Molecular Shape Descriptors.
1. Three-Dimensional Molecular Shape Descriptor

I. Motoc [a, b], G. R. Marshall [a], R. A. Dammkoehler [b], and J. Labanowski [a]

Z. Naturforsch. **40a**, 1108–1113 (1985); received May 25, 1985

Dedicated to Prof. Dr. Oskar E. Polansky in honor of his 66th birthday

The paper presents and illustrates a method which uses numerical integration of the van der Waals envelope(s) to calculate with desired accuracy the molecular van der Waals volume and the three-dimensional molecular shape descriptor defined as the twin-number [OV(α, β); NOV(β, α)], where OV and NOV represent the overlapping and, respectively, the non-overlapping van der Waals volumes of the molecules α and β superimposed according to appropriate criteria.

1. Introduction

The spatial occupancy (shape and size) of molecules can be conveniently defined in the framework of the hard sphere approximation: each atom constituting the molecule is represented by an isotropic sphere centered at the equilibrium position of the atomic nucleus and having a radius equal to the van der Waals (VdW) radius of the atom. A molecular VdW envelope may be uniquely defined as the surface of the intersection of the VdW spheres associated with the atoms in the molecule; consequently, the total volume inside the VdW envelope represents the molecular VdW volume (V_W) – for further discussion concerning molecular VdW area and volume the reader is referred to [1, 2] and references cited therein.

Obviously, for a given molecular geometry, the spatial extension of the molecular model depends to a major degree on the values of the atomic VdW radii considered. Illustratively, Table 1 systematizes four sets of atomic VdW radii: the $r_{W,E}$ values were obtained by Bondi [3, 4] from a comparison of various types of physical properties of molecules; the $r_{W,MM}$ set correspond [5] to a minimum in the VdW function used in molecular mechanics [6]; the $r_{W,QM}$ values were derived by quantum mechanics calculations using a density contour approach [7]; the $r_{W,\sigma}$ values reflect the effective distance of closest approach of non-bonded atoms in crystals [8] – they were determined by comparison of their effects on the predicted sterically allowed conformations of peptides with observed crystal structures [8, 9], and imply a repulsive interaction of approximately 5 kcal/mole at closest approach.

There are two points to be made about Table 1: (i) the effective "size" of an atom or molecule varies with the phenomenon studied; and (ii) in agreement with previous results [10], the correlations collected in Table 2 argue that the (semi) empirically derived sets of VdW radii are highly linearly related. This indicates a similar balance between the values of VdW radii for different atoms within these sets and the less satisfactory correlation $r_{W,\sigma}$ vs. $r_{W,QM}$ reflects a different balance within the $r_{W,QM}$ set.

Table 1. Atomic van der Waals radii, r_W [Å].

Atom	$r_{W,\sigma}$	$r_{W,E}$	$R_{W,MM}$	$r_{W,QM}$
H	1.08	1.20	1.50	1.24
C	1.53	1.70	1.75	1.54
O	1.36	1.50	1.65	1.27
N	1.45	1.55	1.70	1.36
F	1.30	1.47	1.60	1.18
Cl	1.65	1.75	1.95	1.63
Br	1.80	1.85	2.10	1.79
I	2.05	1.98	2.25	2.02
S	1.70	1.80	2.00	1.75
P	1.75	1.80	2.05	1.87

Reprint request to Dr. Ioan Motoc, Department of Physiology and Biophysics, Box 8101, Washington University School of Medicine, St. Louis, Missouri 63110, USA.

[a] Department of Physiology and Biophysics, Washington University School of Medicine, St. Louis, Missouri 63110, USA.

[b] Department of Computer Science, Washington University, St. Louis, Missouri 63130, USA.

0340-4811 / 85 / 1100-1108 $ 01.30/0. – Please order a reprint rather than making your own copy.

Table 2. Results of correlations with Eq. $r_{W,\sigma} = a_0 + a_1 r_{W,x}$.

$r_{W,x}$	a_0	a_1	t_{a0} [a]	t_{a_1} [a]	r [b]	s [c]	F [d]
$r_{W,E}$	−0.434	1.205	−24.38	14.73	0.982	0.053	94.965
$r_{W,MM}$	−0.503	1.116	−28.68	14.96	0.983	0.052	97.947
$r_{W,QM}$	0.139	0.913	4.74	8.66	0.951	0.087	32.803

[a] Student t test for the significance of the regression coefficients. [b] Correlation coefficient. [c] Standard deviation. [d] F test for significance of the correlation.

As the calculated sterically allowed conformations using the $r_{W,\sigma}$ values compare very well [8, 9] with the observations from known protein structures, we choose to use $r_{W,\sigma}$ as a standard measure of the effective size of an atom and refer to these values as r values.

The paper presents and illustrates a method which uses numerical integration of the van der Waals envelope to calculate accurate estimate of molecular van der Waals volume and the three-dimensional, conformation-dependent molecular shape descriptors related to the overlapping and non-overlapping van der Waals volumes of molecules. Also, our method offers a greatly improved computational approach of the descriptors used in Molecular Shape Analysis; and, an extension of the formalism allows the calculation of the three-dimensional version of the Minimal Topological Difference descriptor.

2. Molecular van der Waals Volume

The molecular van der Waals volume (V_W) is a useful parameter in the study of packing density of organic molecules in crystals [11], protein folding [12, 13], properties of condensed phases [1, 4], quantification of bioactivity as a function of molecular structure [14–16].

Within the hard sphere approximation discussed in the previous Section, the locus of points (x, y, z) residing inside the van der Waals envelope of a molecule δ satisfies at least one of the following inequations:

$$(A_I - x)^2 + (B_I - y)^2 + (C_I - z)^2 \leqq r_I^2 ,$$
$$I = 1, 2, \ldots, N_\delta , \tag{1}$$

where (A_I, B_I, C_I) represent the cartesian coordinates of the N_δ atoms constituting the molecule, and r_I is the atomic VdW radius of the atom I.

We turn to numerical integration techniques for obtaining accurate estimate of V_W using (1) to characterize the spatial extension of the molecular model considered. Therefore, one determines the finite, countable set $\boldsymbol{W}$ whose elements $W_j(W_{jx}, W_{jy}, W_{jz})$ are points in the Euclidean 3D-space,

$$\boldsymbol{W} = \{W_j \mid P_1, P_2\} , \tag{2}$$

and satisfy the following properties:

(P_1): W_j, for all j, lie within the parallelipiped which embeds the collections of spheres (1),

$$d_1 \leqq W_{jx} \leqq D_1 , \quad d_2 \leqq W_{jy} \leqq D_2 ,$$
$$d_3 \leqq W_{jz} \leqq D_3 . \tag{3}$$

(P_2): W_j, for all j, are uniformly distributed, independent random points. Next, one determines the set $\boldsymbol{V}^\delta$,

$$\boldsymbol{V}^\delta = \{V_j^\delta \mid P_3, P_4\} \tag{4}$$

whose elements $V_j^\delta(V_{jx}^\delta, V_{jy}^\delta, V_{jz}^\delta)$ satisfy the properties:

(P_3): $\boldsymbol{V}^\delta \subset \boldsymbol{W}$.

(P_4): V_j^δ, for all j, satisfy at least one of the N_δ inequations (1).

As V_W may be regarded as a non-negative, continuous function in the closed bounded domain defined by the VdW envelope, the V_W value is estimates by [17, 18]:

$$V_W = g \text{ card } \boldsymbol{V}^\delta / \text{card } \boldsymbol{W}, \tag{5}$$

where g represent the volume of the parallelipiped (3), and card denotes the cardinal number.

The $\boldsymbol{W}$ set is constructed using either Monte Carlo [17, 18] or strictly deterministic procedures. Within the first procedure $W_j \in \boldsymbol{W}$ is given by

$$W_{jx} = d_1 + (D_1 - d_1)\,\zeta_1 , \quad W_{jy} = d_2 + (D_2 - d_2)\,\zeta_2 ,$$
$$W_{jz} = d_3 + (D_3 - d_3)\,\zeta_3 , \tag{6}$$

where $d_1, \ldots, D_3$ define the parallelipiped which embeds the VdW envelope of the molecule considered; $(\zeta_1, \zeta_2, \zeta_3)$ are uniformly distributed, independent random sequences on the unit interval and were generated using standard algorithms [17b]. The deterministic procedure, which simulates the stochastic approach (6), consists in dividing the

parallelipiped (3) into card $\boldsymbol{W}$ subspaces ("elementary" parallelipipeds) whose centres (W_{jx}, W_{jy}, W_{jz}) define $W_j \in \boldsymbol{W}$.

The accuracy of the estimate (5), for given value of g, is proportional to (card $\boldsymbol{W}$)$^{-1/2}$; this circumstance causes the relatively slow convergence of the procedure and imposes utilization of $\boldsymbol{W}$ sets with cardinality of one to two million when large molecules are considered.

We have computed by numerical integration of the VdW envelope the van der Waals volume (A^3) of ten amino acids (Table 3).

The estimated standard deviation of the V_w values is 0.1. The very good correlation of the calculated V_w values with the corresponding apparent molal volumes (AMV),

$$\begin{aligned} &\mathrm{AMV} = -22.67 + 1.87\, V_w\,, \\ &(t = 8.07)\ (t = 13.98)\,, \\ &(r = 0.98, s = 8.38, F = 85.49)\,, \end{aligned} \tag{7}$$

assumedly reflects that the $r_{w,\sigma}$ atomic VdW radii offer an adequate description of the spatial extension of molecules.

The numerical integration of the VdW envelope, by comparison with geometrical approaches currently used [3, 11, 14], or recently developed [21, 22], yields V_W estimates of significantly greater accuracy (for quantitative comparison see [8]). Therefore, our method is of immediate interest if physically significant estimates of observed data (e.g., the unitary free energy of solvation) are required; then, the accuracy of the molecular van der Waals volume estimate must be maximized.

Table 3. Van der Waals volume of amino acids [a].

Amino acid	V_W [Å^3]	Apparent molal volume [b] [Å^3]
Glycine	54.21	77.50 [c]
Alamine	66.62	105.61
Valine	91.71	158.04
Leucine	104.59	184.45
Serine	74.35	113.48
Threonine	86.78	135.04
Asparagine	92.05	142.12
Glutamine	105.32	157.50
Arginine	130.0	219.01
Methionine	107.8	184.80

[a] The $r_{W,\sigma}$ values used in calculations are reported in Table 1.
[b] Estimated from molecular weight and packing density given in [19].
[c] Taken from [20].

3. Three-Dimensional Molecular Shape Descriptor

The quantitative characterization of molecular shape represents one of the major obstacles in formulating quantitative (chemical) structure – (biological) activity relationships (QSAR) for structurally diverse (non-congeneric) and/or conformationally flexible compounds exhibiting common biological action. The wealth of data that exist for non-congeneric and flexible analogs demands that priority be given to the development of molecular shape descriptors and appropriate strategies to allow for their inclusion in the conceptual framework of the QSAR approach.

The purpose of this section is to introduce the three-dimensional, conformation-dependent molecular shape descriptor which allows quantitative comparison of the molecular shape of widely differing structures. The computational method is discussed in detail and illustrated with a simple example. For the justification of the molecular shape descriptor and its use to generate QSAR's, the reader is referred to [23].

The three-dimensional molecular shape descriptor (3D-MSD) we introduce is pairwise related to the molecular van der Waals volume of the molecules α and β being compared. The descriptor is defined as the twin-number

$$3\,\mathrm{D\text{-}MSD}(\alpha, \beta) = [\mathrm{OV}(\alpha, \beta); \mathrm{NOV}(\beta, \alpha)]\,; \tag{8}$$

$\mathrm{OV}(\alpha, \beta)$ is the overlapping VdW volume of the two molecules, and $\mathrm{NOV}(\beta, \alpha)$ is the non-overlapping VdW volume of molecule β superimposed over molecule α.

The method outlined in Sect. 2 can be extended to obtain accurate estimates of OV and NOV volumes. Therefore, one determines the set $\boldsymbol{W}$,

$$\boldsymbol{W} = \{W_i \mid P'_1, P_2\}\,, \tag{9}$$

where P'_1 reads:

(P'_1): W_i, for all j, lie within the parallelipiped which embeds the collection of $N_\alpha + N_\beta$ spheres (1), which represents the superimposed molecules α and β.

Next, one determines the sets

$$\boldsymbol{V}^{\alpha} = \{V_j^{\alpha} | P_3', P_4'\}, \tag{10}$$

$$\boldsymbol{V}^{\beta} = \{V_j^{\beta} | P_3'', P_4''\}, \tag{11}$$

where P_3', P_3'' and P_4', P_4'' represents P_3 and P_4 with $\delta = \alpha$ or β, respectively, and $\boldsymbol{W}$ is specified by (9).

Utilization of the sets (9), (10), (11) and elementary operations of set algebra allow calculation of the following:

(i) $\mathrm{OV}(\alpha, \beta)$ – the overlapping volume of the molecules α and β,

$$\mathrm{OV}(\alpha, \beta) = g \,\mathrm{card}\,(\boldsymbol{V}^{\alpha} \cap \boldsymbol{V}^{\beta})/\mathrm{card}\,\boldsymbol{W}, \tag{12}$$

(ii) $\mathrm{NOV}(\beta, \alpha)$ – the non-overlapping volume of molecule β superimposed over molecule α,

$$\mathrm{NOV}(\beta, \alpha) = g \,\mathrm{card}\,(\boldsymbol{V}^{\beta} - \boldsymbol{V}^{\alpha})/\mathrm{card}\,\boldsymbol{W}, \tag{13}$$

and

(iii) $\mathrm{NOV}(\alpha, \beta)$ – the non-overlapping volume of molecule α superimposed over molecule β,

$$\mathrm{NOV}(\alpha, \beta) = g \,\mathrm{card}\,(\boldsymbol{V}^{\alpha} - \boldsymbol{V}^{\beta})/\mathrm{card}\,\boldsymbol{W}. \tag{14}$$

Here, g represents the volume of the parallelipiped which embeds the VdW envelopes of the compared molecules α and β;

$$\boldsymbol{V}^{\alpha} - \boldsymbol{V}^{\beta} = \boldsymbol{V}^{\alpha} \backslash \boldsymbol{V}^{\beta} = \{V_j | V_j \in \boldsymbol{V}^{\alpha} \quad \text{and} \quad V_j \notin \boldsymbol{V}^{\beta}\}$$

is the relative complement of $\boldsymbol{V}^{\alpha}$ in $\boldsymbol{V}^{\beta}$, and $\boldsymbol{V}^{\beta} - \boldsymbol{V}^{\alpha} = \boldsymbol{V}^{\beta} \backslash \boldsymbol{V}^{\alpha}$ is the relative complement of $\boldsymbol{V}^{\beta}$ in $\boldsymbol{V}^{\alpha}$. The set (9) is constructed using the procedures described in Section 2, and, obviously, the accuracy of the estimates (12), (13), and (14) shows the same dependence on card $\boldsymbol{W}$ as the estimate (5).

The $[\mathrm{OV}(\alpha, \beta); \mathrm{NOV}(\beta, \alpha)]$ descriptor represents the three-dimensional extension of the topological steric Molecular Descriptor $\mathrm{SMD} = [\mathrm{SMD_c}; \mathrm{SMD_w}]$ which led to excellent correlations for a variety of cases [24] and, also, the quantitative version of the molecular graphic approach referred to as the Active Analog Approach (AAA) [25] which have demonstrated qualitative predictability and wide applicability. It is easily seen that (13) and (14) allow accurate calculation of the three-dimensional version of the Minimal Topological Difference (MTD) descriptor [26, 27]. With the formalism used here, the MTD descriptor is conveniently defined [28] as:

$$\mathrm{MTD} = g \,\mathrm{card}\,(\boldsymbol{V}^{\alpha} \backslash \boldsymbol{V}^{\beta} \cup \boldsymbol{V}^{\beta} \backslash \boldsymbol{V}^{\alpha})/\mathrm{card}\,\boldsymbol{W}. \tag{15}$$

Using (13), (14) and (16),

$$\boldsymbol{V}^{\alpha} \backslash \boldsymbol{V}^{b} \cap \boldsymbol{V}^{\beta} \backslash \boldsymbol{V}^{\alpha} = \Phi \tag{16}$$

it follows that

$$\begin{aligned} \mathrm{MTD} &= g\,(\mathrm{card}\,\boldsymbol{V}^{\alpha} \backslash \boldsymbol{V}^{\beta} + \mathrm{card}\,\boldsymbol{V}^{\beta} \backslash \boldsymbol{V}^{\alpha})/\mathrm{card}\,\boldsymbol{W} \\ &= \mathrm{NOV}(\alpha, \beta) + \mathrm{NOV}(\beta, \alpha). \end{aligned} \tag{17}$$

Further, the Molecular Shape Analysis (MSA) descriptor, V_0, [29, 30], is clearly the $\mathrm{OV}(\alpha, \beta)$ and, therefore, it is a particular case of the 3D-MSD. It should be noted that the MSA procedure to evaluate $V_0(\alpha, \beta)$ overestimates its value and introduces a sizeable error which increases with the branching of molecules considered; this deficiency is eliminated in the 3D-MSD approach.

The 3D-MSD offers a very convenient basis from which the shape (dis)similarity of molecules can be described numerically and compared quantitatively.

Consider the molecules $M_1, M_2, \ldots, M_n$ and specify an internal reference system by choosing the molecule M_r, $1 \leqq r \leqq n$, as reference molecule (i.e., each molecule M_i, $1 \leqq i \leqq n$, is characterized by reference to the same molecule M_r). It follows that in the framework defined above, the shape of each molecule M_i is characterized by the two-dimensional row vector

$$\text{3D-MSD}\,(r, i) = [\mathrm{OV}(r, i); \mathrm{NOV}(i, r)], \quad 1 \leqq i \leqq n$$

and the n vectors are directly comparable. The relation between the shape of the molecules M_p and M_q, $1 \leqq p, q \leqq n$, can be expressed by the familiar Euclidean distance (metric) [31], which in this case reads:

$$\begin{aligned} \mathrm{d}(M_p, M_q) = [&|\mathrm{OV}(r, p) - \mathrm{OV}(r, q)|^2 \\ &+ |\mathrm{NOV}(p, r) - \mathrm{NOV}(q, r)|^2]^{1/2} \geqq 0. \end{aligned} \tag{18}$$

The value of $\mathrm{d}(M_p, M_q)$ is said to be the distance between the molecules M_p and M_q with respect to the shape characteristics {OV, NOV}. Complementary to the notion of distance $\mathrm{d}(M_p, M_q)$ is the idea of similarity $s(M_p, M_q)$ between M_p and M_q with respect to the same set of characteristics {OV, NOV}. The non-negative, real-valued function $s(M_p, M_q)$ is a similarity measure [32, 33] if

$$0 \leqq s(M_p, M_q) \leqq 1, \text{ for } M_p \neq M_q, \; s(M_p, M_p) = 1,$$

and $s(M_p, M_q) = s(M_q, M_p)$.

The metric $\mathrm{d}(M_p, M_q)$ can be used [32, 33] to construct a similarity measure $s(d)$ if $s(d \to \infty) = 0$

and $s(d=0)=1$. We arbitrarily choose as a similarity measure the following single-valued, monotorically-decreasing function of d:

$$s(M_p, M_q) = [1 + d(M_p, M_q)]^{-1}. \quad (19)$$

Since the transformation of d into s is equivocal (e.g., the Gaussian $\exp(-\alpha d^2)$ may also represent the function $s(d)$), the relevant information is the ordering of $\{M_1, M_2, \ldots, M_n\}$ by (19), and, in particular, the degree of similarity of each molecule considered with the chosen reference molecule.

Using (18) and the triangle inequality, it follows that the distance $d(M_p, M_q)$ between the molecules M_p and M_q with respect to the whole set of characteristics considered and the partial distance $d^{\delta}(M_p, M_q)$ defined with respect to some of their characteristics satisfy the inequality

$$d^{\delta}(M_p, M_q) \leqq d(M_p, M_q), \quad (20)$$

and, consequently,

$$s(d^{\delta}) \geqq s(d); \quad (21)$$

e.g., the Molecular Shape Analysis will generally assign higher shape similarity between M_p and M_q than the 3D-MSD because {OV} is a subset of {OV, NOV}.

Fluorene (M_1), anthracene (M_2), and phenanthrene (M_3) are selected to illustrate numerically the three-dimensional shape descriptor (8), the distance function (18), and the similarity measure (19). The shape comparison criterion (i.e., the superposition of the compared molecules over the reference molecule) is illustrated in Figure 1.

The 3D-MSD was calculated using standard molecular geometries, $r_{w,\sigma}$ atomic van der Waals radii, and the Monte Carlo method to generate $\boldsymbol{W}$ sets with card $\boldsymbol{W} = 500{,}000$.

Table 4 collects the values of the 3D-MSD calculated with M_1, M_2, and, respectively, M_3 as reference structure; for comparison, values calculated with card $\boldsymbol{W} = 1{,}000{,}000$ and, respectively, card $\boldsymbol{W} = 500{,}000$ and the $\boldsymbol{W}$ set generated using the deterministic procedure are also listed. Note that

$$\mathrm{OV}(M_r, M_r) = V_{W,M_r} \quad \text{and}$$
$$\mathrm{OV}(M_r, M_i) + \mathrm{NOV}(M_i, M_r) = V_{W,M_i},$$

for all i; inspection of Table 4 shows that the OV and NOV values satisfy these equations with remarkable accuracy.

Table 5 systematizes the values of the distance function $d(M_r, M_i)$, with $r = 1, 2$, and, respectively, 3 and $i = 1, 2, 3$. Using these values with the similarity measure (19) one obtains, as expected, that, for example, with fluorene as reference structure, phenanthrene resembles closer then anthracene the shape of fluorene, i.e., $s(M_1, M_3) = 0.04$ and $s(M_1, M_2) = 0.03$, or, with anthracene as reference structure, $s(M_2, M_1) = 0.030$ and $s(M_2, M_3) = 0.028$

Table 4. The 3D-MSD values [Å^3] for fluorene (M_1), anthracene (M_2), and phenanthrene (M_3) (the $\boldsymbol{W}$ set is generated by Monte Carlo method, and card $\boldsymbol{W} = 500{,}000$).

M_i	Reference structure (M_r)	OV(M_r, M_i)	NOV(M_i, M_r)
M_1	M_1	126.18	0.00
		(126.12) [a]	(0.00) [a]
		(126.19) [b]	(0.00) [b]
M_2		108.67	26.62
		(108.54) [a]	(26.41) [a]
		(108.69) [b]	(26.59) [b]
M_3		114.15	20.99
		(114.26) [a]	(20.90) [a]
		(114.15) [b]	(20.97) [b]
M_1	M_2	108.72	17.49
M_2		135.33	0.00
M_3		110.89	24.34
M_1	M_3	114.23	11.98
M_2		110.98	24.47
M_3		135.29	0.00

[a] The $\boldsymbol{W}$ set is generated using the deterministic procedure, and card $\boldsymbol{W} = 500{,}000$.
[b] The $\boldsymbol{W}$ set is generated using the Monte Carlo procedure, and card $\boldsymbol{W} = 1{,}000{,}000$.

Table 5. Values of distance function [Å^3] for fluorene (M_1), anthracene (M_2), and phenanthrene (M_3) calculated with the 3D-MSD values of Table 4.

M_i	d(M_1, M_i)	d(M_2, M_i)	d(M_3, M_i)
M_1	0.00	31.84	24.23
M_2	31.86	0.00	34.49
M_3	24.19	34.48	0.00

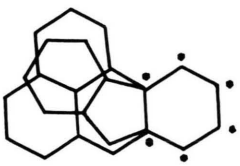

Fig. 1. Shape comparison of fluorene, anthracene, and phenanthrene; the pairs of atoms matched by a least-squares fit in reference and compared molecules are marked by asterisk.

indicate a greater shape similarity between anthracene and fluorene then between anthracene and phenanthrene.

The 3D-MSD approach offers a conformation-dependent, three-dimensional quantitative characterization of molecular shape. The [OV (α, β); NOV (β, α)] descriptor is specifically designed to account for the effect of molecular shape on bioactivity and, therefore, its intended use is generation of quantitative structure-activity relationships (a QSAR investigation of dihydrofolate reductase inhibition based upon 3D-MSD approach is reported in [23]). It is worth noting that since this descriptor is not restricted to bioactive compounds which are structurally congeneric and conformationally rigid, it has the ability of encompassing, within a single QSAR, molecules of widely differing shapes.

Acknowledgement

This work was supported by NIH grant GM24483.

[1] R. S. Pearlman, in Physical Chemical Properties of Drugs, eds. S. H. Yalkowsky, A. A. Sinkula, and S. C. Valvani, Dekker, New York 1980.
[2] M. L. Connolly, J. Appl. Cryst. **16,** 548 (1983).
[3] A. Bondi, J. Phys. Chem. **68,** 441 (1964); ibid. **80,** 3006 (1966).
[4] A. Bondi, Physical Properties of Molecular Crystals, Liquids, and Glasses, Wiley, New York 1968.
[5] N. L. Allinger, Adv. Phys. Org. Chem. **13,** 1 (1976).
[6] U. Burkert and N. L. Allinger, Molecular Mechanics, ACS Monograph **177,** ACS, Washington (D.C.) 1982.
[7] R. J. Boyd, J. Phys. B. **10,** 2283 (1977).
[8] I. Motoc and G. R. Marshall, Chem. Phys. Lett. **116,** 415 (1985).
[9] G. R. Marshall, in preparation.
[10] M. Charton, Top. Curr. Chem. **114,** 57 (1983), J. Amer. Chem. Soc. **101,** 7356 (1979).
[11] A. I. Kitaigorodsky, Molecular Crystals and Molecules, Academic Press, New York 1973.
[12] J. N. Finney, Proc. Roy. Soc., London **A319,** 479 (1970).
[13] B. Lee and F. M. Richards, J. Mol. Biol. **55,** 379 (1971); F. M. Richards, ibid. **82,** 1 (1974).
[14] I. Moriguchi, Y. Kanada, and K. Komatsu, Chem. Pharm. Bull. **24,** 1799 (1976).
[15] M. Tichy, Int. J. Quantum. Chem. **16,** 509 (1979).
[16] M. Charton and I. Motoc, Steric Effects in Drug Design, Springer, Berlin 1983.
[17] a) J. M. Hammersley and D. C. Handscomb, Monte Carlo Methods, Wiley, New York 1964. b) D. E. Knuth, The Art of Computer Programming. Seminumerical Algorithms, **Vol. 1,** Addision-Wesley, Reading (Mass.) 1981.
[18] B. P. Demidovici and I. I. Maron, Computational Mathematics, Mir, Moscow 1973.
[19] G. V. Gurskaya, The Molecular Structure of Amino Acids, Consultants Bureau, New York (N.Y.) 1968.
[20] E. J. Cohn, T. L. McMeckin, J. T. Edsall, and J. H. Weare, J. Amer. Chem. Soc. **56,** 2282 (1934).
[21] R. Pavani and G. Ranghino, Computers and Chemistry **6,** 133 (1983).
[22] R. S. Pearlman, personal communication.
[23] I. Motoc and G. R. Marshall, Z. Naturforsch., in press.
[24] I. Motoc, Steric fit in QSAR. V. Quantitative treatment of the shape of biomolecules, Preprint, Timisoara University, Serie Chimie No. 5, 1980; I. Motoc, Arzneim.-Forsch./Drug Res. **31,** 290 (1981); I. Motoc, I. Muscutariu, and D. Ciubotariu, Kexue Tongbao **26,** 361 (1981); R. Valceanu and I. Motoc, Rev. Roum. Chim. **27,** 225 (1982); I. Motoc, Z. Naturforsch. **38a,** 1342 (1983); I. Motoc, Quant. Struct.-Act. Relat. **3,** 43, 47 (1984).
[25] G. R. Marshall and C. D. Barry, A Functional Representation of Molecular Volume for Computer-aided Drug Design, in Abstr. Meeting Amer. Crystallography Assoc., Honolulu 1979; C. Humblet and G. R. Marshall, Ann. Rep. Med. Chem. **15,** 26 (1980); J. R. Sulfrin, D. A. Dunn, and G. R. Marshall, Mol. Pharmacol. **19,** 307 (1981); G. R. Marshall, in Quantitative Approaches to Drug Design, J. C. Dearden (ed.), Elsevier, Amsterdam 1983; W. E. Klunk, B. L. Kalman, J. A. Ferrendelli, and D. F. Covey, Mol. Pharmacol. **23,** 511 (1983).
[26] Z. Simon, A. Chiriac, I. Motoc, S. Holban, D. Ciubotariu, and Z. Szabadai, Studia Biophys. Berlin **55,** 217 (1976).
[27] A. T. Balaban, A. Chiriac, I. Motoc, and Z. Simon, Steric Fit in QSAR, Lecture Notes in Chemistry, **Vol. 15,** Springer, Berlin 1980; I. Motoc, Top. Curr. Chem. **114,** 93 (1983).
[28] I. Motoc, S. Holban, R. Vancea, and Z. Simon, Studia Biophys. Berlin **66,** 75 (1977); I. Motoc, Math. Chem. **4,** 113 (1978).
[29] A. J. Hopfinger, J. Amer. Chem. Soc. **102,** 7196 (1980).
[30] A. J. Hopfinger, J. Med. Chem. **24,** 818 (1981); C. Battershell, D. Malhotra, and A. J. Hopfinger, ibid. **24,** 812 (1981); A. J. Hopfinger, Arch. Biochem. Biophys. **206,** 153 (1981).
[31] B. S. Duran and P. L. Odell, Cluster Analysis, Springer, Berlin 1974, p. 1–10.
[32] J. Fabian, A. Mehlhorn, and F. Fratev, Int. J. Quantum Chem. **17,** 235 (1980), and references cited therein.
[33] A. Mehlhorn, F. Fratev, O. E. Polansky, and V. Monev, Math. Chem. **15,** 3 (1984).