

Journal of Organometallic Chemistry, 152 (1978) 101–109
© Elsevier Sequoia S.A., Lausanne — Printed in The Netherlands

COMPUTER GRAPHICS DETERMINATION AND DISPLAY OF STEREOISOMERS IN COORDINATION COMPOUNDS

F. CHOPLIN *, R. DORSCHNER, G. KAUFMANN

Laboratoire de Modèles Informatiques appliqués à la Synthèse (ERA 671), ULP, Institut Le Bel, 4, rue Blaise Pascal, 67000 Strasbourg (France)

and W.T. WIPKE

Department of Chemistry, University of California, Santa Cruz, California 95060 (U.S.A.)

(Received December 9th, 1977)

Summary

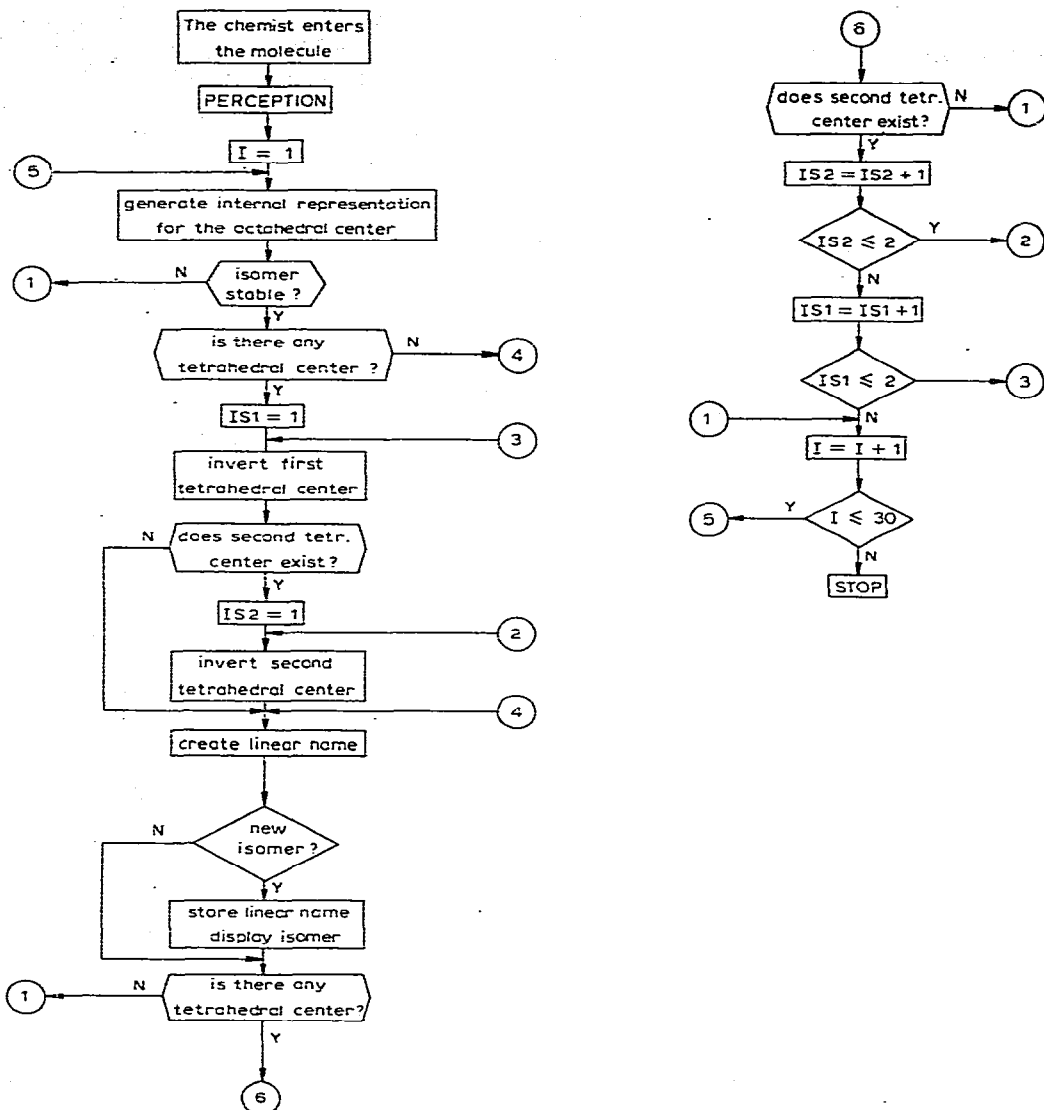
An interactive computer program has been written to determine and display stereoisomers of coordination compounds, which include one trigonal bipyramid or one octahedral stereocenter, and up to six tetrahedral centers. Recognition of isomers is made via a stereodescriptor calculated by means of a naming algorithm.

Introduction

The determination of the number of isomers and their structures in coordination compounds, by using different models or methods [1], is often prone to errors, even in most simple cases. Computer determination appears to be more reliable [2], however programs are limited since they require coding of the input and interpretation of the output by the chemist. It appears interesting to circumvent these problems by use of interactive computer graphics, which has been employed successfully in other areas of chemistry [3], and we describe in this paper an interactive program designed to determine and display stereoisomers in coordination compounds which include one octahedral or trigonal bipyramid stereocenter, and up to six tetrahedral centers.

Method

The present version of the program runs on an UNIVAC 1110 computer, and uses a DEC GT40 as a graphic display; several parts are derived from the version 2.0 of the SECS system [4]. All the operations performed by the program are summarized on Flowchart 1, and are detailed as follows:

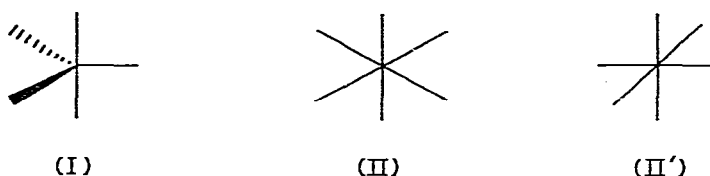


Flowchart 1. Flowchart of the program given in the specific case of a compound with an octahedral center and up to two tetrahedral stereocenters. The flowchart is identical in the case of a TBP center (the only change is in the test for $I \leq 20$), and can be easily modified to accommodate up to six stereocenters.

Input

First, the chemist enters the molecule by drawing it on the display. Bonds are drawn by means of a light pen, and atom types are specified either via the light pen (for C, H, N, O, P, S, X), or with a teletype. As soon as a new atom is entered, it is given an input number by the program. A trigonal bipyramid (TBP) is entered as represented by structure I, whereas an octahedral complex can be drawn as shown by structures II or II'. Stereochemistry for an octahedral center is specified by the user drawing the bonds as represented by II or II',

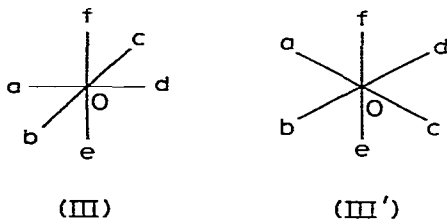
therefore hashed or wedged bonds are not needed on input. For tetrahedral



centers, stereochemistry is specified in the usual way [5], by indicating that a bond is above (represented by \triangleleft : UP) or behind the plane of the display (represented by \triangleleft : DOWN). Then the program builds a computer representation of the molecule, which consists of a connection table [5], and of the set of the display coordinates of the atoms. The connection table contains, for each atom, an attachment list with the input numbers of the connected atoms. A program is used to generate a structural diagram for any connection table, and make the correspondance between the internal representation and the display.

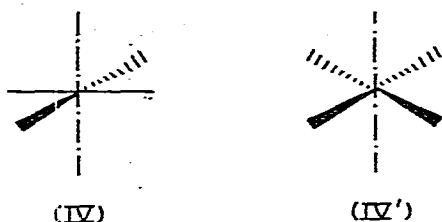
Perception

In the next stage, the program performs a perception of the structure, i.e. an analysis of all relevant topological features: it recognizes the presence and the size of rings [6], and the configurations of the stereocenters. Recognition of the configuration of TBP centers has already been described [7], and we indicate how this operation is performed in the case of an octahedron. Let us consider the octahedral center O in III: the internal representation for this atom includes a list which contains the input numbers of the attached atoms, If a, b, ... e, f are these numbers, the attachment list for O is: O(a, b, c, d, e, f).



When drawing the molecular skeleton III, the user is required to put f, O, and e on an approximate straight line, and one of the three straight lines, a0d, b0c or e0f must be roughly parallel to the O_y vertical axis of the display. In III, we have also a0d parallel to the O_x axis. By using these conventions, the program considers that Oc and Ob in III' are in front of the plane of the display, so is Ob in III and that Oa and Od in III' are behind this plane. This recognition is done by straightforward tests on the coordinates. Once this analysis is ended, the structure is redisplayed as represented by IV and IV', in order to show the correct stereochemistry. The bonds parallel to O_y are displayed in a special manner, to differentiate them from other bonds in complex diagrams. This holds also for axial bonds in TBPs.

The attachment list of O is reordered to represent correctly the configuration of this atom: attachments are associated by pairs, each one containing the numbers of two atoms which are on an axis passing through O and *trans* with respect



to the plane defined by 0 and the other four atoms. Atoms connected to 0 by wedged bonds are given $+Z$ coordinate by the program, and those connected by hashed bond are given $-Z$; then, using elementary vector analysis on the 3D structure, the program reorders the attachment list of 0 in the following way: one goes in clockwise order from first atom of pair 1 toward first atom of pair 2, when looking from the second toward the first atom of pair 3. Thus, if we consider structure III', the attachment list of 0 becomes:

$0(a, b, c, d, e, f) \rightarrow 0(ac, db, fe) \rightarrow 0 < ac, bd, fe > *$

The last list which is ordered to represent the configuration, is equivalent to 23 other lists. The described configuration is unchanged as far as the sum of the permutations between pairs and inside pairs is even.

Tetrahedral atoms are treated in an analogous way [5], and at the end of this step, the connection table of the molecule represents accurately the configurations of the stereocenters.

Generation of isomers

The next step is the generation of the isomers. The internal representation of each isomer is created by reordering a copy of the initial connection table. For a TBP or an octahedral center, this is done by using a table of all the possible isomers, which describe the permutations to be performed on the attachment list of this center, as detailed by Bennett [2]. Each tetrahedral stereocenter is inverted separately, by exchanging two entries in the corresponding attachment list.

Stability

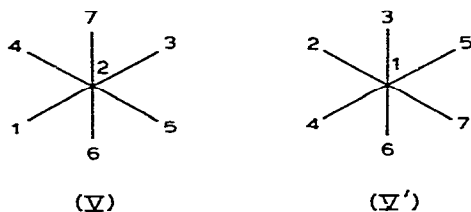
The stability of the current isomer is checked, by using standard rules: in an octahedral complex, chelating groups must be *cis*; in a TBP, rings with less than seven atoms cannot span two equatorial positions. Other restrictive rules can be included in the program if necessary.

Comparison of isomers

The connection table of the current isomer has been generated, but it is necessary to compare it to already obtained isomers. This is not directly possible by comparison of the connection tables, since they do not include any information about equivalence between ligands. The program creates, for each isomer, a linear name which allows fast comparison between structures. This name is obtained by use of the SEMA algorithm [7,8] and it gives a full description of the molecular topology. Since we deal with stereoisomers, we keep only the

* We use the parentheses to indicate an unordered list, and brackets for an ordered list.

part of the name describing the stereocenters. In order to represent the configurations, the algorithm calculates, for each stereocenter, a stereodescriptor which is equal to 1 or 2 for a tetrahedral center, and varies from 1 to 20 for a TBP [7]. We describe briefly how such descriptors are obtained for an octahedral structure. The aim of the naming algorithm is to find an atom numbering independent of the input numbering: this is done by creating a name for each possible numbering, and keeping the one corresponding to the lowest name, regarded as a number. A stereodescriptor is included in the name, and is obtained in the following way: let us consider an octahedral complex V, with the input atom numbers as indicated. The attachment list for atom 2 is (45, 13, 76). This list can be written by use of a SEMA numbering (such as given by V') and becomes

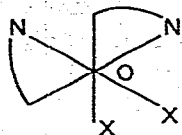


(27, 45, 36). This list is then reordered into a canonical form by replacing each number by its relative rank in the list, which gives (16, 34, 25). By use of an even number of permutations, the list is reordered so that atom 1 occupies the first entry, and atom 2 occupies the third entry. If atom 2 occupies the second entry, then atom 3 must be in third entry. We thus obtain (16, 25, 43), which is compared to each item of a reference array describing the 30 theoretical isomers of an octahedron with the same conventions. The rank of the item identical to the list under scrutiny is returned as a stereodescriptor to the octahedral center. All the descriptors obtained from the trial numberings are stored in a set S^* , which thus gathers all the descriptors corresponding to a given configuration. Then the algorithm picks up the lowest member d of S as a stereodescriptor for the given center ($1 \leq d \leq 30$), and stores it in the linear name. This one is then compared to those of already found isomers, and it is deleted if it is identical to another name.

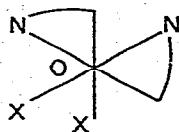
A specific problem arises to recognize an enantiomeric relationship between two isomers. In order to allow the program to perform this operation, the reference array is built so that the difference between the ranks of two items describing enantiomers is equal to 15.

To illustrate the usefulness of such a convention, let us consider two enantiomers VIa and VIb. For VIa, the SEMA algorithm leads to the set $S_a = \{11, 13, 21, 23\}$, and retains $d_a = 11$ as stereodescriptor for 0 (vide supra for the definition of S). When examining VIb, the program builds the set $S_b = \{6, 8, 26, 28\}$, retains $d_b = 6$, and compares d_b to the previously obtained descriptors. In order to determine if a structure (of descriptor d) and VIb are enantiomers, the program checks if $d + 15$ (or $d - 15$ if $d > 15$) belongs to S_b : in this case, when comparing VIa and VIb, we have: $11 + 15 = 26 \in S_b$, and the program concludes that VIa and VIb are enantiomers. It must be noted that the S sets are not saved:

* A set is internally represented by a 36 bits word. See ref. 4b.



(VIa)



(VIb)

only the current one is used, and it is replaced by a new one as soon as another isomer is generated.

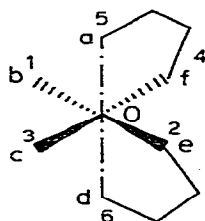
In a similar way, an inactive center can be detected by checking that its descriptor d and $d + 15$ (or $d - 15$ if $d > 15$) belong to the same set S .

It must be noted that faster ways can be used to compare isomers in simple cases [2]. However, the present system is very general, and the naming part of the program does not need to be modified were several stereocenters involved. Only the part dealing with the generation of the isomers has to be changed for this purpose.

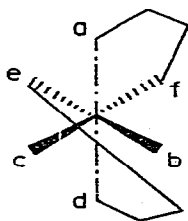
Display of isomers

At this point, the current stereoisomer is a new one, and it can be displayed. Initial atom coordinates are saved at the beginning of the process, and they are not modified for the atoms which are not directly bonded to the octahedral or TBP center, so that these atoms are always displayed at the same place. For tetrahedral centers, the new stereochemistry is indicated by the program which simply modifies the UP and DOWN characters of the attached bonds. The ligands of the central atom are given the coordinates of the new position they occupy.

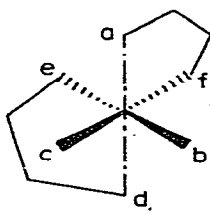
For example, let us consider structure VII, where atom numbers are indicated by letters a, b, ... e, f, and the positions they occupy on the display are referenced



(VII)



(VIII)



(VIII')

by numbers, which correspond also to the entries in the attachment list of atom O:

1	2	3	4	5	6
(b	e	c	f	a	d)

Another stereoisomer will be described by the following list:

1	2	3	4	5	6
(e	b	c	f	a	d)

which means that atom e occupies now position 1, etc... This new stereoisomer will be displayed as represented by structure VIII, which is equivalent to VIII'. Generally, structures such as VIII are interpreted very easily by the chemist. In

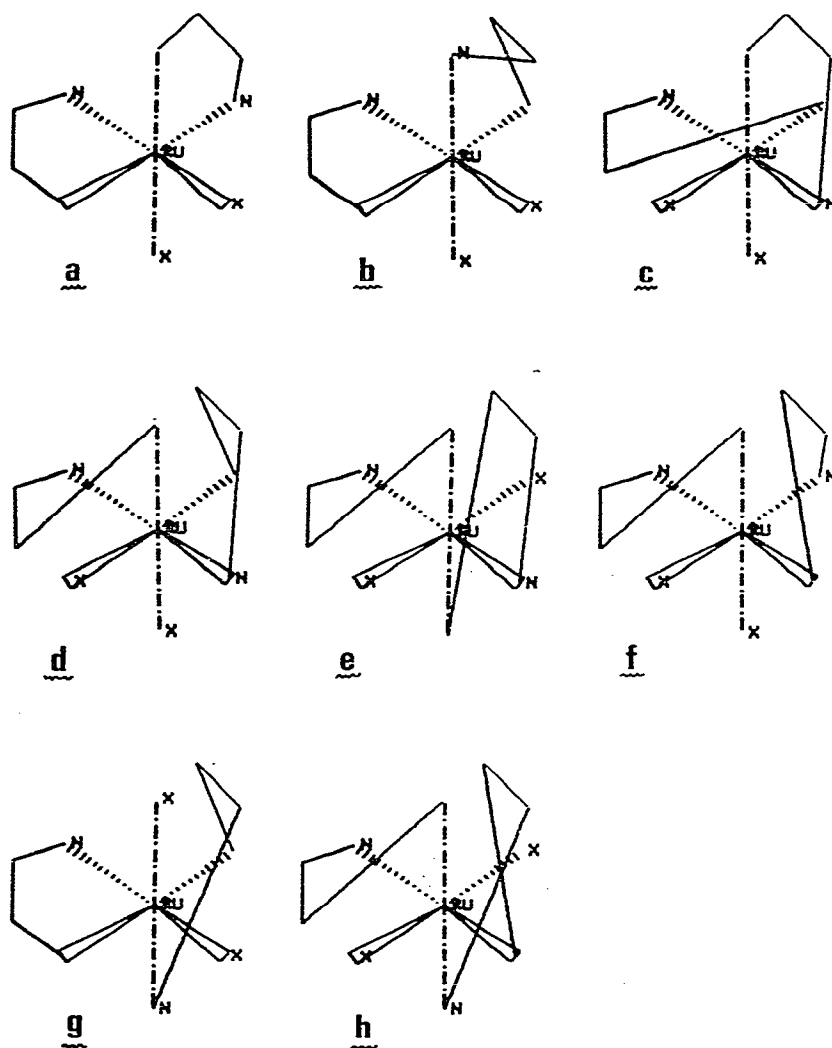


Fig. 1. Stable stereoisomers of a ruthenium complex, displayed by the GT40 ($X = CO$). The program recognizes the following pairs of enantiomers: (a,f), (b,g), (c,d).

case the drawing is too intricate, he can move the atoms other than the attachments to the central atom with the light pen, and obtain a diagram such as VIII'.

Results

The results obtained with our program are in agreement with those given by Bennett for a set of octahedral complexes [2]. We present the output for two molecules: one octahedral complex of Ru [9] (Fig. 1), and a pentacoordinated phosphorane with two carbon stereocenters [10] (Fig. 2). Total computer time

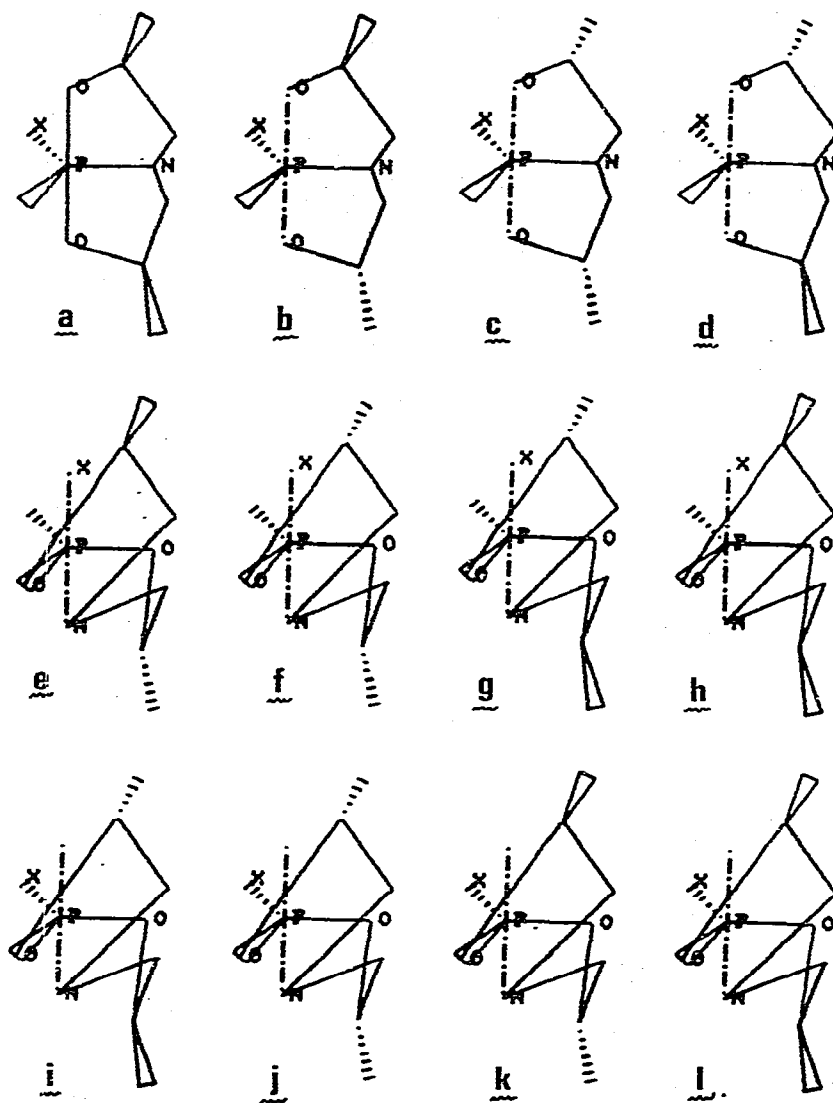


Fig. 2. Stereoisomers of a phosphorane displayed by the GT40 ($X = H$). The program recognizes the following pairs of enantiomers: (o,d), (e,g), (i,k). Hydrogens are implicit, except when attached to a pentacoordinated phosphorus [7].

to generate and display these isomers is respectively 16 and 15.8 sec (CPU time is 7 and 6.7 sec.), but it can increase drastically for highly symmetrical structures. This time could be reduced, however by use of fairly complex symmetry short-cuts which are not implemented at the present time. For all the examples which have been tested, the program is very convenient to use, and rapidly provides easily interpretable diagrams without the intervention of the chemist.

Acknowledgement

The authors (F.C., R.D., G.K.) thank the CNRS for a grant of computer time. One of us (T.W.) wishes to acknowledge support from NIH Research Resources grant RR01059.

References

- 1 (a) W.C. Fernelius and B.E. Bryant, *J. Amer. Chem. Soc.*, 75 (1953) 1735. (b) R.F. Trimble, Jr., *J. Chem. Educ.*, 31 (1954) 176. (c) B.A. Kennedy, D.A. McQuarrie and C.H. Brubaker, Jr., *Inorg. Chem.*, 3 (1964) 265. (d) J.C. Baylar, Jr., *J. Chem. Educ.*, 34 (1957) 334, 626.
- 2 W.E. Bennett, *Inorg. Chem.*, 8 (1969) 1325.
- 3 E.J. Corey, W.T. Wipke, R.D. Cramer III and W.J. Howe, *J. Amer. Chem. Soc.*, 94 (1972) 421.
- 4 *Simulation and Evaluation of Chemical Synthesis program*, (a) T.M. Dyott, Ph.D. Thesis, Princeton University, 1973. (b) W.T. Wipke, in *Computer Representation and Manipulation of Chemical Information*, Wiley Interscience, N.Y. 1974, p. 147.
- 5 W.T. Wipke and T.M. Dyott, *J. Amer. Chem. Soc.*, 96 (1974) 4825.
- 6 W.T. Wipke and T.M. Dyott, *J. Chem. Info. and Comput. Sc.*, 15 (1975) 140.
- 7 F. Choplin, R. Marc, G. Kaufmann and W.T. Wipke, to be published.
- 8 W.T. Wipke and T.M. Dyott, *J. Amer. Chem. Soc.*, 96 (1974) 4834.
- 9 R.L. Bennet, M.I. Bruce, B.L. Goodall, M.Z. Iqbal and F.G.A. Stone, *J. Chem. Soc. Dalton*, (1972) 1787.
- 10 D. Houalla, T. Mouheich, M. Sanchez and R. Wolf, *Phosphorus*, 5 (1975) 229.