match no. 8 pp. 323 - 332 1980

SECTOR PARTITION OF THE MOLECULAR VAN DER WAALS SPACE AS A MEASURE OF STERIC EFFECTS

I.Moţoc [a], O.Dragomir-Filimonescu [a] and I.Muşcutariu [b]

a) Chemistry Research Centre,
Bulev.M.Viteazu 24, 1900 Timişoara 5

b) Department of Physics, The University,
Str.V.Pârvan 4, 1900 Timişoara 5
Romania

(received: March 1980)

Summary. The paper presents a Monte Carlo method developed to perform the sector partition (i.e. octants, quadrants etc.) of the molecular Van der Waals space. The resulted sector volumes are useful as steric parameters in QSAR. The complete listing of the program which implements the described method (NØVA program) and an illustrative application are also included.

1. Introduction

The quantitative treatment of steric effects is one of the most difficult problems in the framework of quantitative structure - activity relationships (QSAR) : "The problem of steric interactions represents the ultimate in difficulty of the drug designer" (C. Hansch [1]).

We propose here to use as steric parameters in QSAR the sector volumes (i.e. octants, quadrants etc.) of the drug molecules. The general strategy is to :

1) Consider a series of N bioactive compounds possessing the same type of biological activity, elicited via the same mechanism.

2) Compute the octant volumes of the molecules 1, 2,...,N. The starting cartesian coordinates of the intersection point P between three reference planes $P(x_o, y_o, z_o)$ are specified (fig. 1). We denote the resulted octant volumes by V_1, V_2, ..., V_8. Obviously, a convenient summation of the V_i, i = 1 - 8, values results to the quadrant etc. volumes (see Figure 2).

3) Compute the equation (1) :

$$BR_i = a + \sum_{\substack{\text{sector} \\ \text{volumes}}} b_j V_{ji} \quad , \qquad i = 1, 2, \ldots, N \qquad (1)$$

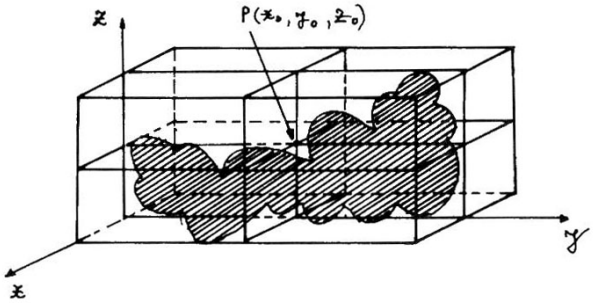


Figure 1. The octant partition of the Van der Waals molecular space.

The definition of octants

V_j	x	y	z	Binary code		
V_1	x_o	y_o	z_o	0	0	0
V_2	x_o	y_o	z_o	0	0	1
V_3	x_o	y_o	z_o	0	1	0
V_4	x_o	y_o	z_o	0	1	1
V_5	x_o	y_o	z_o	1	0	0
V_6	x_o	y_o	z_o	1	0	1
V_7	x_o	y_o	z_o	1	1	0
V_8	x_o	y_o	z_o	1	1	1

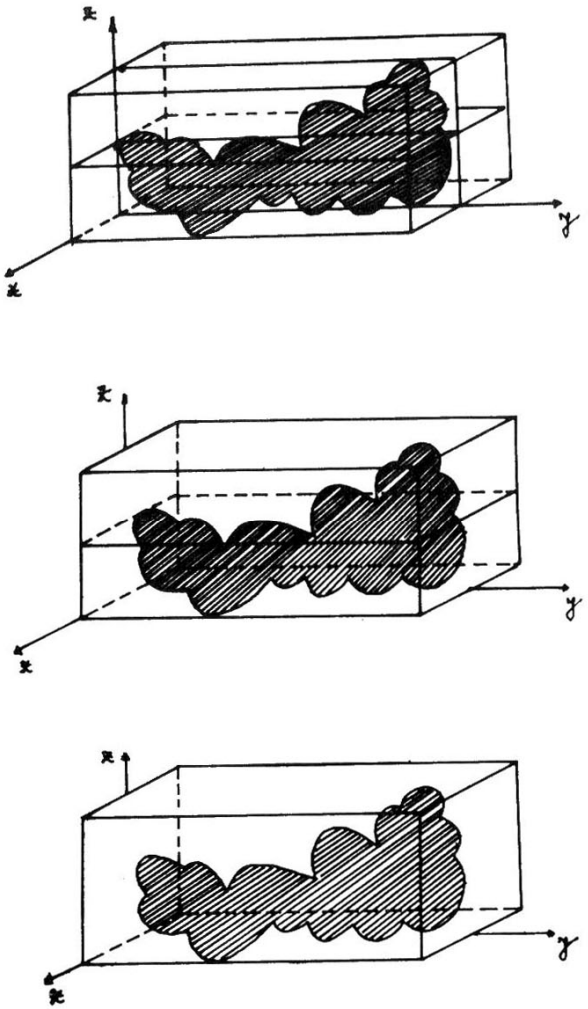


Figure 2 :Other sector partition derived from the octant partition.

where BR_i stands for the biological response elicited by the bioactive molecule i, $1 \leq i \leq N$; a and b_j are determined by means of the least square method. Equation (1) is characterized by the correlation coefficient (r), Fisher statistics (F) and Student t-statistics relative to the b_j regressional coefficients.

4) Change with a prescribed increment, $n.\Delta x$, $n.\Delta y$ and $n.\Delta z$ ($n = 1, 2, \ldots$ and $\Delta x = \Delta y = \Delta z = 1$ Å, or 0.5 Å etc.) the coordinates of $P(x_o, y_o, z_o)$, namely $P(x_o + n.\Delta x, y_o + n.\Delta y, z_o + n.\Delta z)$. The coordinates vary in the following succession : $x > y > z$. Recompute the V_i values and the equation (1).

5) From among the equations (1), computed for the various sector partitions of the Van der Waals molecular space, select the best one, according to r or F values. The statistically significant b_j regressional coefficients with the V_j in the best obtained equation show the positions where the steric effects are important ($b_j < 0$ denotes repulsive steric potential, and $b_j > 0$ an attractive steric potential). The statistically irrelevant b_j coefficients (according to Student statistics) indicate the positions where the steric effects are not important.

Finally, in order to obtain a QSAR for the considered series of bioactive compounds, one completes the selected equation in terms of V_j to account for the other effector/receptor interactions, namely electronic, hydrophobic etc.

$$BR_i = a + \sum_{\substack{\text{sector} \\ \text{volumes}}} b_j V_{ji} + \sum_k c_k D_{ki} \qquad (2)$$

where D_k are substituent constants which code the electronic ($D : \sigma, \sigma^+, \sigma_I, \sigma_R$, F, R etc.), hydrophobic ($D : \pi$, f) etc. interactions.

The method here introduced may be regarded as a generalization of the method recently proposed [2] by Testa and Purcell, and it is related to OVA method introduced by one of us [3,4].

2. The Computation of the Octant Partition of the Molecular Van der Waals space

In the framework of the Monte Carlo techniques [5], it is an easy task to perform the octant partition computations. One proceeds as follows :

1) Each molecule is described by the Cartesian coordinates and Van der Waals radii of the atoms.

2) One determines the parallelepiped of volume a, which circumscribes the considered molecule. Within this parallelepiped, one defines the octants i = 1, 2, ..., 8.

3) One generates N_t uniform random points P(x, y, z) within the octant i. From among the N_t random points, one selects the N_i points which satisfy the system of inequations (3) :

$$(x_{IJ} - x)^2 + (y_{IJ} - y)^2 + (z_{IJ} - z)^2 < R_{IJ}^2 \qquad (3)$$

where (x_{IJ}, y_{IJ}, z_{IJ}) are the Cartesian coordinates of the atoms (index I) in molecule J, and R_{IJ} stands for the Van der Waals radius of the atom I in molecule J.

The V_i values are computed according to relation (4) :

$$V_i = a\, N_i / N_t \qquad (4)$$

The V_i, i = 1, 2, ..., 8 values allow the computation of other partitions (they are systematised in Figure 2).

3. The NOVA Program

The NOVA program implements the above described method to compute the octant partition of the Van der Waals molecular space. The program is written in FØRTRAN IV and it was run on the FELIX C-256 computer.

The necessary input data are systematized below.

The input data are given by cards, and they contain :

a) 1 to 64 : Molecule Description Sets (MDS)

b) error specification (EPS)

c) an unlimited number of Run Control Data Sets (RCDS).

The end of MDS's and RCDS's will be marked by EØF cards.

(a)

MDS contains :

1) A control card, FØRMAT (5A8, I1, F8, 4) :

Column	Content
1 - 32	name of the molecule
33 - 40	the author of the Van der Waals radii to be used (BOYD or PAULING)
41	INPUT - an indicator variable
42 - 49	Y - biological activity of the considered molecule.

If INPUT = 1, the molecule is described by the coordinates of its atoms and by Van der Waals radii (one card for each atom). FØRMAT (4 F 8.4). In this case the current MDS will be ended with an EØF card.

If INPUT = 0, the atomic coordinates are computed by the program, using the following data.

i) one card containing : NØAT,(IZAT (I), I = 1,3), KWIK, RL2, R23, THETA, FØRMAT (5I3, 3F14.7).

ii) NØAT - 3 cards which contain : NA, NB, NC, ND, IZAT(ND), ILAZY, RCD, TH BCT, PA BCD, FØRMAT (6I4, 3F14.7).

The meaning of the above variables is given in ref.7 (technical documentation of CØØRD program).

(b)

One card which specifies the relative value of the expected error EPS, FØRMAT(F12.6).

(c)

The RCDS consists of a set of three cards :

i) card one, FØRMAT (80I1) :

Column	Content
1 - 3	IFCR(3), indicating the directions which are fixed in an iterative regression.
4	IREGRESS, a regression index
5 - 80	IMCØ(64), a vector which indicates the molecules considered in the current run.

ii) card two, FØRMAT(3 F15.7) :

Column	Content
1 - 45	XC(3), starting point coordinates

iii) card three, FØRMAT (F15.7) :

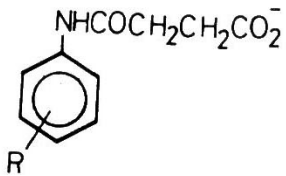
Column	Content
1 - 15	PAREG, the value of increment (x, y, z).

The listing of the NØVA program is given in the Appendix.

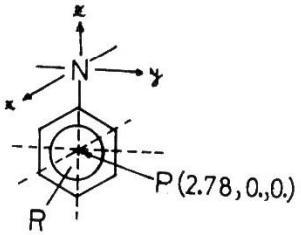
4. Applications

As an illustration, let us consider the eighteen phenylsuccinates I collected in Table 1 (data from ref. 6).

The octant partition computations were performed as shown in formula II.



I



II

The resulted values are systematized in Table 1.

It is interesting to note that the sector volumes may be used to express the steric features as "series expansion" :

$$-\Delta F^{o}_{rel} = -4.318 + 0.028\ WV \quad (r = 0.366,\ s = 0.614,\ F = 1.162)$$

$$-\Delta F^{o}_{rel} = -2.914 - 0.135\ V_{a_1} + 0.028\ V_{a_2} \quad (r = 0.543,\ s = 0.554,\ F = 1.947)$$

$$-\Delta F^{o}_{rel} = -3.439 + 0.167\ WV - 0.326\ V_{a_1} - 0.129\ V_{a_2}$$

$$(r = 0.781,\ s = 0.412,\ F = 5.071)$$

ΔF^{o}_{rel} stands for the hapten - antibody relative free energy of combination, r is the correlation coefficient, s is the standard deviation and F signifies the Fisher statistics.

Acknowledgement

We are grateful to Professor Dr.A.T.BALABAN, for valuable comments concerning the present paper.

References

1. C.HANSCH, in "Drug Design", vol.1 (E.J.ARIENS ed.), Academic Press, New York. 1970, p.271.
2. B.TESTA and W.P.PURCELL, Eur.J.Med.Chem., 13, 509 (1978).
3. I.MOTOC, Math.Chem., 4, 113 (1978).
4. I.MOTOC, I.COTAESCU and R.VANCEA, Nouv.J.Chim. (in press).
5. Yu.A.SCHREIDER, ed., "The Monte Carlo Method", Pergamon Press, New York, 1970, p.271.
6. C.HANSCH and P.MOSER, Immunochem., 15, 535 (1978).
7. J.J.RAY, QCPE Program No. 226, Indiana University, Bloomington, Indiana.

Table 1. Phenylsuccinamate haptens : experimental data and various partitions of the molecular Van de Waals space

No.	R	$-\Delta F^{o}_{rel}$	V_{a1}	V_{a2}	V_{b1}	V_{b2}	V_{b3}	V_{b4}
1.	H	-1.90	7.397	74.570	20.280	20.334	20.854	20.493
2.	2-CH_3	-2.75	8.892	88.670	20.280	20.007	20.854	28.481
3.	3-CH_3	-1.72	8.892	104.567	28.567	28.007	28.734	28.701
4.	4-CH_3	-1.73	10.471	93.287	20.280	30.940	20.854	31.634
5.	2-Cl	-1.96	7.397	96.684	20.280	31.469	20.854	31.478
6.	3-Cl	-1.59	7.397	90.095	20.280	28.110	20.854	28.248
7.	4-Cl	-0.81	7.397	90.670	24.366	24.338	24.987	24.376
8.	2,4-di-Cl	-1.71	8.892	104.775	24.366	32.011	24.987	32.303
9.	2,5-di-Cl	-1,33	7.397	97.081	25.995	25.914	26.522	26.047
10.	2-Br	-1.94	14.281	100.263	20.280	36.323	20.854	37.087
11.	3-Br	-1.43	7.397	107.129	20.280	34.668	20.854	36.810
12.	4-Br	-0.74	7.397	107.652	28.700	28.449	29.170	28.730
13.	2-I	-2.02	12.132	88.328	20.280	29.399	20.854	29.927
14.	3-I	-0.43	10.915	86.265	18.226	30.199	18.418	30.337
15.	4-I	-0.18	7.397	93.546	25.103	25.025	25.565	25.250
16.	2-NO_2	-1.74	8.847	89.018	20.280	28.112	20.854	28.619
17.	3-NO_2	-1.17	7.397	90.528	20.280	28.448	24.527	28.343
18.	4-NO_2	-0.61	7.397	90.841	23.850	24.894	24.894	24.967

($-\Delta F^{o}_{rel}$ are taken from ref. 6).

Table 1. (continued)

No.	V_1	V_2	V_3	V_4	V_5	V_6	V_7	V_8	WV
1.	1.756	19.098	2.049	18.444	1.678	18.602	1.914	18.420	81.77
2.	1.756	19.098	2.732	25.689	1.678	18.602	2.726	25.281	97.31
3.	1.756	26.978	2.732	25.689	1.678	26.889	2.726	25.281	113.46
4.	1.756	19.098	3.569	28.065	1.678	18.602	3.468	27.472	103.44
5.	1.756	19.098	2.049	29.429	1.678	18.602	1.914	29.555	103.44
6.	1.756	19.098	2.049	26.199	1.678	18.602	1.914	26.196	97.24
7.	1.756	23.231	2.049	22.327	1.678	22.688	1.914	22.424	97.24
8.	1.756	23.231	2.732	29.571	1.678	22.688	2.726	29.285	113.46
9.	1.756	24.766	2.049	23.998	1.678	24.317	1.914	24.000	103.44
10.	1.756	19.098	5.429	31.658	1.678	18.602	5.418	30.905	114.21
11.	1.756	19.098	2.049	34.761	1.678	18.602	1.914	34.668	114.21
12.	1.756	27.414	2.049	26.681	1.678	27.022	1.914	26.535	114.21
13.	1.756	19.098	4.395	25.532	1.678	18.602	4.303	25.096	100.21
14.	3.187	15.231	2.288	28.049	3.244	14.982	2.196	28.003	100.21
15.	1.756	23.809	2.049	23.201	1.678	23.425	1.914	23.111	100.21
16.	1.756	19.098	2.744	25.875	1.678	18.602	2.669	25.443	97.62
17.	1.756	19.098	2.049	26.294	1.678	18.602	1.914	26.534	97.62
18.	1.756	22.771	2.049	22.918	1.678	22.172	1.914	22.980	97.62

$WV = \sum_{i=1}^{8} V_i$; $V_{a1} = V_1+V_3+V_5+V_7$; $V_{a2} = V_2+V_4+V_6+V_8$; $V_{b1} = V_5+V_6$; $V_{b2} = V_7+V_8$

$V_{b3} = V_1+V_2$; $V_{b4} = V_3+V_4$