match no. 4 pp. 113 - 120 1978

A NEW STERIC PARAMETER, OV, AND MINIMAL STERIC DIFFERENCE, MSD

Ioan Moţoc
Chemical Research Centre
M. Viteazu 24, 19oo Timişoara
Romania

(Revised version received: August 1978)

Content : In the frame of the set theory, a new steric parameter, OV (overlapping volumes), is defined. General formalism and two theorems concerning OV are reported here. Finally, the paper reports a new definition of MSD (minimal steric difference) and some topological considerations concerning the initial formalism of MSD parameter.

1. Introduction

Several kinds of steric parameters have been proposed and applied for structure - chemical activity relationships. The steric parameters defined by Taft,[1] Hancock,[2] Palm,[3] Charton,[4] Becker[5] and other are widely used.

In molecular biology the steric parameters of Amoore,[6] Allinger,[7] Rao,[8] Simon[9-12] among other, are used. (For review with applications, see Chapman and Shorter[13]).

This paper deals with the definition, the computing method and two theorems concerning a new steric parameter, OV.

A set theory consideration of minimal steric difference (MSD) leads to a relation between OV and MSD. Finally a topological consideration of MSD shows that MSD induces a Hansdorf space structure on the set M of molecules considered.

2. General formalism

The formalism of the set theory allows to define a new steric parameter. Let us consider :

Definition 1 : a given set is named a W_2 - set if :

$$W_2 = \{ w/P,\ P' \} \qquad (1)$$

where : $w = (w_x, w_y, w_z)$

P is the property which implies : "w are random points with uniform repartition" and P' implies : w fulfills the equation

$$\begin{cases} (w_x-A_{IJ})^2 + (w_y-B_{IJ})^2 + (w_z-C_{IJ})^2 < R_{IJ}^2 & (2) \\ \qquad I = 1,\dots, N_J \\ (w_x-L_{PQ})^2 + (w_y-M_{PQ})^2 + (w_z-N_{PQ})^2 < R_{PQ}^2 & (3) \\ \qquad P = 1,\dots, N_Q \end{cases}$$

A_{IJ}, B_{IJ}, C_{IJ} are cartesian co-ordinates of atom i from molecule J ; R_{IJ} is the Van der Waals radius of atom I in molecule J ; N_J is the number of atoms of the molecule J.

L_{PQ}, M_{PQ}, N_{PQ} are cartesian co-ordinates of atom P from molecule Q ; R_{PQ} is the Van der Waals radius of atom P in molecule Q ; N_Q is the number of atoms of the molecule Q.

<u>Definition 2</u> : a given set is named a W_1 - set if :

$$W_1 = \{ w/P,\ P_1,\ P_2 \} \qquad (4)$$

where : P_1 is the property which implies : "w is within a parallelepiped Y, defined by P_2".

P_2 is as : $W_2 \subset W_1$

<u>Definition 3</u> : a given set is named a W^j - set, j = 1, 2 if :

$$W^1 = \left\{ m_1/P, (m_{1x}-A_{IJ})^2 + (m_{1y}-B_{IJ})^2 + (m_{1z}-C_{IJ})^2 < R_{IJ}^2 \right\} \qquad (5)$$

$$I = 1,\dots, N_J$$

respectively

$$W^2 = \left\{ m_2/P, (m_{2x}-L_{PQ})^2 + (m_{2y}-M_{PQ})^2 + (m_{2z}-N_{PQ})^2 < R_{PQ}^2 \right\} \qquad (6)$$

where $m_j = (m_{jx}, m_{jy}, m_{jz})$ and P is the property : m_j are random points with uniform repartition.

Since the points m_j are obtained with a random number generator their densities in the molecules J and Q respecti-

vely should tend to become equal. Therefore W^1 and W^2 are relative measures of the Van der Waals volumina of the corresponding molecules if the W^j - sets are large enough.

3. The steric parameter OV

Definition 4 : The steric parameter OV is given by relation (7) :

$$OV = \alpha \, |W_2| \, / \, |W_1| \qquad (7)$$

where $|W_1|$ and $|W_2|$ denote the cardinal numbers of the sets W_1 and W_2 respectively ; α is the volume of the parallelepiped Y.

The following conclusions may be easily seen :

(i) the dimension of OV is L^3;

(ii) the Monte Carlo technique is the computing method of OV;

(iii) OV depends as well on the shape and structure of the molecules J and Q as on the distance between them .

In a future note, we will report LFER computations , using this steric parameter.

Assuming that elements of $W_{i=1,2}$ and $W^{j=1,2}$ are obtained with a random number generator with the period X such :

$$X = \max_{i,j} \left\{ |W_i| \, , \, |W^j| \right\} \qquad (8)$$

we have the following theorems :

Theorem 1 : $W^1 \cap W^2$ is a measure of OV for a given set W_1.

By definition, we have :

$$OV = \alpha \, |W_2| \, / \, |W_1| = |W^1 \cap W^2| \, / \, |W_1| = \beta \, |W^1 \cap W^2|$$

where $\beta = \alpha \, / |W_1|$ is a constant, because both $|W_1|$ and α are constants if W_1 is given. Q. E. D.

Finally, let us observe that OV is symmetrical function of the structure and shape of the molecule J and Q because :

$$W^1 \cap W^2 = W^2 \cap W^1$$

4. MINIMAL STERIC DIFFERENCE

In the initial formulation,[14,15] the minimal steric difference (MSD) is an approximative method for numerical evaluation of the steric fit of (or difference between) a molecule with respect to a standard. The method consists in superimpo-

sing the two molecules to be compared. MSD is the sum of the nonoverlapping Van der Waals volumes of the two molecules, but in the initial version, the (approximately) nonsuperposable atoms are counted and taken as MSD.

The improved version[16,17] of MSD consists in the exact evaluation of nonoverlapping volumes of molecules by means the Monte Carlo technique.

We used MSD as steric parameter in multilinear structure - biologic activity correlations, with good results.[18-20]

5. MSD : Sets theory considerations

Let us denote the complementary set of W^1 in W^2 by $\overline{W}^1$ and of W^2 in W^1 by $\overline{W}^2$, $\overline{W}^1 = \{ w \ / \ w \in W^2, \ w \notin W^1 \}$, $\overline{W}^2 = \{ w \ / \ w \in W^1, \ w \notin W^2 \}$. For a given set W_1 (this means that W^1 and W^2 are specified, because W_1, W^1, W^2 are obtained with the same random generator with the period $X \geqslant |W_1|$) we have the following :

Proposition 1 : $\overline{W}^1 \cup \overline{W}^2$ is a measure of MSD.

Using the MSD definition[14] and the Monte Carlo method[21] we obtain following expression : $MSD = \alpha |\overline{W}^1 \cup \overline{W}^2| \ / \ |W_1|$ where α is the volume of parallelepiped Y.

Now, because W_1 is given (say Card W_1 = 100000), α is a constant and it results that $\overline{W}^1 \cup \overline{W}^2$ is a measure of MSD.

$$MSD + 2OV = \alpha \left(|W^1| / |W_1| + |W^2| / |W_1| \right)$$

6. Relation between OV and MSD

For a given distance between the two molecules and for $|W^1|$, $|W^2|$ large enugh, the connection between MSD and OV is given by

Theorem 2 : $\overline{W}^1 \cup (W^1 \cap W^2)$ is a measure of the Van der Waals volume of the molecule Q = 2.

$\overline{W}^2 \cup (W^1 \cap W^2)$ is a measure of the Van der Waals volume of the molecule J = 1.

Indeed,

$\overline{W}^1 \cup (W^1 \cap W^2) = (\overline{W}^1 \cup W^1) \cap (\overline{W}^1 \cup W^2) = (W^1 \cup W^2) \cap W^2 = W^2$ because $\overline{W}^1 \subset W^2 \rightarrow \overline{W}^1 \cup W^2 = W^2$ and $\overline{W}^1 \cup W^1 = W^1 \cup W^2$;

$\overline{W}^2 \cup (W^1 \cap W^2) = (\overline{W}^2 \cup W^1) \cap (\overline{W}^2 \cup W^2) = W^1 \cap (W^1 \cup W^2) = W^1$ because

$\overline{W}^2 \subset W^1 \Rightarrow \overline{W}^2 \cup W^1 = W^1$ and $\overline{W}^2 \cup W^2 = W^1 \cup W^2$.

7. MSD : topological considerations

Let be the set :[22]

$$M = \{ m_i / m_i = \text{molecules} \}_{i \in N}$$

N is the set of natural numbers.

We call M the fundamental space, and m_i the space points.

Assumptions considered in the initial, atom-per-atom, formulation[14] of the MSD allow the following :

Definition 3 : The points of a cubic lattice L in Euclidian E_3 space are the topological co-ordinates of every $m_i \in M$. (Each atom of m_i corresponds to a point of L).

Lemma 1 : For every $m_i \in M$ there is a sublattice $l_i \in L$ which describes the m_i structure.

The lemma can be proved easily, leaving it for the reader.

Let us denote the set of all sublattices obeying Lemma 1 by P(L), then :

Lemma 2 : M and P(L) are isomorphic : $M \simeq P(L)$. Indeed, M and P(L) are numerable.

Corollary : It is equivalent to work with $m_i \in M$ or with $l_i \in L$.

The corollary is proved by Lemma 2, and it will be largely used in the following.

The entries of $l_i \in L$ are W^i_{rpq} - approximative Van der Waals volumes of the atoms located on the points $(rpq) \in l_i$.

Definition 4 : The composition law

$$\ominus : L \times L \rightarrow R_+$$

is given by relation :

$$l_i \ominus l_j = \sum_{\text{all } r,p,q} \left| W^i_{rpq} - W^j_{rpq} \right| \text{ where } l_i, l_j \in P(L)$$

$$r = 1,\ldots,R \; ; \; p = 1, \ldots , P \; ; \; q = 1,\ldots, Q \; ;$$

$$R, P, Q \in N$$

From (14) is obvious that :

$$l_i \ominus l_j = l_j \ominus l_i$$

and expressions such as :

$$(l_i \ominus l_j) \ominus l_k \; ; \; l_i \ominus (l_j \ominus l_k) \; ; \; \ldots$$

are meaningless.

Thus :

$$MSD(m_i, m_j) = MSD(m_j, m_i) = l_i \ominus l_j = l_j \ominus l_i$$

$$m_i, m_j \in M \; ; \; l_i, l_j \in P(L)$$

<u>Proposition 2</u> : MSD induces a Hausdorff space structure on M.

Indeed, we can define the function ϱ :

$$\varrho : P(L) \times P(L) \longrightarrow R_+$$

as follows :

$$\varrho(l_i, l_j) = \sqrt{(l_i \ominus l_j)^2}$$

ϱ has the following properties :

$$\varrho(l_i, l_j) \geqslant 0, \; \varrho(l_j, l_i) = \varrho(l_i, l_j) \; , \; \forall \, l_i, l_j \in P(L)$$

$$\varrho(l_i, l_i) = 0 \qquad , \; \forall \, l_i \in P(L)$$

$$\varrho(l_i, l_k) \leqslant \varrho(l_i, l_j) + \varrho(l_j, l_k), \forall l_i, l_j, l_k \in P(L)$$

Thus, ϱ is a metrics, and $(P(L), \varrho)$ is a metric space. Acording to a well-known theorem[23] in topology $(P(L), \varrho)$ is a Hausdorff space. From $P(L) \simeq M$ we have that (M, ϱ) is a Hausdorff space also, and the theorem is proved.

<u>Disscusion</u>

Dugundji, Ugi et al.[24-26] using FIEM sets, have obtained the metrics :

$$D(B,E) = \sum_{i,j} |e_{ij} - b_{ij}|$$

where D(B,E) represents the constitutional distance between two molecules. Our metrics $\varrho(l_i, l_j)$, represents the structural distance between two molecules.

ACKNOWLEDGMENTS

I would like to thank Prof. Dr. O. E. Polansky for valuable comments concerning the present paper.

References

1. R. W. TAFT, in "Steric Effects in Organic Chemistry", ed. M. S. NEWMAN, Wiley, New York, 1956, p.556.
2. C. K. HANCOCK, E. A. MEYERS and B. J. YAGAR, J. Amer. Chem. Soc., 83, 4211 (1961).
3. V. A. PALM, "Fundamentals of the Quantitative Theory of Organic Reactions", Khimiya, Leningrad, 1967, c. 1o.
4. M. CHARTON, J. Amer. Chem. Soc., 97, 1552, 3691, 3694 (1975).
5. F. BECKER, Z. Naturforch., 14a, 547 (1959) ; ibid., 15b, 251 (196o) ; ibid., 16b, 236 (1961).
6. J. E. AMOORE, G. PALMIERI and E. WANKE, Nature, 216, 1o84 (1967).
7. N. L. ALLINGER, Proc. Int. Congress Pharmacol., 5, 5763 (1973).
8. S. T. RAO, M. G. ROSSMANN, J. Mol. Biol., 76, 241 (1973).
9. Z. SIMON, Z. SZABADAI, Studia Biophys. (Berlin), 39, 123 (1973).
1o. Z. SIMON, Angew. Chem. Internat. Edit., 13, 27 (1974).
11. I. MOTOC, ST. HOLBAN, D. CIUBOTARIU and Z. SIMON, Preprints, Timişoara Univ., 1/1975, 14/1975, 16/1975.
12. I. MOTOC, Match, this issue.
13. N. B. CHAPMAN and J. SHORTER, "Advances in Linear Free Energy Relationships", Plenum Press, London, 1972.
14. Z. SIMON and Z. SZABADAI, Studia Biophys. (Berlin), 39, 123 (1973).
15. Z. SIMON, Angew. Chem., 13, 719 (1974).
16. I. MOTOC, ST. HOLBAN, D. CIUBOTARIU and Z. SIMON, Preprint, Univ. Timişoara, Serie Chimie 17/1975.
17. I. MOTOC, ST. HOLBAN, R. VANCEA and Z. SIMON, Studia Biophys., (Berlin), 66, 75 (1977).
18. Z. SIMON, Studia Biophys., (Berlin) 51, 49 (1975).
19. Z. SIMON, I. MOTOC, ST. HOLBAN and D. CIUBOTARIU, Preprint, Univ. Timişoara, Serie Chimie 16/1975.
2o. I. MOTOC and A. VLAD, Studia Biophys. (Berlin), in press.

21. a) YU. A. SHREIDER, "The Monte Carlo Method", Pergamon Press, London, 1966.
b) H. KAHN, "Applications of Monte Carlo", Rand Memorandum RM-1237, 1956.
22. I. MOTOC and Z. SIMON, Preprint, Univ. Timişoara, Serie Chimie 15/1975.
23. N. BOURBAKI, "Éléments de Mathématique. Topologie Générale", Hermann, Paris 1964.
24. J. DUGUNDJI and I. UGI, Fortschr. chem Forsch., 39, 19 (1973).
25. I. UGI, P. GILLESPIE and C. GILLESPIE, Trans. N. Y. Acad. Sci., 34, 416 (1972).
26. J. DUGUNDJI, P. GILLESPIE, D. MARQUARDING, I. UGI and F. RAMIREZ, in "Chemical Applications of Graph Theory", ed. A. T. BALABAN, Academic Press, London, 1976, p. 1o7-175.