A Line-Formula Notation System
for Coordination Compounds

PATRICIA M. MCDONNELL*

Office of Research and Development, U. S. Patent Office, Washington, D. C. 20231

and

ROBERT F. PASTERNACK**

Information Technology Division, National Bureau of Standards, Washington, D. C. 20234

The development of notation systems for representing organic structures in linear form has been a subject of considerable interest during recent years.[1-3] However, the development of similar methods for uniquely and unambiguously representing inorganic structures has received little attention. This paper describes an initial attempt to develop such a method for handling inorganic coordination compounds which will be compatible, insofar as possible, with existing practices in organic notation systems and which will be extensible to other types of inorganic compounds. The approach described here is independent of any particular set of seniority rules; for purposes of illustration in this paper, the principles of the Hayward system for organic structures have been utilized.

The class chosen for initial investigation in the development of a notation system for inorganic structures was coordination complexes. This class includes discrete aggregates—cations, anions, and neutral molecules—rather than endless chains, lattices, or networks. The handling of these aggregates requires the resolution of problems which are similar to those which must be solved in the development of methods for treating other types of inorganic structures but which do not occur in the handling of organic structures. Coordination centers (termed *Group M atoms* in this paper) exhibit variable coordination numbers and symmetries. Unique and unambiguous representation of these complexes requires the development of a method for expressing not only the relation between each attached group (generally termed a *ligand*) and the coordination center but also the relations among and within the various ligands themselves. Furthermore, if several Group M atoms are present in a given aggregate, their relations, one to another, must also be described.

REFERENCE–STRUCTURE APPROACH

Because of these requirements, it was felt that a stereochemical approach would prove to be most fruitful, and a table of reference structures (Table I) was prepared. This table contains idealized representations of postulated structures and shows the highest-order axes of rotational

symmetry and planes perpendicular to those axes in which ligand attachments lie. Each proposed coordination structure is assigned a two-character symbol: the coordination number of the fully coordinated Group M atom; and an uppercase letter, as the *symmetry designator*, which refers to the particular geometric distribution of attachments. Furthermore, each position of ligand attachment is assigned a single lowercase letter, termed a *locant designator*. The symmetry designators, although assigned somewhat arbitrarily, reflect frequency of occurrence of individual configurations. The assignment of locant designators to the reference structures was made in accordance with symmetry characteristics and with reference to highest-order axes of rotation.

The use of these locant designators in the actual ciphering of a given aggregate depends upon a set of arbitrary seniority rules. With the aid of such rules, the relative seniorities of ligands about each Group M atom are determined. The appropriate reference structure is chosen to act as a type of space-coordinate system. The aggregate is then allowed to undergo any form of motion permitted for a rigid body until it is aligned with the reference structure. Insofar as possible, the most senior ligand attachment is made to coincide with locant designator a; the next most senior, with locant designator b; etc. By treating the coordinated bonds as points on a rigid body in this manner, it is possible to describe fully in the cipher all types of isomerism pertaining to coordination centers.
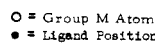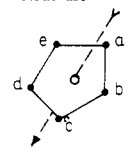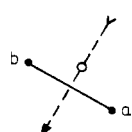
CIPHERING RULES AND EXAMPLES
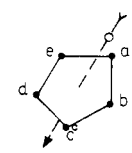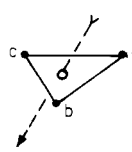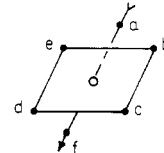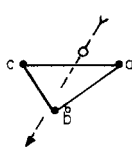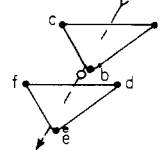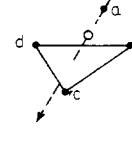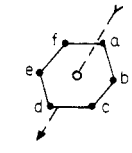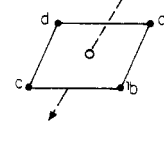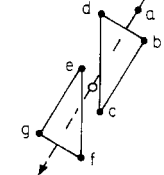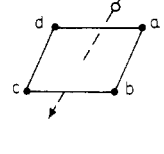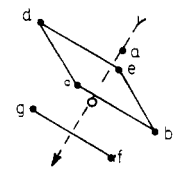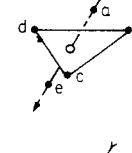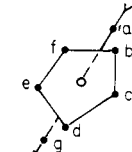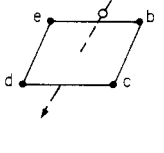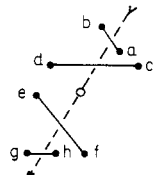
The approach to ciphering inorganic complexes is illustrated in this section through the utilization and adaptation of the Hayward enumeration and notation rules for organic compounds but is by no means restricted to implementation with this particular system.

Since several coordination centers may be present in one aggregate, the Group M atoms are assigned index numbers according to some arbitrary method, *e.g.*, by a set of rules similar to the Hayward rules for enumeration of parent hydrocarbons. Group M atoms are ciphered first, ligands are ciphered next as independent entities, and bonds between Group M atoms are ciphered last. Various types of stereochemical information and enclosing marks are introduced. Membership marks, $\in \ni$, are used for information concerning Group M atom index numbers,

---

Table I
Table of Reference Structures

Structure	Cipher symbol	Common name	Structure	Cipher symbol	Common name
O = Group M Atom ● = Ligand Position				5C	Pentagonal plane
	2A	Linear chain		5D	Pentagonal pyramid
	2B	Bent chain		6A	Octahedron
	3A	Trigonal plane		6B	Trigonal prism
	3B	Trigonal pyramid		6C	Hexagonal plane
	4A	Tetrahedron		7A	Octahedron with seventh position beyond one face
	4B	Tetragonal or square plane		7B	Trigonal prism with seventh position beyond one face
	4C	Tetragonal pyramid		7C	Pentagonal bipyramid
	5A	Trigonal bipyramid		8A	Dodecahedron with triangular faces
	5B	Tetragonal pyramid			

Table I
Table of Reference Structures (Continued)

Structure	Cipher symbol	Common name	Structure	Cipher symbol	Common name
	8B	Tetragonal antiprism		8D	Cube
	8C	Trigonal prism with extra bonds at extremities of unique axis		9A	Trigonal biantiprism

coordination numbers, and symmetry designators. Ciphers for ligands and for connections between Group M atoms are enclosed within double pointed brackets, « ». Single brackets, [ ], are inserted in ligand ciphers immediately following atoms through which bonding occurs and contain the index numbers of the appropriate Group M atoms and the correct locant designators as determined from Table I.

The following structures and ciphers are presented to illustrate the approach to ciphering inorganic coordination complexes described above. The principles employed correspond to the revised Hayward rules which are to be published soon. According to these rules, ciphers for ions are enclosed in square brackets and followed. by the encircled appropriate sign and a number indicating the magnitude of the charge. Just as for substituents in the Hayward method the ligands are ciphered in order of increasing seniority in the following examples.

Example 1. Dioxopentafluorouranate(VI) Ion, $[UO_2F_5]^{3-}$.— This complex has a pentagonal bipyramid configuration, corresponding to reference structure 7C; the uranium atom is in a plane with five fluorine atoms, and an oxygen atom is located on each side of the plane[4] (Figure 1). According to the revised Hayward rules, if two functional substituents have the same number of atoms connected in the same manner, that substituent is senior which has the atom of higher atomic number on the occasion of first difference upon leaving the skeletal portion of the molecule; therefore, the monoatomic ligand F, atomic
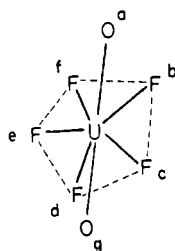
number 9, is senior to the monoatomic ligand O, atomic number 8. However, the only possible alignment of the complex with the reference diagram is such that O atoms correspond to positions a and g and F atoms with b through f. The Hayward cipher symbols for uranium and monocoordinated oxygen are ur and V, respectively. The proposed cipher is

$$[ur\in1,7C\ni\ll V[1a,1g]\gg2\ll F[1b,1c,1d,1e,1f]\gg5]\ominus3 \qquad (1)$$

Example 2. Dichlorodiammineplatinum $[PtCl_2(NH_3)_2]$. There are *cis* and *trans* forms of this square-planar complex[5] (Figure 2). According to the revised Hayward rules,
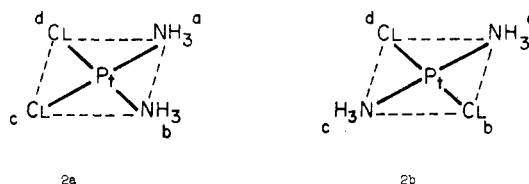
Figure 2.

that functional substituent is senior which contains the larger number of atoms; therefore, polyatomic NH_3 is senior to monoatomic Cl. The cipher symbols for chlorine and platinum are cl and pt, respectively. The proposed ciphers, which differ only with regard to locant designators, are

$$pt\in1,4B\ni\ll cl[1c,1d]\gg2\ll N[1a,1b]H3\gg2 \qquad (2a)$$

$$pt\in1,4B\ni\ll cl[1b,1d]\gg2\ll N[1a,1c]H3\gg2 \qquad (2b)$$

Example 3. [(Ethylenedinitrilo)tetraacetato]cobaltate-(III)Ion, $[Co(C_{10}H_{12}N_2O_8)]^-$.— At least three isomers of this ion (which is generally considered to be an inorganic complex even though it contains a large organic moiety) have been reported: two mirror-image octahedral arrangements and one trigonal prism configuration[6] (Figure 3). In order to obtain ciphers which are unique as well as unambiguous, the assignment of locant designators must be based upon seniority of ligand atoms through which connection to the central atom occurs. According to the revised Hayward

Figure 1.

Figure 3.

rules, a skeletal heteroatom is senior to an atom of a functional substituent; therefore, the N skeletal hetero-atoms are senior to the O functional substituent atom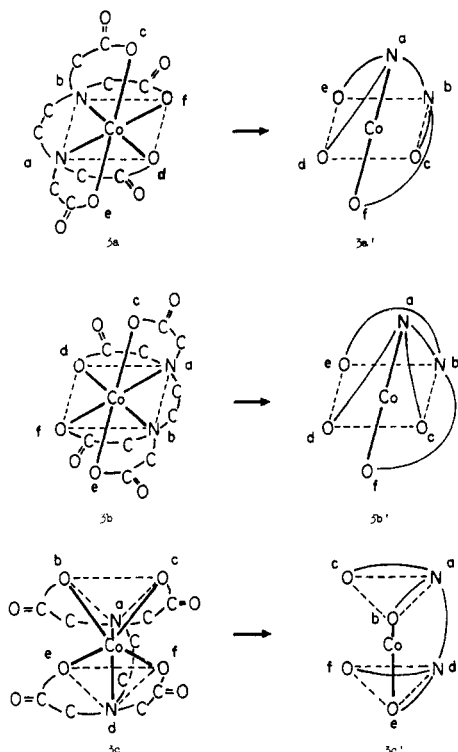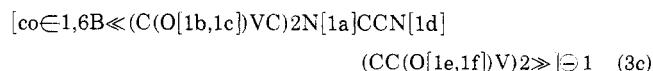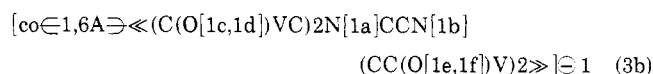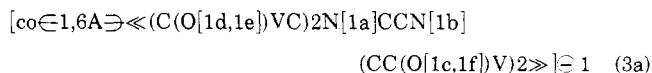s. In the assignment of locant designators, the complexes as drawn must be rotated until they are aligned with the appropriate reference structures such that, insofar as possible, the most senior ligand connection coincides with locant designator a; the next most senior, locant designator b; etc. Structures 3a', 3b', and 3c' illustrate the orientations into which the original representations were rotated when the locant designators were assigned. In the cases of the octahedral complexes 3a and 3b, the structures can be rotated so that either equivalent N atom of the symmetrical ligand coincides with locant designator a and the other equivalent N atom with locant designator b of reference structure 6A. Since the N atoms in the trigonal prism configuration 3c lie in different planes with respect to the highest-order axis of rotational symmetry, one of them is made to coincide with locant designator a and the second then coincides with locant designator d of reference structure 6B. The Hayward symbol for cobalt is co and the cipher for the $(OOCCH_2)_2NCH_2CH_2N(CH_2COO)_2^{4-}$ anion is $[(C(O)VC)2NCCN(CC(O)V)2[\ominus 4$. The ciphers for the complexes are

$[co\in1,6A\ni\ll(C(O[1d,1e])VC)2N[1a]CCN[1b]$

$(CC(O[1c,1f])V)2\gg]\ominus 1$ (3a)

$[co\in1,6A\ni\ll(C(O[1c,1d])VC)2N[1a]CCN[1b]$

$(CC(O[1e,1f])V)2\gg]\ominus 1$ (3b)

$[co\in1,6B\ll(C(O[1b,1c])VC)2N[1a]CCN[1d]$

$(CC(O[1e,1f])V)2\gg]\ominus 1$ (3c)

DISCUSSION

The approach outlined in this paper permits the coding of coordination complexes in a linear cipher suitable for computer storage. The cipher, based on the Table of Reference Structures and an arbitrary set of seniority rules for ligands, allows for the inclusion of all known stereochemical information. In extending this system to those cases in which no such information but only the line formula is given, the cipher symbols for the Group M atoms and ligands are noted in the same order as a fully determinate structure but without insertion of any symmetry or locant designators. For example, if only line formulas has been given in each of the above illustrations, the ciphers are

1. $[ur\ll V\gg 2\ll F\gg 5]\ominus 3$
2. $pt\ll cl\gg 2\ll NH3\gg 2$
3. $[co\ll(C(O)VC)2NCCN(CC(O)V)2\gg]\ominus 1$

The same computer program can therefore be used to search for a single isomer of a particular coordination compound and also for any complex containing a specified Group M atom and/or ligand. The latter type of search is accomplished by simply programming the computer to ignore the occurrence and contents of membership marks and single brackets.

It should also be noted that the cipher for an organic ligand takes the same form as the cipher for the same connected set of atoms occurring as an independent entity; it differs only by the insertion within single brackets of information concerning the stereochemistry of the coordination complex. This similarity permits one to use the same computer program to search for substructures in organic compounds and to search for substructures in organic ligands contained in inorganic coordination complexes.

CONCLUSION

This approach to ciphering inorganic complexes is independent of any particular notation system. The reduction of the approach to actual practice requires the utilization of arbitrary seniority rules for the assignment of locant designators and the adoption of a set of symbols for representing the atoms and the connections between atoms. The above examples illustrate how this approach can be utilized with the symbols and conventions of the Hayward system and with a set of seniority rules based upon that system. The approach is, however, readily adaptable to use with other notation systems.

Additional seniority and ciphering rules based upon the principles of the Hayward system are currently under development for the treatment of more complicated coordination complexes, e.g., polynuclear complexes, and will appear in another report. Consideration is also being given to the use of the Table of Reference Structures in the treatment of other types of inorganic structures, such as boron hydrides and heteropoly anions.

REFERENCES

(1) "Rules for I.U.P.A.C. Notation for Organic Compounds," John Wiley and Sons, Inc., New York, N. Y., 1961.

(2) H. W. Hayward, "A New Sequential Enumeration and Line Formula Notation System for Organic Compounds," Patent Office Research and Development Report No. 21, Department of Commerce, Washington, D. C., 1961.

(3) W. J. Wiswesser, "A Line-Formula Chemical Notation," Thomas Y. Crowell Co., New York, N. Y., 1954.

(4) F. A. Cotton, and G. Wilkinson, "Advanced Inorganic Chemistry: A Comprehensive Text," Interscience Publishers, Inc., New York, N. Y., 1952, p. 525.

(5) Ref. 4, p. 519.

(6) D. W. Cooke, Y. A. Im, and D. H. Busch, *Inorg. Chem.*, 1, 14 (1962).

PROCEEDINGS IN PRINT

The Aerospace Section of the Science-Technology Division of Special Libraries Association is undertaking publication of a new reference service for scientists, engineers, and librarians. Entitled "Proceedings in Print," each bimonthly issue of approximately 50 pages will provide full bibliographic data about proceedings in as wide a subject range and span of years as is pertinent to current technology in the aerospace fields.

Subscription orders for "Proceedings in Print" are now being accepted at $20 per year. For volume one, prepayment is required. Checks should be made payable to Aerospace Section, Special Libraries Association, and mailed to Maurice Rahilly, Business Manager, "Proceedings in Print," Avco Corp., Research and Advanced Development Division, 201 Lowell St., Wilmington, Mass.

ENGINEERING GRADUATES

Job opportunities for last year's crop of engineering graduates were remarkably bright in view of employment difficulties for some categories of experienced engineers. This was the highlight of a new survey just published by the Engineering Manpower Commission of Engineers Joint Council in which colleges were asked to provide information on the post-graduation plans of engineering graduates who have just received their bachelor's degrees.

The steady increase in the number of graduates going on for advanced degrees is also part of the over-all picture. In the last survey, conducted in 1961, 14% were reported entering graduate school compared with 17% in the current survey. An additional 5% of the graduates were known to be intending to pursue post-graduate training in one form or another while employed. Combined, 21.7% will pursue graduate studies.

The survey was based on responses from 134 institutions representing over one-half of the total number of engineering schools in the nation. The report, "The Placement of Engineering Graduates—1964," is available through the Engineering Manpower Commission, 345 E. 47th St., New York, N. Y. 10017, at $1.00 per copy.

SLA DIRECTORY OF MEMBERS

Special Libraries Association has published its biennial "SLA Directory of Members as of July 15, 1964." Copies are available from the Association at a member's price of $2.50 and a non-members' price of $10.

The 144-page Directory is an alphabetical listing of approximately 5,962 members of the Association with addresses, titles, company names or home addresses, and other data.

TESTING OF DOCUMENT RETRIEVAL SYSTEMS

Herner and Company has formed an evaluation section to undertake tests of the effectiveness of document retrieval systems utilizing techniques of the type developed in England by Cyril Cleverdon.

Test programs are carried out under the immediate supervision of F. W. Lancaster, former member of Cleverdon's Cranfield team.

Any type of document retrieval system can be tested with techniques which involve measurement of response to a number of "synthetic" and "real-life" questions.

System effectiveness is measured in terms of the percentage of known relevant documents in the collection retrieved in searching, and the percentage of total search output judged by the original questioner.

Test programs are designed to permit analysis which can identify defects and sources of failure, as well as indicating system efficiency. A detailed report summarizing findings on system behavior, drawing useful comparisons with other systems, and making recommendations on means of improving the system or its operation, is the final product of a Herner test program.

Supplementing its test programs, Herner and Company is organizing a series of one-week courses on systems evaluation, beginning in the spring of 1965. Lecturers will include Cyril Cleverdon, consultant to Herner and Company, and Robert Fairthorne, Saul Herner, and F. W. Lancaster, of the Herner and Company staff.

SOVIET–UNITED STATES EXCHANGE OF SPECIAL LIBRARIANS

A $27,100 grant by the National Science Foundation has been made available to Special Libraries Association for an exchange of visits between Soviet and United States special librarians in the science-technology field. This exchange is one of the 19 scientific and technical exchange programs provided for under an agreement signed by the United States and the Soviet Union in Moscow, Feb. 22, 1964.

Six SLA members and an interpreter will leave for the Soviet Union, probably in the spring of 1965: John P. Binnington, Brookhaven National Laboratory, Upton, N. Y.; William S. Budington, John Crerar Library, Chicago, and SLA President; Mrs. Irma Johnson, Charles Hayden Memorial Library, Massachusetts Institute of Technology, Cambridge; Dr. Frank E. McKenna, Air Reduction Company, Inc., Murray Hill, N. J.; Gordon Randall, Thomas J. Watson Research Center, IBM, Yorktown Heights, N. Y.; Winifred Sewell, National Library of Medicine, Bethesda, Md.; and interpreter, Boris Gorokhoff, Washington, D. C.