To whom correspondence should be addressed.

allo, altro, gluco, . . .

anti, syn

cis, trans

d, l, dl

endo, exo

R, S

Figure 1. Examples of prefixes used as stereochemical descriptors in Registry I.

dl-aldosterone

brucine

10-bromomorphinone

16α-iodocholesterol

d-isopentylamine fumarate

Figure 2. Examples of substance names used as stereochemical descriptors in Registry I.

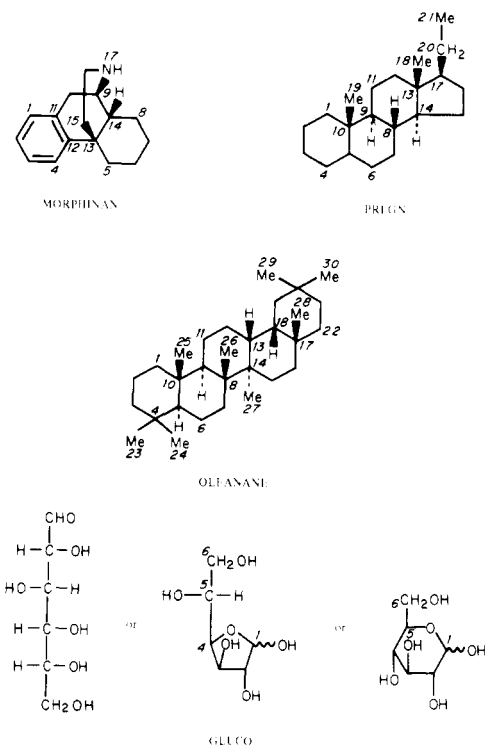


Figure 3. Examples of Registry II natural product terms and their structural equivalents.

system of stereochemical descriptors. It is this detailed system, and operations associated with it, that is the major topic of this paper.

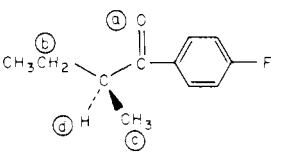
STEREOCHEMICAL REPRESENTATIONS IN THE CURRENT SYSTEM

In 1970, work was begun on analyzing the Registry II descriptors. Many of the descriptors were found to be quite complex and to consist of more than one type of information. In most cases these complex descriptors had been added to the file in association with a code indicating that the machine edits for valid text descriptor terms were to be bypassed. Whenever one of these descriptors was encountered, machine registration could not be completed until a chemist had reviewed that descriptor. The explicit identification of each type of information was thought to be a means for allowing more complete machine editing and providing better control of the terminology. With this as a goal, it was decided to continue developing the stereochemical descriptor approach and to introduce for Registry III a system that would edit and validate all stereochemical descriptors and greatly expand automated registration of stereoisomers.

Identifier	Type
1:	Absolute
2:	Relative
3:	Optical
4:	Stereoparent
5:	Amino acid/carbohydrate
6:	Trivial name
7:	Coordination compound

Acceptable combinations - 1&2, 2&3, 3&4, 3&6, 3&7

Figure 4. Types of stereochemical information identified for use in Registry III.

Structural diagram	Stereochemical descriptor
	1:S

CA Index Name and Registry Number

1-Butanone, 1-(4-fluorophenyl)-2-methyl-
(*S*)- [35066-94-5]

Figure 5. Example of an absolute descriptor.

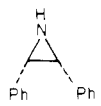
The various types of information allowed in Registry III stereochemical descriptors are given in Figure 4. The stereochemical descriptors used in Registry III are based on the system adopted for *Chemical Abstracts* during the 9th Collective Period. Since the rules and definitions of this system have been published in considerable detail,[7,8] only a general description of each type of information will be given in this paper.

In addition to the seven specific descriptor types given in Figure 4, an "NS" descriptor is used when no specific stereoisomer is reported and an asterisk (*) is used when the stereochemistry cannot be expressed by the descriptor conventions. An identifier is associated with each descriptor type to allow the application of different edits to different types of descriptors. The identifier also allows an automatic edit to determine if acceptable combinations of descriptors have been used for complex stereochemistry. The seven types of descriptors are discussed to provide a basis for explaining the automatic editing.

1. **Absolute descriptors** are based on the chiralty symbols *R* and *S* as defined by Cahn, Ingold, and Prelog in the Sequence Rule.[9] Figure 5 shows a stereoisomer of known absolute configuration which has only one chiral center.

The chirality is determined by first assigning priorities, according to a set of established rules, to the four groups directly attached to the chiral center. In this example, the highest priority group is labeled *a* and the lowest is labeled *d*. The molecule is then visualized so that the group of lowest priority, *d*, projects away from the observer. The directional sequence of the other three groups is determined, proceeding from highest priority group to lowest, or in this case, from *a*

Structural diagram | Stereochemical descriptor



2:CIS

CA Index Name and Registry Number

Aziridine, 2,3-diphenyl-
cis- [*1605-06-7*]

Figure 6. Example of a relative descriptor.

Structural diagram | Stereochemical descriptor

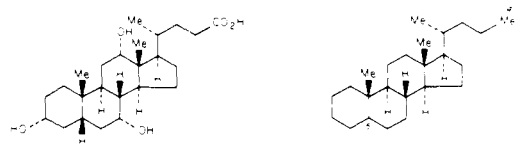
$CH_3(CH_2)_5CH{=}C{=}CHCO_2H$ 3:(-)

CA Index Name and Registry Number

2,3-Decadienoic acid
(-)- [*36143-66-5*]

Figure 7. Example of an optical descriptor.

Structural diagram | Stereoparent



Stereochemical descriptor - **4**:3α,5β,7α,12α.CHOL

CA Index Name and Registry Number

Cholan-24-oic acid, 3,7,12-trihydroxy-
(3α,5β,7α,12α)- [*81-25-4*]

Figure 8. Example of a stereoparent descriptor.

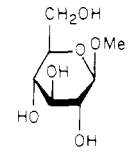
to *b* to *c*. If the direction is clockwise, the substance is labeled *R*; if the direction is counterclockwise, as in this example, it is designated *S*. Note that the stereochemical descriptor and the CA Index Name are based on the same system and indicate the same degree of stereochemical detail. This allows a verification, via computer edit, that the stereochemistry cited in the *Chemical Abstracts* Chemical Substance Indexes is the same as that recorded in the Registry stereochemical descriptor for a given stereoisomer.[10] The Registry Number assigned to this specific stereoisomer is shown in brackets following the name.

2. **Relative descriptors** are used to indicate diastereoisomerism. The appropriate descriptor is determined by comparing groups in a molecule to a reference plane or group in that same molecule. In Figure 6, the groups to be compared are the phenyls, and the reference plane is the plane containing the three-membered ring. The descriptor "cis" is used to indicate that the two reference groups are on the same side of the reference plane.

3. **Optical descriptors** indicate the direction of rotation of plane-polarized light. Optical rotation is not used as a descriptor when the absolute stereochemistry is known. In Figure 7, the allene unit is a chiral element, but since its absolute configuration is not known, the minus sign indicating negative rotation is used to identify the specific enantiomer reported.

4. **Stereoparent descriptors** are based on common names of classes of natural products. For example, "morphinan" is the stereoparent term used to describe the morphine family of alkaloids. Similarly, cholic acid, the stereospecific substance shown in Figure 8, is described by comparing it to the stereoparent cholane. The term "CHOL" implies the configurations shown in the stereoparent diagram, and the additional stereochemistry needed to describe the substance completely is expressed as a prefix.

Structural diagram | Stereochemical descriptor



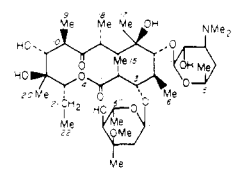
5:β-D-GLUCO

CA Index Name and Registry Number

β-D-Glucopyranoside, methyl [*709-50-2*]

Figure 9. Example of a carbohydrate descriptor.

Structural diagram | Stereochemical descriptor



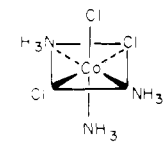
6:ERYTHROMYCIN

CA Index Name and Registry Number

Erythromycin [*114-07-8*]

Figure 10. Example of a trivial name descriptor.

Structural diagram | Stereochemical descriptor



7:OC-6-21

CA Index Name and Registry Number

Cobalt, triamminetrichloro-
(*OC*-6-21)- [*36294-67-4*]

Figure 11. Example of a coordination compound descriptor.

5. The **amino acid and carbohydrate descriptors** are based on the well-established terms used in these specialty areas. In Figure 9, "β-D-GLUCO" defines the absolute configurations of the five centers shown in the structural diagram.

6. **Trivial name descriptors** are used for natural products having complex stereochemistry. Generally, these substances do not belong to a family of compounds as is true for the type 4 stereoparents. In Figure 10, "erythromycin" implies the stereochemistry of the macrocyclic ring as well as the carbohydrate moieties. Prefixes are not used to describe stereochemical modifications of type 6 stereoparents.

7. The **coordination compound descriptors** are based on a technique first proposed by Petrarca and Rush.[4] In Figure 11, the descriptor consists of two parts, the system indicator "OC-6", and the geometry number "21". The system indicator "OC-6" means that an octahedral system of coordination number 6 is present. The geometry number "21" indicates the ligand arrangement about the central atom. A complete description of this system has been published by Brown, Cook, and Sloan.[8]

Additional information on all types of descriptors is given in Appendix A.

STEREOCHEMICAL DESCRIPTOR PROCESSING

A simplified flow diagram, shown in Figure 12, illustrates the processing steps for stereochemical descriptors in Registry

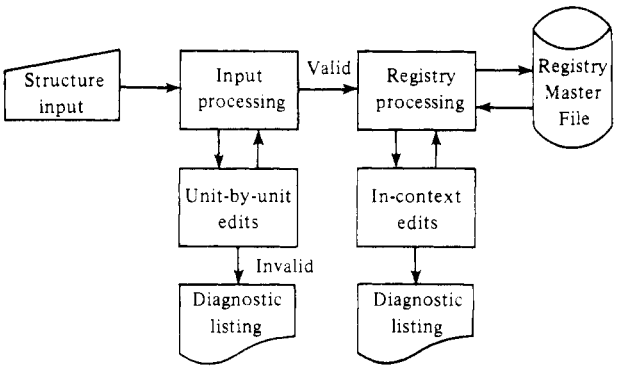


Figure 12. Processing of stereochemical descriptors in Registry III.

Input descriptor	Diagnostic message	Corrected descriptor
1:R3:(-)	Invalid Combination	1:R
1:1R3:CIS	Invalid Combination	1:1R2:CIS
2:1α.2α.3β	Punctuation Error	2:1α,2α,3β
2:(1α,3β,4β(R*))	Format Error	2:1α,3β,4β(R*)
3:ENDO	Incorrect Term	2:ENDO
4:3β,ANDROST	Punctuation Error	4:3β.ANDROST
4:311α.PREGN	Incorrect Locant	4:3β,11α.PREGN
5:β-D-GLCO	Incorrect Term	5:β-D-GLUCO
5:E-THREO-3(E)	Incorrect Prefix	5:D-THREO-3(E)

Figure 13. Types of errors detected during "unit-by-unit" editing.

III. The structural diagram and stereochemical descriptor are input during the first step. The input techniques used for this step have been described in detail by Dayton and Zamora.[11] During the input processing, the unit-by-unit edits are performed on individual descriptors. These edits verify that the format used is acceptable, check whether the terms used are valid for the indicated descriptor types, check the locants to ensure they fall within acceptable ranges, and check whether valid combinations of types are used for complex descriptors. A Backus–Naur Form (BNF) description of each of the descriptor types is used as a basis for the unit-by-unit edits.[12] This form of definition was chosen because it allows precise communication between the chemist and the computer specialist. A Backus–Naur Form definition of the type 5 descriptor is given in Appendix B.

Although the unit-by-unit edits verify that a descriptor is valid, they do not determine if it is the correct one for the substance. Valid descriptors are passed on to the Registry processing step. Descriptors containing errors detected during input editing are listed for a chemist to review, and the corrected descriptors are sent back through the input edits. Figure 13 shows a variety of descriptors containing errors detected by the unit-by-unit edits.

During registration, Unique Chemical Registry Records (UCRR) are generated and matched against the Registry Master File. In cases where the topology is matched, the stereochemistry must then be matched. Both those UCRR's which have topologies new to the file and those UCRR's which match on both topology and stereochemistry are completely processed by machine. Those UCRR's which match on topology but not on stereochemistry go to the In-Context Editing program. During this editing step the input descriptor is compared with all of the descriptors already on file for substances having the matched topology. The in-context edits determine descriptor-type compatibility as well as descriptor-content compatibility. Results of descriptor-type edits are illustrated by the first two examples in Figure 14. By comparing only the descriptor-type identifiers, shown in boldface type, it is determined that the type 1 descriptor input

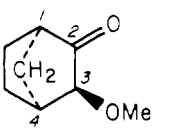
- Descriptor-type compatibility

Input	On file	Comment
1:R	**5**:D	Incompatible
2:ENDO	**2**:CIS	Compatible

- Descriptor-content compatibility

Input	On file	Comment
1:**1R**	1:**R**	Incompatible
1:**R**	1:**S**	Compatible
2:**ENDO**	2:**CIS**	Incompatible

Figure 14. Types of errors detected during "in-context" editing.



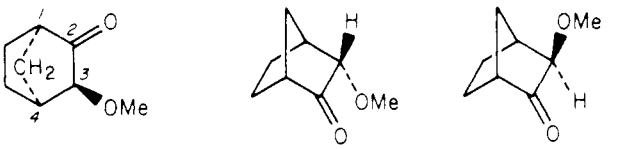
2:1α,3β,4α

Temporary identifier	Input descriptor	Diagnostic message
	⋮	
6791645Z	2:1α,3β.4α	Format error 0002
	⋮	

Figure 15. Report from "unit-by-unit" editing.

Problem number	Action code	Temporary identifier	Registry number	Stereo descriptor
71-1	A K M T U	6791645Z		2:1α,3β,4α
71-2	M T		53329-05-8	2:ENDO
71-3	M T		53329-06-9	2:EXO

DESCRIPTOR INCONSISTENCY



6791645Z [*53329-05-8*] [*53329-06-9*]

Figure 16. Report from "in-context" editing.

is not compatible with the type 5 descriptor on file, while the type 2 descriptor input is compatible with the same type on file.

Results of descriptor-content compatibility editing are illustrated by the last three examples in Figure 14. In the first example, the data shown in boldface type are incompatible owing to the inconsistency in the use of locants. This inconsistency arises from the fact that the input descriptor uses the locant "1" and the descriptor on file has no locant. In the second example the descriptors are compatible and are processed completely by machine. The descriptor contents in the third example are determined to be incompatible because the terms are derived from different stereochemical systems.

To illustrate these edits further, consider the hypothetical registration of a substance for which errors in the descriptor have been made intentionally. Figure 15 shows the structural diagram of the substance, its stereochemical descriptor, and a list from the unit-by-unit edits. The format error detected by the edits was corrected by changing the period between 3β and 4α to a comma. After the error has been corrected, the record passes the unit-by-unit edits and enters the registration process. When comparing the input descriptor to those on file for substances having the same topology, the edit detects an incompatibility. The report listing this incompatibility for a

Descriptor Type	Frequency†
1	65,871
2	166,981
3	34,549
4	135,894
5	141,012
6	4,716
7	77,274
* (unassignable)	132,377
Total	758,674

†as of November 1976

Figure 17. Frequency of stereoisomer descriptor types.

chemist to review is given in Figure 16.

The structural diagrams do not appear on the report, but are included in the table to facilitate the discussion. By comparing the input descriptor (1α,3β,4α) with those on file, the chemist determines that the input descriptor is synonomous with the ENDO descriptor already on file. To indicate that the input structure is identical with one on file, the chemist simply encircles action code T and Registry Number of the matched structure, as shown in Figure 16. Following input of this transaction, machine registration of the substance will be completed.

The CAS Registry files contain records for over 3.6 million chemical substances, of which more than 750 000 are specific stereoisomers. The frequencies of the different types of stereoisomers in the Registry files are given in Figure 17. For a routine production run, over 99% of the stereochemical descriptors are passed by the unit-by-unit edits. Approximately 75% are then passed by the in-context editing and are completely processed by machine without a chemist review.

SUMMARY

The techniques and stereochemical descriptors used in Registry III for registration of stereoisomers have been presented. In this system all stereochemical descriptors are edited and validated, and more stereoisomers can be registered without chemist intervention than was possible with previous versions of Registry.

ACKNOWLEDGMENT

CAS gratefully acknowledges the support of the National Science Foundation for portions of the research efforts that led to the development of these techniques.

APPENDIX A. REGISTRY III STEREOCHEMICAL TERMS AND SYMBOLS

Appendix A is a general description of the various stereochemical descriptors used in Registry III. All seven types of descriptors are defined in Backus–Naur Form. The Backus–Naur Form definition of allowed type 5 descriptors is given in Appendix B.

Type 1. Absolute Configuration

Symbols	R, S
Descriptors	Absolute configurational descriptors are derived by assigning an R or S symbol to the chiral element and associating a locant with the symbol when necessary. Only one symbol is used for the type 1 descriptor.
Examples	1:R; 1:S; 1:3S; 1:4(R)

Type 2. Relative Configuration

Terms and Symbols	E, Z, α, β, R*, S*, cis, trans, exo, endo, syn, anti
Descriptors	Relative configurational descriptors are assigned according to the procedures given in ref 6. Only certain combinations of terms and symbols are allowed for describing complex stereochemistry. For example, various combinations of exo, endo, syn, and anti are allowed; however, these terms may not be used in conjunction with α and β. Machine edits determine if the combinations are valid type 2 descriptors.
Examples	2:TRANS; 2:2-EXO,3-ENDO; 2:1α,3β(R*),4β

Type 3. Optical Rotation

Symbols	(+), (−), (±)
Descriptors	The sign of optical rotation is only used as a descriptor when the absolute stereochemistry is not given. It may be used in combination with types 2, 4, 6, and 7 terms.
Examples	3:(+); 2:TRANS3:(+); 3:(±)4:5α.ANDROST

Type 4. Stereoparent

Terms	Androst, Cholest, Gibbane, Ursane, Yohimban, . . . (there are presently 133 approved type 4 stereoparent terms; more can be added).
Descriptors	The terms are used either alone or with prefixes and suffixes to allow the specification of additional and modified stereochemistry as type 4 descriptors. The terms, prefixes, and suffixes are verified by machine edits.
Examples	4:.CHOLEST; 4:3β.YOHIMBAN; 4:3β,17β.ANDROST.3(β-D-GLUCO)

Type 5. Amino Acid/Carbohydrate

Terms and Symbols	D, L, DL, ALLO, ALTRO, . . . (the complete list of terms and the Backus–Naur Form definition of the type 5 descriptor are given in Appendix B).
Descriptors	See Appendix B.
Examples	See Appendix B.

Type 6. Trivial Name

Terms	Carpaine, Ergochrome-A, Erythromycin, 9S-Erythromycin, Picromycin, 8α-Remeramine, . . . (there are 543 approved type 6 terms; this list is updated periodically to accommodate additional natural products as their stereochemistry becomes established).
Descriptors	All natural products are not assigned type 6 descriptors. Most can be assigned either type 4 or types 1 and 2 descriptors. When a type 6 descriptor is used, only the terms, exactly as they appear on the dictionary, are allowed. Stereoisomeric forms of type 6 natural products are described by combining the natural product

Table I[a]

Amino acid/carbohydrate descriptors (AAC)		Examples
<TYPE 5>	::= 5 : <AAC>	
<AAC>	::= ALL, <DL>	ALL,L
	| <ITEMLIST>	
	| MESO	
<ITEMLIST>	::= <ITEM>	
	| <ITEMLIST>, <ITEM>	L,D,D
<ITEM>	::= <DL>	L
	| <DL>-(<ALPH>)	D-(R)
	| <CARBO>	GLUCO
	| <CARBOC>	
	| <CARBOSUF>	
<CARBO>	::= <TERM>	GLUCO
	| <PREFIX 1> <TERM>	α-D-GLUCO
	| <PREFIX 2> <TERM>	D-GLUCO
<PREFIX 1>	::= A - <PREFIX 2>	β-D-
	| B - <PREFIX 2>	α-D-
<PREFIX 2>	::= <DL> -	D-
<CARBOC>	::= <CARBO> - <CARBO>	D-GLYCERO-D-GLUCO
<CARBOSUF>	::= <CARBO> - <SUFFIX>	D-THREO-3(R)
	| <CARBOC> - <SUFFIX>	D-GLYCERO-D-GLUCO-3(R)
	| <CARBOSUF>, <SUFFIX>	D-THREO-3(R), 5(Z)
<SUFFIX>	::= <NUMBER> (<ALPH>)	3(R)
<ALPH>	::= R | S | E | Z	
<DL>	::= DL | D | L	
<TERM>	::= ALLO	
	| ALTRO	
	| ARABINO	
	| ERYTHRO	
	| CIS	
	| GALACTO	
	| GLUCO	
	| GLYCERO	
	| GULO	
	| IDO	
	| LYXO	
	| MANNO	
	| RIBO	
	| TALO	
	| THREO	
	| TRANS	
	| XYLO	

[a] 1. Any <ITEMLIST> whose only <ITEM>s are four or more identical values of <DL> is invalid, i.e., L, L, L, L should be ALL,L. 2. If several <SUFFIX> terms appear in a <CARBOSUF>, each <NUMBER> must be different and they must occur in ascending order. 3. The Greek characters α and β are represented by letters A and B, respectively, in Registry III stereochemical descriptors.

term with an appropriate stereochemical symbol and then adding the combination to the dictionary of approved type 6 terms.

Examples 6:ERYTHROMYCIN; 6:9R-ERYTHROMYCIN; 6:9S-ERYTHROMYCIN; 6:COBIN; 6:COBIN-F(β-D-RIBO)

Type 7. Coordination Compounds

Symbols T-4, SP-4, TB-5, SP-5, OC-6, TP-6

Descriptors Coordination compounds are described by selecting the appropriate symbol to indicate the basic geometry (T-4 = tetrahedral; TB-5 = trigonal bipyramidal; OC-6 = octahedral, etc.) and then adding a suffix to that symbol to indicate the relative and absolute nuclear stereochemistry and any ligand stereochemistry.[8]

Examples 7:T-4-R; 7:SP-5-24-C; 7:OC-6-32. (CIS)

APPENDIX B. BACKUS–NAUR FORM DEFINITION OF TYPE 5 DESCRIPTOR

In the Backus–Naur Form notation, the symbol "::=" means "is defined as", and the symbol "|" means "or". The relationship "x followed by y" is represented by "xy". The symbols "⟨" and "⟩" are used to enclose the names of classes or categories; any symbol not so enclosed indicates a literal occurrence of the symbol. Using these conventions, a one-digit or two-digit number can be defined as follows:

⟨number⟩::=⟨digit⟩|⟨digit⟩⟨digit⟩
⟨digit⟩::=0|1|2|...|9

Examples are given in Table I.

REFERENCES AND NOTES

(1) P. G. Dittmar, R. E. Stobaugh, and C. E. Watson, "The Chemical Abstracts Service Chemical Registry System. I. General Design", *J. Chem. Inf. Comput. Sci.*, **16**, 111–121 (1976).
(2) D. P. Leiter, Jr., H. L. Morgan, and R. E. Stobaugh, "Installation and Operation of a Registry for Chemical Compounds", *J. Chem. Doc.*, **5**, 238–242 (1965).
(3) A. E. Petrarca, M. F. Lynch, and J. E. Rush, "A Method for Generating Unique Computer Structural Representations of Stereoisomers", *J. Chem. Doc.*, **7**, 154–165 (1967).
(4) A. E. Petrarca and J. E. Rush, "Methods for Computer Generation of Unique Configurational Descriptors for Stereoisomeric Square-Planar and Octahedral Complexes", *J. Chem. Doc.*, **9**, 32–37 (1969).
(5) H. L. Morgan, "The Generation of a Unique Machine Description for Chemical Structures—A Technique Developed at Chemical Abstracts Service", *J. Chem. Doc.*, **5**, 107–113 (1965).
(6) A. M. Patterson, L. T. Capell, and D. F. Walker, "The Ring Index", 2nd ed, American Chemical Society, Washington, D.C., 1960.
(7) J. E. Blackwood and P. M. Giles, Jr., "*Chemical Abstracts* Stereochemical Nomenclature of Organic Substances in the Ninth Collective Period (1972–1976)", *J. Chem. Inf. Comput. Sci.*, **15**, 67–72 (1975).
(8) M. F. Brown, B. R. Cook, and T. E. Sloan, "Stereochemical Notation in Coordination Chemistry. Mononuclear Complexes", *Inorg. Chem.*, **14**, 1273–1278 (1975).
(9) R. S. Cahn, C. K. Ingold, and V. Prelog, "Specification of Molecular Chirality", *Angew. Chem., Int. Ed. Engl.*, **5**, 385–415 (1966).
(10) G. G. Vander Stouw, "Computer Programs for Editing and Validation of Chemical Names", *J. Chem. Inf. Comput. Sci.*, **15**, 232–236 (1975).
(11) A. Zamora and D. L. Dayton, "The Chemical Abstracts Service Chemical Registry System. V. Structure Input and Editing", *J. Chem. Inf. Comput. Sci.*, **16**, 219 (1976).
(12) P. Naur, Ed., "Revised Report on the Algorithmic Language ALGOL 60", *Commun. ACM*, **6**, 1–17 (1963).