match no. 4 pp. 121 - 130 1978

LA MOLÉCULE TÉTRAÉDRIQUE CHIRALE: UN FORMALISME ORIENTÉ VERS L'ORDINATEUR ET LE PROGRAMME "STEREO-1"

O. Dragomir-Filimonescu et I. Moţoc *

(Revised version received:August 1978)

Le formalisme algébrique dévelloppé dans cet article petmet le traitement de la notion de configuration de la molécule tétraédrique chirale, ce problème devenant implémentable sur l'ordinateur. On présente la description du programme STEREO-1 qui reconnait le nom de la configuration d'une molécule tétraédrique chirale (au cadre du système R/S).

1. Introduction

On utilise dans cet ouvrage le système (supposé connu) de Cahn, Ingold et Prelog. On fait ensuite une succinte description du programme, on donne le listing du programme principal, deux ordinogrammes et une application.

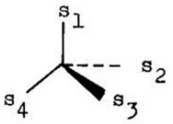
STEREO-1 est le premier d'une série développée pour le traitement, à l'aide de l'ordinateur, des problèmes de stéréochimie. Il reconnait seulement la configuration. Il peut opérer seulement sur les ligands atomiques ou sur les chaines organiques linéaires. Une version améliorée (STEREO-2) peut distinguer aussi les énantiomères (le listing complet paraîtra dans la Revue Roumaine de Chimie). Une autre version, plus complète (STEREO-3), se trouve dans un état avancé de travail. Ce programme pourra opérer aussi sur les chaines organiques ramifiées.

2. Théorie

Soit la molécule tétraédrique (1), où s_i, i=1,..,4

* Adresser correspondance à cet auteur à l'adresse indiquée dans l'article précédent.

sont des ligands supposés différents et qui appartiennent[1] à l'espace des ligands S.



(1)

Pour déterminer la configuration de la molécule (1) à l'aide de l'ordinateur, il est nécessaire de faire une description convenable des éléments suivants:

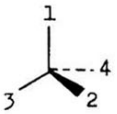
i) la position spatiale des ligands s_i (pour faire possible l'utilisation des règles de chiralité);

ii) la structure des ligands s_i (pour faire possible l'utilisation des règles de succession).

Ainsi:

2.1. La description de la position spatiale relative des ligands autour du tétraèdre

Soit l'indexation (2) des sommets du tétraèdre:



(2)

que nous allons considérer standard.

Nous désirons obtenir une description ayant la forme:
- à la position "1" correspond le ligand s_4;
- à la position "2" correspond le ligand s_1;
- etc...

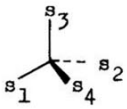
Ayant donnée l'indexation des ligands, il est facile de voir que l'utilisation des substitutions[2] convient à ce but.

Donc, les substitutions $h:\{1,2,3,4\} \rightarrow \{1,2,3,4\}$ seront les objets mathématiques qui décrivent la position relative des ligands.

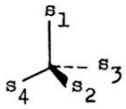
Soit par exemple:

$$h_1=\begin{pmatrix}1 & 2 & 3 & 4\\ 3 & 4 & 1 & 2\end{pmatrix} \qquad \text{et} \qquad h_2=\begin{pmatrix}1 & 2 & 3 & 4\\ 1 & 2 & 4 & 3\end{pmatrix}$$

qui décrivent les molécules:



(h_1) et

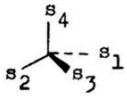


(h_2)

ainsi que

$$h_3=\begin{pmatrix}1 & 2 & 3 & 4\\ 4 & 3 & 2 & 1\end{pmatrix}$$

décrit la molécule:



On peut observer que les permutations dans "h", correspondantes à un nombre pair de transpositions, décrivent des rotations du tétraèdre, tandis que les permutations correspondantes à un nombre impair de transpositions décrivent des positions chirales.

Les rotations ne modifient pas la configuration; elles seront donc utilisées pour sa détermination.

2.2. La description de la structure des ligands

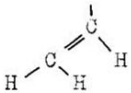
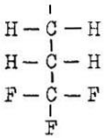
Pour décrire la constitution et la topologie des ligands s_i, on va utiliser les doublets de matrices (X,M) où: $X=[x_{ij}]$, $j=1,2,3$; $i=1,..,k\in N$;
$M=[m_{ij}]$, ibid.

X est la matrice des ordres de liaison, x_{ij} représentant l'ordre de la liaison i-j arrondi au plus proche numéro entier.

M est la matrice des numéros d'ordre et/ou de massedes atomes correspondants aux positions (ij) de X.

STEREO-1 opérant seulement sur les ligands-atomes ou sur les chaînes organiques linéaires, les matrices X et M sont limitées à ces cas. Les deux matrices sont construites d'une manière similaire: - la première ligne correspond à l'atome attaché au centre du tétraèdre; la deuxième ligne correspond aux atomes attachés à l'atome précédent, etc...,.

On va fixer ces idées avec les exemples suivants:

ligand	matrice X	matrice M
	$\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 9 & 0 & 0 \end{bmatrix}$
	$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$	$\begin{bmatrix} 6 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$
	$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 6 & 0 & 0 \\ 8 & 1 & 0 \end{bmatrix}$
	$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 6 & 0 & 0 \\ 6 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$
	$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$	$\begin{bmatrix} 6 & 0 & 0 \\ 1 & 6 & 1 \\ 1 & 6 & 1 \\ 9 & 9 & 9 \end{bmatrix}$

Observation: les matrices

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{et} \quad \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

sont équivalentes respectivement aux matrices

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad \text{et} \quad \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \end{bmatrix}$$

Cette observation sera utilisée dans le programme STEREO-1 pour comparer les matrices.

2.3. Les règles de succession

STEREO-1 utilise la règle de succession suivante: "le numéro d'ordre plus grand précède le numéro d'ordre plus petit".

Par l'utilisation du doublet (X,M), cette règle devient:

Soit s_p un ligand avec son doublet:

$$X^p=[x^p_{ij}]\ ;\quad M^p=[m^p_{ij}]$$

et s^q un ligand avec son doublet:

$$X^q=[x^q_{ij}]\ ;\quad M^q=[m^q_{ij}]$$

On dit:

$$s_p \succ s_q$$

si et seulement si:

$$\mathrm{MAX}\left\{\sum_{j=1}^{3} c^p_{i_o j}, \sum_{j=1}^{3} c^q_{i_o j}\right\} = \sum_{j=1}^{3} c^p_{i_o j} \quad \text{unique}$$

et aussi

$$\sum_{j=1}^{3} c^p_{ij} = \sum_{j=1}^{3} c^q_{ij} \qquad \text{pour } \forall i=1,\ldots,i_o-1$$

On a noté par c_{ij} l'élément de la matrice:

$$C=[c_{ij}]=[x_{ij}\cdot m_{ij}] = X\cdot M \quad \text{(produit direct)}$$

où: j=1,2,3; i=1,...,k N.

2.4. Les règles de chiralité

Etant donnée la substitution:

$$h : \{1,2,3,4\} \rightarrow \{1,2,3,4\}$$

notée par:

$$h = \begin{pmatrix} i \\ h(i)=j \end{pmatrix}, \quad i=1,2,3,4$$

on remplace h(i) avec le rang du ligand "j" dans une séquence spécifiée de quatre ligands. Donc, on obtient la

substitution:

$$h' = \begin{pmatrix} i \\ h(i)=r_{s_j} \end{pmatrix} \quad , \quad i=1,2,3,4$$

où r_{s_j} symbolise le rang du ligand s_j.

On fait des permutations circulaires dans h' jusqu'à ce qu'on obtienne une substitution dans laquelle

$$h(4)=4$$

On néglige cette collonne de la substitution h' et le résultat est une substitution "g" ayant la forme:

$$g = \begin{pmatrix} i \\ g(i)=r_{s_j} \end{pmatrix} \quad , \quad i=1,2,3$$

g peut avoir l'une des formes suivantes:

$$g_o=\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \; g_1=\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \; g_2=\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$$

ou les transposées des ceux ci-dessus.

Mais: $(\{g_o, g_1, g_2\}=G,*)$ où "*" est la loi de composition des substitutions, est un groupe cyclique d'ordre 3; aussi, l'indexation (2) précise le concept de configuration R (rectus).

On peut facilement prouver les deux affirmations suivantes:

i) Etant donnée "g", g décrit une configuration R si et seulement si:

$$\exists\, g_i \in G \quad : \quad g*g_i=g_o \qquad (3)$$

ii) Si (3) n'est pas satisfaite, g décrit une configuration S.

3. Le programme STEREO-1

3.1. L'algorythme

a) on lit "h" et les quatre doublets (X,M);
b) on détermine h';
c) on détermine g;
d) à partir de (3) on détermine la configuration.

3.2. Présentation du programme

- nom	STEREO-1
- ordinateur	FELIX C-256
- système d'opération	SIRIS 3 V15A
- language de programmation	FORTRAN IV
- ressources nécessaires	CR, PR, 32K mém.

Les ordinogrammes des séquences les plus importantes du programme sont présentées dans les figures 1 et 2.

La figure 1 représente le schéma logique du programme principal. Le sousprogramme PRODMATR execute le produit direct entre les matrices X et M, pour obtenir C. Le sousprogramme COMPAR évalue le nombre des matrices identiques (NRMIDENT); si on a une molécule chirale on appelera le sousprogramme CONFIG qui déterminera la configuration de la molécule à partir des quatre matrices C (figure 2).

A la fin on présente en exemple la séquence principale de STEREO-1.

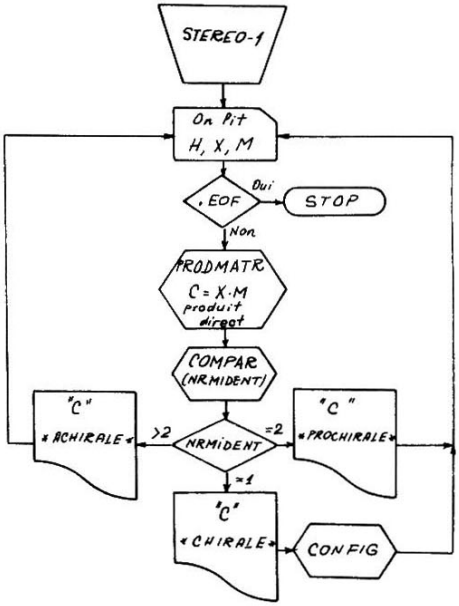


Fig. 1

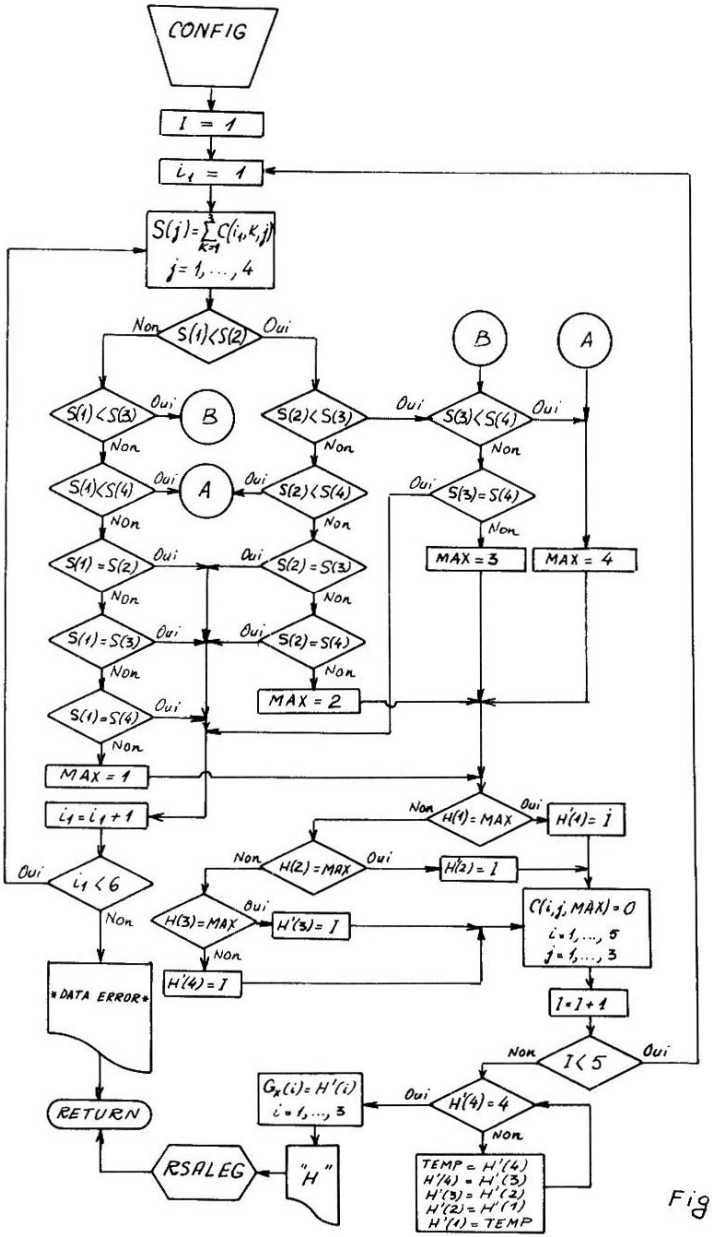

CONFIG
I = 1
i1 = 1
S(j) = Σ C(i1,k,j)
k=1
j = 1, ..., 4
S(1) < S(2)
Non
Oui
S(1) < S(3)
S(1) < S(4)
S(1) = S(2)
S(1) = S(3)
S(1) = S(4)
S(2) < S(3)
S(2) < S(4)
S(2) = S(3)
S(2) = S(4)
S(3) < S(4)
S(3) = S(4)
A
B
MAX = 1
MAX = 2
MAX = 3
MAX = 4
i1 = i1 + 1
i1 < 6
* DATA ERROR *
RETURN
H(1) = MAX
H(2) = MAX
H(3) = MAX
H'(1) = j
H'(2) = I
H'(3) = I
H'(4) = I
C(i,j,MAX) = 0
i = 1, ..., 5
j = 1, ..., 3
I = I + 1
I < 5
H'(4) = 4
TEMP = H'(4)
H'(4) = H'(3)
H'(3) = H'(2)
H'(2) = H'(1)
H'(1) = TEMP
Gx(i) = H'(i)
i = 1, ..., 3
"H"
RSALEG

Fig. 2

```
      INTEGER G0(3),G1(3),G2(3),GX(3),PRGXG0(3),PRGXG1(3),
     *PRGXG2(3),TABELRS(12),C(5,3,4),CLUCRU(5,3,4),H(4),
     *HTEMP(4),SUMA(4),MCHIRALE(5,3,4,12),X(5,3,4),M(5,3,4)
      DATA G0(1),G0(2),G0(3),G1(1),G1(2),G1(3),G2(1),G2(2),
     *G2(3)/1,2,3,2,3,1,3,1,2/
      L=1
      IPAG=0
    1 READ(105,115,END=1000)H
  115 FORMAT(4I1)
      IPAG=IPAG+1
      IF(IPAG.EQ.7)GOTO 333
    2 READ(105,116,END=1000)(((X(I,J,K),J=1,3),I=1,5),K=1,4),
     *(((M(I,J,K),J=1,3),I=1,5),K=1,4)
  116 FORMAT(15I4)
      CALL PRODMATR(X,M,C)
      CALL COMPAR(NRMIDENT,C)
      IF(NRMIDENT.EQ.1)GOTO200
      IF(NRMIDENT.EQ.2)GOTO400
      WRITE(108,20)(((C(I,J,K),J=1,3),K=1,4),I=1,5)
   20 FORMAT(' '////4X,3I3,4X,3I3,4X,3I3,4X,3I3,
     *'     ** MOLECULE ACHIRALE **'
     */(4X,3I3,4X,3I3,4X,3I3,4X,3I3))
      GOTO1
  333 WRITE(108,777)
  777 FORMAT('1'//)
      IPAG=IPAG-6
      GOTO2
  400 WRITE(108,30)(((C(I,J,K),J=1,3),K=1,4),I=1,5)
   30 FORMAT(' '////4X,3I3,4X,3I3,4X,3I3,4X,3I3,
     *'     ** MOLECULE PROCHIRALE **',
     */(4X,3I3,4X,3I3,4X,3I3,4X,3I3))
      GOTO1
  200 DO 10 K=1,4
      DO 10 I=1,5
      DO 10 J=1,3
   10 MCHIRALE(I,J,K,L)=C(I,J,K)
      WRITE(108,35)(((C(I,J,K),J=1,3),K=1,4),I=1,5)
   35 FORMAT(' '////4X,3I3,4X,3I3,4X,3I3,4X,3I3,
     *'     ** MOLECULE CHIRALE **'
     */(4X,3I3,4X,3I3,4X,3I3,4X,3I3))
      DO 40 K=1,4
      DO 40 J=1,3
      DO 40 I=1,5
   40 CLUCRU(I,J,K)=C(I,J,K)
      CALL CONFIG(CLUCRU,H,TABELRS,L,GX,G0,G1,G2)
      L=L+1
      GOTO1
 1000 STOP
      END
```

LE LISTING DU PROGRAMME PRINCIPAL

....

Les auteurs présentent leurs remerciements les plus respectueux au Proffesseur O. E. Polansky pour ses très utils conseils concernant l'ouvrage.

Bibliographie

1. I. Moţoc, Rev. Roumaine Chim., 20, 1341(1975)
2. I. Moţoc, Rev. Roumaine Chim., 20, 1335(1975)

Adresses des auteurs:

O. Dragomir-Filimonescu — Institut de Recherches pour la Technique de Calcul, 9 rue Gh. Lazăr, 1900 Timişoara, ROUMANIE

I. Moţoc — Département de Chimie Physique, Université, 1900 Timişoara, ROUMANIE