Spectral Moments of the Edge-Adjacency Matrix of Molecular Graphs. 2. Molecules Containing Heteroatoms and QSAR Applications

Ernesto Estrada†

Departamento de Diseño de Fármacos, Centro de Bioactivos Químicos, Universidad Central de Las Villas, Santa Clara 54830, Villa Clara, Cuba

Received July 5, 1996⊗

A novel substructural approach to structure−property and structure−activity relationships studies is proposed. The method is based on the computation of spectral moments of the edge-weighted adjacency matrix. In the present work, bond distances are used as bond weights in the diagonal entries of edge matrix. Spectral moments of the edge-weighted adjacency matrix are expressed in terms of structural fragments present in molecules via algebraic expressions. Boiling points of 58 alkyl halides were well described by the present approach, and this property was expressed as a linear combination of structural fragments of the studied molecules. QSAR applications of the novel approach are explored by studying the antifungal activity of benzyl alcohols. An equation expressing contributions of each substituent in benzyl alcohols to the antifungal activity is obtained. The methodology proposed here permits the rational design of novel compounds with improved activity, which is illustrated for benzyl alcohol derivatives.

1. INTRODUCTION

The graph theoretical approach appears to be an important alternative to computer-aided molecular design methods.[1] The most used and very well-known graph theoretical methods are those based on topological indices.[2] These theoretical indices are numbers that describe the structural information of molecules through graph theoretical invariants.[3] Description of the molecular structure by a simple number can appear as a loss of information.[4] However, the use of topological indices in the successful description of physicochemical and biological data proves the usefulness of these methods.[5]

At present, one of the most important criticisms of topological indices is related with the physical meaning of some of these descriptors.[6] Moreover, several models obtained by using a combination of topological indices are very convoluted, and their physical significance remains occult. For instance, Sablijc and Trinajstić[5] have remarked that the complexity of some equations obtained with molecular connectivity indices is forbidding.

Another approach used in chemical graph theory is the representation of molecules not by simple numbers but using vectors or sequences of numbers. These vectors can be built by using different graph invariants that produce sequences of numbers with structural information about the molecular graph. Among the invariants that produce sequences of numbers we can mention the spectra of the vertex-adjacency matrix,[7] the count of path numbers in the graph,[8] the number of self-avoiding paths,[9] random walks,[10] self-returning walks,[11] basic graphs,[12] etc.

In the previous paper[13] of this series we proposed a new graph theoretical invariant that produces a sequence of numbers for a molecular graph. This invariant is based on the spectral moments of the edge-adjacency matrix. The use of spectral moments of this edge matrix for the description of physical and physicochemical properties of alkanes has produced very significant models from the statistical point of view. These models have also been interpreted in terms of the different fragments of molecular graphs, permitting a direct relation between the studied properties and the structural features of molecules.[13]

In closing, this novel graph theoretical invariant has shown two interesting features to be used in quantitative structure−property (QSPR) and structure−activity relationships (QSAR) studies. These features are (i) very good correlation with physical and physicochemical properties of organic molecules and (ii) a direct structural interpretation in terms of molecular fragments.

A necessary step further in the development of any graph theoretical approach to chemical structure is to include the differentiation of heteroatoms present in molecules. The neglect of heteroatoms and spatial features of molecules has been considered by Randić as severe limitations of graph theoretical models.[14]

Several approaches have been reported in the literature to avoid the problem of differentiation of heteroatoms. Molecular connectivity indices[15] have been modified to discriminate among heteroatoms using different ways. For instance, Kier and Hall introduced the concept of valence connectivity,[15] and Kupchik used the van der Waals radii of heteroatoms to substitute the previously used simple connectivity of atoms in molecules.[16] Quantum chemical parameters, such as bond orders[17,18] and electron charge densities[19] have also been used in molecular connectivity schemes to account for heteroatoms differentiation and stereochemical features of molecules. On the other hand, Randić has used variable weights in the diagonal entries of adjacency matrix in order to calculate connectivity indices[14] and weighted path numbers[20] which produce discrimination of heteroatoms. Estrada[21] used different weights in the nondiagonal entries of the edge-adjacency matrix as a way to differentiate heteroatoms in the edge connectivity index.[22]

In this work we outline the inclusion of heteroatoms' differentiation into the formalism of the spectral moments of the edge adjacency matrix. This approach is based on

† E-mail address: farmaco@cbqvc.sld.cu.

⊗ Abstract published in *Advance ACS Abstracts,* October 15, 1996.

S0095-2338(96)00113-8 CCC: $14.00

Table 1. Average Bond Distances for Different Bonds in Organic Molecules[a]

bond[b]	distance (Å)	bond[b]	distance (Å)	bond[b]	distance (Å)	bond[b]	distance (Å)
C4−C4	1.54	C3−C3	1.46	C2−C1	1.63	C2≡C2	1.20
C4−C3	1.52	C3−C2	1.45	C3═C3	1.34	C2≡N1	1.16
C4−C2	1.46	C3−N3[c]	1.40	C3═C2	1.31	C3═C3[e]	1.40
C4−N3	1.47	C3−N2	1.40	C3═N2	1.32	C3═N2[e]	1.34
C4−N2	1.47	C3−O2	1.36	C3═O1	1.22	C3−F[e]	1.32
C4−O2	1.43	C3−F	1.33	C2═C2	1.28	C3−Cl[e]	1.71
C4−S2	1.81	C3−Cl	1.73	C2═N2	1.32	C3−Br[e]	1.89
C4−F	1.40	C2−C2	1.38	C2═O1	1.16	C3−I[e]	2.10
C4−Cl	1.76	C2−N3	1.33	N3═O1[d]	1.24	C3−O2[e]	1.36
C4−Br	1.94	C2−O2	1.33	N2═O1[d]	1.22	C3−N3[e]	1.48
C4−I	2.13	C2−F	1.30	N2═N2	1.25	C3−S2[e]	1.81

[a] Taken from ref 28. [b] The atomic symbol is followed by a number which indicates the number of neighbor atoms to which it is bonded. [c] 1.32 Å for N−C═O group. [d] For partial double bonds in NO and NO_2. [e] C3 in a benzene ring.

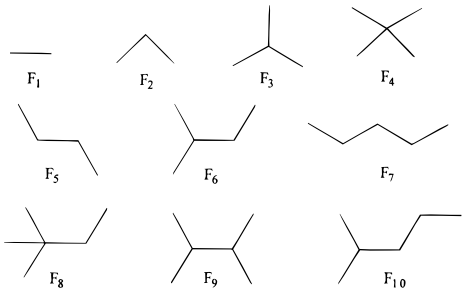

Figure 1. Pictorial representation of some structural fragments of simple molecular graphs.

the use of bond distances as weights in the diagonal entries of the edge matrix.

2. OUTLINE OF THE APPROACH

The edge adjacency matrix **E** of a (molecular) graph $G = (V, E)$ is a square and symmetric matrix, whose elements e_{ij} are ones or zeros if the edges i and j are adjacent or not, respectively. In this approach, like in many others graph theoretical representation of chemical structures, the vertices of the molecular graph represent atoms and edges represent bonds in molecules. The edge-adjacency matrix has been explicitly defined in the chemical graph theory literature,[23] but has received little attention in both chemical and mathematical literature until recently.[23] The rediscovery by the present author of the edge-adjacency matrix (**E**) as an important source of graph theoretical invariants has produced a series of useful molecular descriptors.[18,21,22] We are also pointed out that the edge-adjacency matrix of a molecular graph G is identical to the vertex-adjacency matrix of the line graph of G.[24]

Spectral moments of the **E** matrix have been defined as follows:[13]

$$\mu_k = \mu_k(G, \mathbf{E}) = tr(\mathbf{E}^k) \tag{1}$$

In this expression, $\mu_k(G, \mathbf{E})$ is the kth spectral moment of the **E** matrix, *tr* is the trace of the corresponding matrix, and k is an integer exponent. The analogous concept of spectral moments of the vertex-adjacency matrix **A**, μ_k (G, **A**), has been studied in the chemical literature with different objectives.[25]

In order to account for the discrimination of heteroatoms we propose the introduction of weights in the diagonal entries of the edge adjacency matrix. There are many forms of weighting the edges in a molecular graph. In the present approach we consider the use of bond distances as an effective way to provide the heteroatoms's differentiation in the molecular graph. Bond distances have a series of interesting features that make them attractive as edge weights in this type of approach. They are quantities that can be measured experimentally or estimated from theoretical calculations with acceptable accuracy. Molecular mechanics[26] and quantum chemical geometry optimization[27] are two of the most important methods used to estimate bond distances.

Spectral moments $\mu_k^d = \mu_k^d(G, \mathbf{E})$ of the edge-weighted adjacency matrix can be straightforwardly calculated by using an expression analogous to the eq 1. These spectral moments can be expressed in terms of structural fragments of the corresponding molecular graph in full analogy to the equations obtained for moments of the simple edge adjacency matrix.[13] If we denote by $d_i = d(e_i)$ the distance corresponding to the edge i, i.e., $d(e_i) = d(v_u, v_t)$ in which v_u and v_t are incident to e_i, the first four spectral moments of the edge-weighted adjacency matrix for acyclic graphs can be expressed as

$$\mu_0^d = |F_1| = m \tag{2}$$

$$\mu_1^d = \sum_i d_i \tag{3}$$

$$\mu_2^d = 2|F_2| + \sum_i (d_i)^2 \tag{4}$$

$$\mu_3^d = 6|F_3| + \sum_i (d_i)^3 + 3\sum_i F_2^i d_i \tag{5}$$

$$\mu_4^d = 2|F_2| + 12|F_3| + 24|F_4| + 4|F_5| + \sum_i (d_i)^4 + 4\sum_i F_2^i (d_i)^2 + 8\sum_i F_3^i d_i + 4\sum_r (d_i d_j)_r \tag{6}$$

In eqs 2−6 the terms are as follows, $|F_k|$ is the number of fragments of type k in the graph G, m is the number of edges in G, F_i^i is the number of fragments F_k containing the bond i, and the sum $\sum_r$ in eq 6 is carried out for all adjacent edges in the graph. Fragments $F_k (k = 1,2,...,18)$ were depicted in the first paper of this series[13] and some of them are illustrated in Figure 1.

In the present study we use standard average bond distances for the different bonds in molecules.[28] Some of

Table 2. Spectral Moments

$$\mu_1^d = 1.54|Fd_1| + 1.40|Fd_2| + 1.76|Fd_3| + 1.94|Fd_4| + 2.13|Fd_5|$$
$$\mu_2^d = 2|F_2| + 2.37|Fd_1| + 1.96|Fd_2| + 3.10|Fd_3| + 3.76|Fd_4| + 4.54|Fd_5|$$
$$\mu_3^d = 3.65|Fd_1| + 2.74|Fd_2| + 5.45|Fd_3| + 7.30|Fd_4| + 9.66|Fd_5| + 8.40|Fd_6| + 8.82|Fd_7| + 9.48|Fd_8| + 10.02|Fd_9| + 10.59|Fd_{10}| + 10.56|Fd_{11}| + 9.90|Fd_{12}| + 11.10|Fd_{13}| + 11.67|Fd_{14}| + 11.64|Fd_{15}| + 10.44|Fd_{16}| + 12.21|Fd_{17}| + 12.78|Fd_{18}| + 11.01|Fd_{19}| + 9.24|Fd_{20}| + 6|F_3|$$
$$\mu_4^d = 5.62|Fd_1| + 3.84|Fd_2| + 9.59|Fd_3| + 14.16|Fd_4| + 20.58|Fd_5| + 25.52|Fd_6| + 27.93|Fd_7| + 32.07|Fd_8| + 35.77|Fd_9| + 39.91|Fd_{10}| + 39.18|Fd_{11}| + 34.72|Fd_{12}| + 43.08|Fd_{13}| + 47.54|Fd_{14}| + 47.14|Fd_{15}| + 38.50|Fd_{16}| + 51.72|Fd_{17}| + 56.46|Fd_{18}| + 42.76|Fd_{19}| + 30.46|Fd_{20}| + 47.84|Fd_{21}| + 46.72|Fd_{22}| + 49.60|Fd_{23}| + 51.04|Fd_{24}| + 52.56|Fd_{25}| + 48.48|Fd_{26}| + 49.92|Fd_{27}| + 51.44|Fd_{28}| + 51.36|Fd_{29}| + 52.80|Fd_{30}| + 54.32|Fd_{31}| + 54.24|Fd_{32}| + 55.76|Fd_{33}| + 57.28|Fd_{34}| + 55.68|Fd_{35}| + 57.20|Fd_{36}| + 52.48|Fd_{37}| + 57.12|Fd_{38}| + 58.64|Fd_{39}| + 53.93|Fd_{40}| + 60.16|Fd_{41}| + 55.44|Fd_{42}| + 50.72|Fd_{43}| + 60.08|Fd_{44}| + 55.36|Fd_{45}| + 61.60|Fd_{46}| + 56.88|Fd_{47}| + 52.16|Fd_{48}| + 58.40|Fd_{49}| + 53.68|Fd_{50}| + 45.60|Fd_{51}| + 54.24|Fd_{52}| + 58.56|Fd_{53}| + 63.12|Fd_{54}| + 48.96|Fd_{55}| + 24|F_4| + 4|F_5|$$

[a] Fragments F_k and Fd_k are illustrated in Figures 1 and 2, respectively.

these bond distances which will be used in the present work are depicted in Table 1.

Polynomial expressions 2−6 can be transformed to linear combinations of structural fragments present in molecules. We can find a quantitative relation between a property P and the spectral moments of **E** having, for instance, the following appearance:

$$P = a_0 + a_1\mu_1 + a_2\mu_2 + a_3\mu_3 + a_4\mu_4 \qquad (7)$$

From eq 7 property P can be expressed in terms of the different fragments present in these molecules via an expression of the following form:

$$P = a_0 + \sum_i (b_{CX}|CX|)_i + \sum_j (b_{C(XY)}|C(XY)|)_j + \sum_k (b_{C(XYZ)}|C(XYZ)|)_k + a_4(24F_4 + 4F_5)$$

In this expression indexes i, j and k run over the different one-, two-, and three-bond fragments in the molecules, respectively. For instance, $(|CX|)_1 = |CC|$, $(|CX|)_2 = |CF|$, ..., $(|C(XY)|)_1 = |C(CC)|$, $(|C(XY)|)_2 = |C(FF)|$, and so forth. Coefficients b_{CX}, $b_{C(XY)}$, and $b_{C(XYZ)}$ can be determined by the following expressions:

$$b_{CX} = a_1 d_{CX} + a_2(d_{CX})^2 + a_3(d_{CX})^3 + a_4(d_{CX})^4 \qquad (8)$$

$$b_{C(XY)} = 2a_2 + 3a_3(d_{CX} + d_{CY}) + a_4\{2 + 4[(d_{CX})^2 + (d_{CY})^2] + 4(d_{CX}d_{CY})\} \qquad (9)$$

$$b_{C(XYZ)} = 6a_3 + a_4[12 + 8(d_{CX} + d_{CY} + d_{CZ})] \qquad (10)$$

In these expression d_{CX} is the distance for the bond C−X and the a_i's are the coefficients in eq 7. As an example we will construct the spectral moments of edge-weighted adjacency matrix of alkyl halides by using this substructural approach. Expressions for the first four spectral moments of alkyl halides are depicted in Table 2. Structures of the different fragments Fd_i given in these expressions are illustrated in Figure 2.

3. QUANTITATIVE STRUCTURE−PROPERTY STUDIES

To consider a theoretical invariant as a molecular descriptor it needs to have several desirable attributes.[29] For practical purposes, it is convenient that candidates to molecular descriptors have good correlations with at least one physical property. In the present work we have selected boiling points of a series of alkyl halides to carry out this test. This kind of compound was selected because the compounds have great structural variability, which include different heteroatoms and branching of the alkyl chain. This series was composed of 58 methyl and ethyl halides having any combination of halogen atoms in the structure.

The first six spectral moments of a weighted-edge adjacency matrix were calculated for the series of alkyl halides. These values were used to describe normal boiling points (BP) of 47 compounds forming the calibration data set. The remaining 11 compounds chosen at random were placed in an external prediction set. Compounds in the prediction set were never used in the development of the quantitative models during this work; they were only used to validate thd model found.

The best linear regression equation for the description of normal boiling points of alkyl halides is illustrated below:

$$\text{BP (K)} = 167.01(\mu_1^d - \mu_0^d) + 1.01 \times 10^{-2}\mu_2^d - 5.99\mu_3^d + 0.69\mu_4^d + 154.82 \qquad (11)$$

$$n = 43 \quad r = 0.9922 \quad \text{rms} = 7.86 \text{ K} \quad F = 600.2$$

Experimental and calculated boiling points of the 58 alkyl halides are given in Table 3.

In the development of the quantitative model for the description of normal boiling points of the calibration data set, four compounds were detected as statistical outliers. Outliers detection was carried out by using the following standard statistical tests:[30] residual, standardized residuals, studentized residual and Cooks' distance. The four compounds were bromoiodomethane, iodomethane, 1,2-dichloro-1,2-diiodoethane, and 1,1,1-dichlorofluoroethane. There seems to be no distinct structural relation among these compounds.

Statistical parameters in eq 11 suggest a high quality of the model found. The correlation coefficient r is over 0.99, and the root of medium square error (rms) is only 7.86 K. The plot of residual values against calculated boiling points showed no observable pattern. The last validation procedure carried out for eq 11 was the test for its predictive ability by using the 11-member external prediction set. The rms error

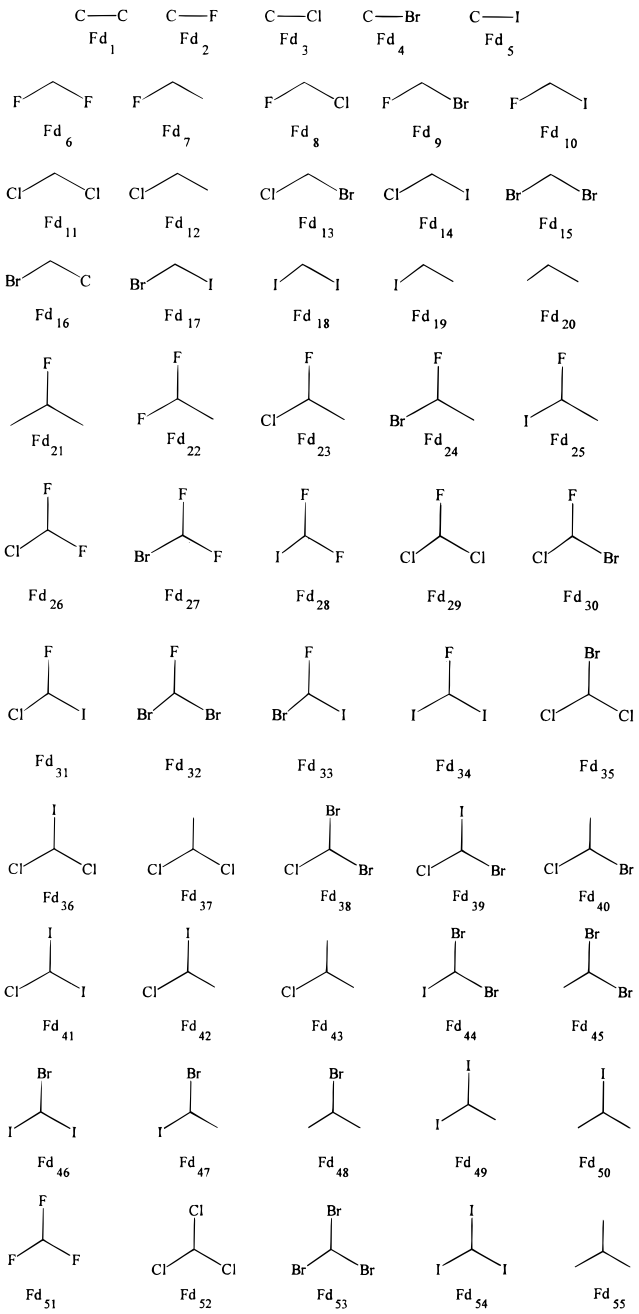

Figure 2. Structural fragments present in alkyl halides which are contained in the first four spectral moments of the edge-weighted adjacency matrix.

in the prediction set was only 8.91 K, which evidences the good predictive power of the model found.

In eq 11, and in all models found in this work, we have preferred the use of a combination of two spectral moments of the edge-weighted matrix. This combination was justified by the fact that both indices are highly collinear; they are the count of edge and weighted edges in the graphs, respectively. The analytical form of this combination, $(\mu_1^d - \mu_0^d)$, was suggested by the coefficients of each index in equations when they were used as independent variables. This kind of combination of molecular descriptors has been widely used in the literature.[31] However, the problem of collineality between independent variables, i.e., spectral moments, is not important in the present approach. Spectral moments can be expressed as linear combinations of structural fragments, and then identical fragments are summed up forming new variables. These new variables are the independent structural fragments of the corresponding molecular graph. In the final expression, the property *P* is expressed in terms of such structural fragments, and, as a consequence, the collineality of spectral moments has no importance at the end of this study. This final expression has a similar appearance to those obtained by using Smolenskii[32] or embedding frequencies approaches.[33]

By substituting spectral moments of the **E** matrix into eq 11 for its expressions in terms of structural fragments given in Table 2, we can obtain the quantitative contribution of each fragment to the boiling points in alkyl halides. The general quantitative expression that relates normal boiling points of alkyl halides to the structural fragments in molecules is given in Table 4.

An analysis of this equation shows interesting features of alkyl halides. Coefficients of all different bond contributions to boiling points are positive. All two-bond fragments, X−C−Y, show negative coefficients with a slight variation of absolute value of the coefficient. Finally, fragments C(XYZ) show positive and negative contributions to the boiling points of alkyl halides. There are some resemblances in the general trends of the values of coefficients for the different fragments in such compounds. The molecular weight of the fragment makes a great contribution to the values of the coefficients. For instance, the greatest contribution of a single bond to the boiling point of alkyl halides is carried out by the C−I bond, which has the greatest molecular weight. Similar results are observed for the two- and three-bond fragments, for which the greatest contributions are made by fragments CI_2 and CI_3. However, there are some influences of the number of fluorine atoms present in the fragments which diminish the absolute value of the coefficients. For instance, the coefficient of fragment C−F is smaller than that of C−C which has a smaller molecular weight. This trend is also observed in the coefficients of fragments C(XY) and C(XYZ). For the three-bond fragment the contribution of the molecular weight MW accounts for the 88% of the variance in the coefficients, while the combination of molecular weight and the number of fluorine atoms in the fragment N_F, account for more than the 98% of the variance in the coefficients. Equation 12 express quantitatively this relation:

$$\text{coeff} = 3.05 \times 10^{-2}\,\text{MW} - 1.398N_F - 1.544 \quad (12)$$

$$r = 0.991 \quad s = 0.46 \quad F = 894$$

The fragments C(XYZ) can be related to the branching of the molecular framework. As a consequence, the coefficients for these fragments should be expected to be negative similarly to those obtained for alkane hydrocarbons. This "normal" contribution is observed for the *tert*-butyl fragment $C(C)_3$ in Table 4, fragment Fd_{55}. However, contributions due to molecular weight or due to the presence of fluorine atoms in the structural fragment produce positive and negative deviations from the "normal" behavior observed in alkanes.[13,34]

The results obtained here point out that bond distances are excellent edge weights for the calculation of spectral moments of the **E** matrix. On the other hand, "real" distances for the C−F bond diminish with the increase of fluorine atoms in a molecular fragment, e.g., C−F distances in CH_3F,

Table 3. Experimental and Calculated Values of Normal Boiling Points of Alkyl Halides

no.	compd	BP (K)	BP (K)	res.[b]	no.	compd	BP (K)	BP (K)	res.[b]
1	CH_3Br	276.75	277.85	−1.10	30	CH_2ClCH_2I[c]	413.15	404.03	9.12
2	CH_2BrCl	341.25	341.89	−0.64	31	CH_3CHBr_2	381.15	365.72	15.43
3	$CHBrCl_2$	363.15	371.91	−8.76	32	CH_2BrCH_2Br	404.45	403.68	0.77
4	CBr_2Cl_2[c]	423.35	410.96	12.39	33	$CH_2BrCBrCl_2$	451.45	451.12	0.33
5	CH_2BrI[d]	412.25			34	$CHBrClCHBrCl$[d]	468.15		
6	CBr_2F_2[c]	297.65	312.42	−14.77	35	$CBrF_2CBrF_2$	319.55	314.49	5.06
7	$CBrCl_3$	377.85	385.67	−7.82	36	$CClF_2CF_3$[c]	235.15	222.97	12.18
8	CH_3Cl	248.95	255.72	−6.77	37	CH_3CHCl_2	330.45	321.69	8.76
9	$CClBr_3$	431.65	432.98	−1.33	38	CH_3CCl_2F[d]	305.15		
10	CH_2Br_2	370.15	363.57	6.58	39	CCl_2FCF_3	309.15	309.15	0.00
11	$CHCl_2I$[c]	405.15	394.88	10.27	40	CF_2ClCCl_3[c]	364.65	368.51	−3.86
12	$CFBr_2Cl$	353.45	359.91	−6.46	41	CH_2ClCH_2Cl	356.65	360.66	−4.01
13	CH_2Cl_2	319.15	320.30	−1.15	42	$CH_2ClCClF_2$	319.95	309.08	10.87
14	CH_2F_2	221.55	228.09	−6.54	43	$CH_2ClCHClF$	346.85	346.25	0.60
15	CH_2BrF[c]	293.15	295.63	−2.48	44	$CFCl_2CFCl_2$	364.65	365.26	0.39
16	CH_2I[d]	455.15			45	CH_2ICH_2I	452.65	447.10	5.20
17	CHI_2F	375.65	391.55	−15.90	46	CH_3CH_2F	235.45	246.41	−10.96
18	CH_2FI	326.55	316.86	9.69	47	CH_3CH_2I	345.45	335.45	10.00
19	CH_3I	315.55	299.86	15.69	48	$CHCl_2CCl_3$	435.15	443.76	−8.61
20	CCl_3I	415.15	409.70	5.45	49	$CFCl_2CCl_3$[c]	409.15	417.09	−7.94
21	CCl_4	349.65	362.18	−12.53	50	$CHBr_2CHBr_2$	516.65	515.39	1.26
22	CBr_4[c]	462.15	458.52	3.63	51	CH_2ClCCl_3	403.65	406.25	−2.60
23	$CHCl_3$	334.85	349.87	−15.02	52	$CHCl_2CHCl_2$	419.35	427.41	−8.06
24	$CHBr_3$[c]	422.65	417.02	5.63	53	$CH_2BrCHBr_2$	462.05	458.10	3.95
25	CHI_3	491.15	484.02	7.13	54	CH_3CCl_3	347.25	332.54	14.71
26	CH_3CH_2Br	311.55	313.98	−2.43	55	CCl_3CF_3	318.98	317.56	1.39
27	CH_2BrCH_2Cl	380.15	382.20	−2.05	56	$CHCl_2CH_2Cl$	386.95	392.66	−5.71
28	CH_2BrCH_2F	344.65	336.11	8.54	57	$CHFClCHCl_2$[c]	376.15	382.11	−5.96
29	CH_3CH_2Cl	385.45	392.50	−7.05	58	CCl_2FCF_2Cl	320.85	316.85	4.00

[a] Normal boiling point, taken from Weast, R. D. CRC Handbook of Chemistry and Physics, 67th ed.; 3rd printing; CRC: Boca Raton FL, 1987. [b] Observed − calculated. [c] Compound in external prediction set. [d] Compound detected as an outlier in the training set.

Table 4. Normal Boiling Points of Alkyl Halides Expressed in Terms of Structural Fragments[a]

$$\begin{aligned}\mathrm{BP\ (K)} = {} & 72.21|Fd_1| + 53.04|Fd_2| + 100.92|Fd_3| + 123.07|Fd_4| + 145.08|Fd_5| - 32.69|Fd_6| - \\ & 33.53|Fd_7| - 34.62|Fd_8| - 35.33|Fd_9| - 35.87|Fd_{10}| - 36.21|Fd_{11}| - 35.32|Fd_{12}| - \\ & 36.73|Fd_{13}| - 37.08|Fd_{14}| - 37.16|Fd_{15}| - 35.96|Fd_{16}| - 37.42|Fd_{17}| - 37.59|Fd_{18}| - \\ & 36.43|Fd_{19}| - 34.31|Fd_{20}| - 2.93|Fd_{21}| - 3.70|Fd_{22}| - 1.72|Fd_{23}| - 0.72|Fd_{24}| + \\ & 0.33|Fd_{25}| - 2.49|Fd_{26}| - 1.50|Fd_{27}| - 0.45|Fd_{28}| - 0.50|Fd_{29}| + 0.49|Fd_{30}| + \\ & 1.54|Fd_{31}| + 1.49|Fd_{32}| + 2.53|Fd_{33}| + 3.58|Fd_{34}| + 2.48|Fd_{35}| + 3.53|Fd_{36}| + \\ & 0.27|Fd_{37}| + 3.47|Fd_{38}| + 4.52|Fd_{39}| + 1.26|Fd_{40}| + 5.57|Fd_{41}| + 2.31|Fd_{42}| + \\ & -0.94|Fd_{43}| + 5.52|Fd_{44}| + 2.26|Fd_{45}| + 6.56|Fd_{46}| + 3.31|Fd_{47}| + 0.05|Fd_{48}| + \\ & 4.36|Fd_{49}| + 1.10|Fd_{50}| - 4.48|Fd_{51}| + 1.49|Fd_{52}| + 4.47|Fd_{53}| + 7.61|Fd_{54}| - \\ & 2.16|Fd_{55}| + 16.56|F_4|\ 2.76|F_5| + 154.82\end{aligned}$$

[a] Fragments F_k and Fd_k are illustrated in Figures 1 and 2, respectively.

CH_2F_2, CHF_3, and CF_4 are 1.385, 1.358, 1.332, and 1.325, respectively. As a consequence the use of these "real" bond distances as edge weights in the present approach should produce significant improvements in QSPR results.

4. QUANTITATIVE STRUCTURE−ACTIVITY STUDIES

The study of quantitative relationships between chemical structure and biological activity of organic compounds follows the same general methodology as the QSPR studies.[4,35] However, the development of the biological activity by certain molecules is a very complex process that includes several physicochemical stages, such as partition between different phases and formation of drug-receptor complex through hydrophobic, electronic, or steric interactions.[35] This fact has justified the use of several approaches to the study of quantitative structure−activity relationships (QSAR) in different series of bioactive compounds. Among these approaches we can mention the free-energy scheme of Hansch,[36] additivity models of Free−Wilson and Fujita−Ban,[37] pattern recognition techniques,[38] comparative molecular field analysis (CoMFA),[39] semiempirical or *ab initio* quantum chemical approaches,[40] and the use of graph theoretical concepts, specially that of topological indices.[4,15,41,42] A systematization of structure−activity relationships was carried out by Trinajstić, Randić, and Klein.[43] They classified the different existing approaches as structure-cryptic, structure-implicit, and structure-explicit schemes.

In the present study we propose a novel approach to QSAR studies based on the use of spectral moments of edge-weighted adjacency matrix of molecular graphs. This approach, as all the structure-explicit schemes,[43] uses a well-defined mathematical invariant that has a direct structural interpretation. In this case, the structural interpretation is given in terms of substructural fragments of the molecules which have some resemblance with the additivity schemes of Free−Wilson and Fujita−Ban.[37]

To perform the present QSAR study we have selected a series of 19 benzyl alcohols with antifungal activity measured by the inhibitory power against *Aspergillus niger*.[44] This data set was used by Kier[45] as an example to test the usability

Table 5. Observed and Calculated Antifungal Activity of Benzyl Alcohol Derivatives for *Aspergillus niger*

substituent	log (1/*C*) obsd[a]	log (1/*C*) calcd[b]	residual
none	1.51	1.29	0.22
4-Cl	2.07	2.06	0.01
2,4-Cl_2	3.07	2.82	0.25
3,4-Cl_2	3.07	2.82	0.25
2,4,5-Cl_3	3.32	3.57	−0.25
3,4,5-Cl_3	3.63	3.57	0.06
2-Br	2.15	2.30	−0.15
4-Br	2.27	2.32	−0.05
4-I	2.75	2.61	0.14
4-CH_3	1.79	1.79	0.00
2,4-$(CH_3)_2$	2.14	2.27	−0.13
3,5-$(CH_3)_2$, 4-Cl	3.05	3.02	0.02
3,5-$(CH_3)_2$, 4-I	3.42	3.57	−0.15
2-NO_2	2.49	2.03	0.46
4-NO_2	2.00	2.04	−0.04
4-CN	1.67	1.85	−0.18
2-OH	1.39	1.54	−0.15
3-OH	1.39	1.55	−0.16
4-OH	1.39	1.55	−0.16

[a] Taken from ref 45. [b] Calculated from expression 13.

Table 6. Antifungal Activity of Benzyl Alcohol Derivatives in Terms of Structural Fragments[a]

$$\log(1/C) = 1.294 + 1.13|S_1| + 1.41|S_2| + 1.74|S_3| + 0.83|S_4| + 0.72|S_5| + 0.24|S_6| + 0.75|S_7| + 0.35|S_8| + 0.58|S_9| - 0.096|S_{10}| - 0.107|S_{11}| - 0.121|S_{12}| - 0.085|S_{13}| - 0.083|S_{14}| - 0.073|S_{15}| - 0.060|S_{16}| - 0.081|S_{17}| - 0.070|S_{18}| - 0.077|S_{19}| - 0.149|S_{20}| - 0.154|S_{21}| - 0.159|S_{22}| - 0.145|S_{23}| - 0.144|S_{24}| - 0.134|S_{25}| - 0.143|S_{26}| - 0.141|S_{27}| - 0.012|F_5|$$

[a] Statistical parameters are the same as in eq 13.

and structural meaning of the path cluster connectivity index of fourth order $^4\chi_{PC}$. The equation obtained by this author using a combination of $^1\chi^v$ and $^4\chi_{PC}$ indices accounted for the 88% of the variance in the data. The model obtained by us is illustrated below:

$$pC = 1.634(\mu_1^d - \mu_0^d) - 3.099 \times 10^{-3}\mu_4^d - 3.078 \quad (13)$$

$$r = 0.967 \quad s = 0.201 \quad F = 115$$

Expression 13 accounts for more than the 93% of the variance in the inhibitory activity of substituted benzyl alcohols against *Aspergillus niger*. The calculated values of biological activities for the 19 compounds are given in Table 5 together with the experimental values and the residuals. The greater deviation from the experiment is observed for the calculated activity of compound 14, for which the residual is 0.46. This compound has a nitro group in the ortho position which can interact via an intramolecular hydrogen bond with the hydroxyl group, forming a classical six membered ring. This kind of interaction has not been considered in the present study.

With the present approach we have obtained a model with two variables, similar to that obtained by Kier,[45] with improved statistical parameters of the regression. However, the main advantage of the spectral moment approach is not only concerning the statistical quality of models found but it is related to the easy structural interpretation of results. In this sense, we have obtained the contribution of each substructure in benzyl alcohol derivatives to the antifungal activity. This equation is illustrated in Table 6 and Figure 3.

The equation shown in Table 6 can be transformed to account for the contribution of each substituent in benzyl

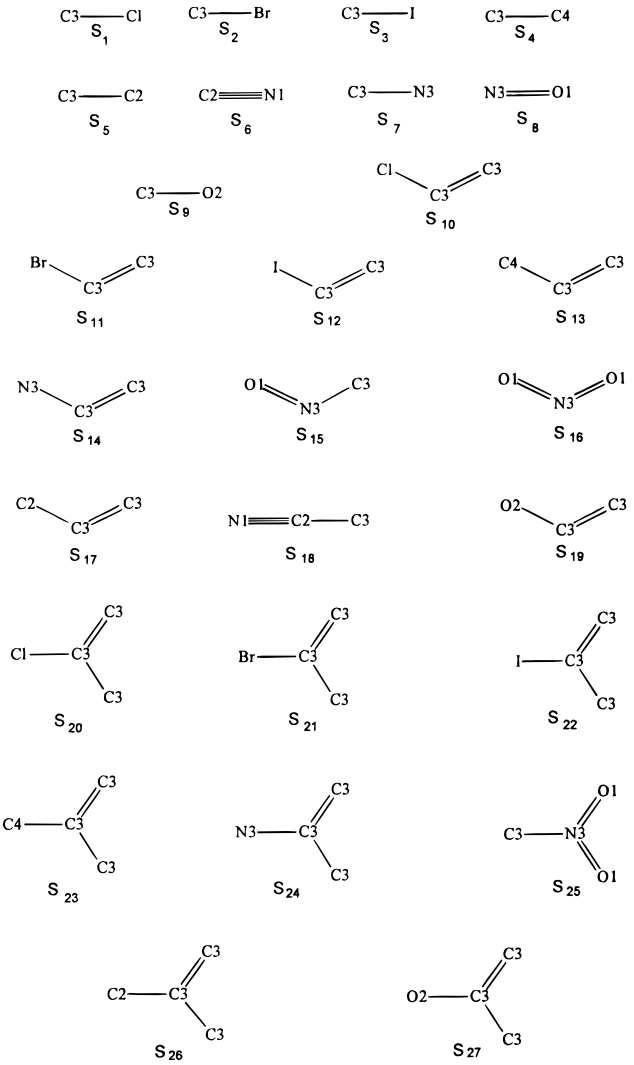

Figure 3. Structural fragments contained in the first four spectral moments of benzyl alcohol derivatives. Atomic symbols as in Table 1.

alcohol derivatives to the biological activity. For instance, the contribution of each chlorine atom to the antifungal activity is obtained by summing the coefficients of fragments S_1, S_{10}, and S_{20}; contribution of each nitro group is obtained in a similar way by summing the coefficients of fragments S_7, S_8, S_{14}, S_{15}, S_{16}, S_{24}, and S_{25}. Following this procedure for all substituents present in the studied benzyl alcohols we obtain eq 15. This equation expresses the particular influence of each substituent in benzyl alcohol derivatives to the inhibitory potency against *Aspergillus niger*.

$$pC = 1.294 + 0.789|Cl| + 1.042|Br| + 1.339|I| + 0.515|CH_3| + 0.585|CN| + 0.800|NO_2| + 0.285|OH| - 0.012|F_5| \quad (14)$$

In this equation each coefficient represent the contribution of the corresponding substituent to the biological activity. These coefficients contain no information about the influence of the substituent's position in the phenyl ring. However, the term $|F_5|$ is directly related to contributions from different sites of substitution in the aromatic ring. In Table 7 we give the values of $|F_5|$ for the different positions of substitution in the studied compounds.

As can be appreciated eq 15 permits a completely rational design of novel derivatives of benzyl alcohols with improved

Table 7. Contribution of Substituent's Position in the Aromatic Ring to the Antifungal Activity of Benzyl Alcohol Derivatives

position of substituents	contribution to $\|F_5\|$	position of substituents	contribution to $\|F_5\|$
2	−0.036	2,3,5	−0.096
3	−0.024	2,3,6	−0.108
4	−0.024	2,4,5	−0.096
2,3	−0.072	2,4,6	−0.096
2,4	−0.060	3,4,5	−0.096
2,5	−0.060	2,3,4,5	−0.144
2,6	−0.072	2,3,5,6	−0.144
3,4	−0.060	2,3,4,6	−0.144
3,5	−0.048	2,3,4,5,6	−0.192
2,3,4	−0.108		

antifungal activity, which is the main objective of any QSAR studies.[35] For instance, if we want to synthesize benzyl alcohols with two substituents attached directly to phenyl ring, we need to select the iodine at positions 3 and 5 (see Table 7) in order to obtain the greatest biological activity.

Another important objective of QSAR studies is related to the interpretation of models found in terms of possible drug-receptor interaction mechanisms.[35] There are many ways in which substituents in the studied benzyl alcohols can influence the interaction of such compounds with the biological receptors. The nature of this influence can be hydrophobic, electronic, or steric or a combination of them. We can relate the physicochemical nature of substituents's contributions to the biological activity by expressing the coefficients in eq 15 as a function of substituent constants. In this sense we explore the relation between coefficients and the following substituent constants:[46] π hydrophobic constants, Swain and Lupton electronic constants, molar refractivity MR, and STERIMOL parameters L, B_1, and B_2. The best regression coefficients were obtained for equations using MR and B_1 parameters: $r = 0.958$ and $r = 0.992$, respectively. These results prove the steric nature of contributions carried out by substituents of the antifungal activity of benzyl alcohols and the usefulness of bond distances as edge weights to account for this kind of contribution.

5. CONCLUDING REMARKS

There are some resemblances between the final results obtained with the present approach and those that can be obtained by procedures which combine the use of substructural molecular descriptors and multiple regression analysis, such as the Free−Wilson and Fujita−Ban[37] schemes. For the application of the Free−Wilson model it is necessary to solve a system of linear equations that can be easily resolvable through linear regression analysis. However, to have enough degrees of freedom, we need more than $G + 5$ compounds in the calibration set. The value of G is obtained by adding the number of substituents to each site of substitution in the studied set of molecules plus one. For instance, in the example of benzyl alcohols studied in the present work there are four sites of substitution with 6, 4, 8, and 3 substituent, respectively. This gives a value of $G = 22$ and we need 27 compounds to carry out a Free−Wilson study.

In most of the QSAR studies we need to use an available data set and adapt our theoretical methodologies to it. In this sense the present approach represents a very important scheme for QSAR studies, because we can obtain substructural models without the limitations of Free−Wilson schemes in respect to the number of compounds in the data set.

On the other hand, an equation statistically similar to eq 14 can be obtained by using a most simple approach such as the Hansch−Fujita scheme.[36] This equation is given below:

$$\log(1/C) = 0.977\mathrm{SB}_1 = 2.615 \qquad (15)$$

$$r = 0.955 \quad s = 0.227 \; F = 176$$

In eq 15, SB_1 is the sum of B_1 constants for all substituents in the benzyl alcohol derivative. This equation was suggested by the relationships found by us between the coefficients of each substituent in eq 14 and B_1. We recall that in this example we have four sites of substitution and seven substituent constants. In this case, if we use the traditional Hansch procedure,[36] we need to explore 7175 equations with up to three variables to obtain an equation like eq 15. Moreover, the present approach can be used in such cases to which the data set does not permit the use of a Hansch-type approach.[35] The model found in the present work to describe the boiling points of alkyl halides is a good example of such cases.

The use of spectral moments of the edge-weighted adjacency matrix in QSPR and QSAR studies has strong and weak points like any other theoretical approach to molecular structure. In this sense, the present approach cannot be considered as a universal scheme to be applied successfully to any structure−property or structure−activity study.

The indiscriminate use of topological indices in QSPR and QSAR studies has caused severe criticisms among theoretical drug designers.[47] This fact and the identification of graph theory with the Hückel molecular orbital (HMO) method has produced some misunderstanding of the role of chemical graph theory in theoretical chemistry.[48] However, this indiscriminate use of theoretical approaches is also observed in other branches of theoretical chemistry, such as in the misuse of ab initio quantum chemical methods.[49] On the other hand, we need to recall that chemical structures are very complex and as Dewar has stated[50] *"We do not know, and probably never know, what molecules are "really" like. Our understanding of them is based on models that reproduce their properties well enough to be useful"*. The present approach can be considered as a model of chemical structure that reproduces some properties related to it, and, as a consequence, it is useful.

Use of bond distances as edge weights in the present approach can be considered an example of a general methodology in which different weights can be used to account for different structural properties of organic molecules. Hence, the use of bond charges,[51] bond dipoles,[52] bond energies,[53] or the average of some properties of atoms defining the bond such as atomic charges, electronegativities,[53] electrotopological states,[54] and so forth appears to be prominent edge weights in the present methodology. However, bond distance is an observable quantity which can be obtained with adequate precision by theoretical methods. On the other hand, many other bond properties are defined or directly related to bond distances. For instance, diamagnetic susceptibility depends only on the effective radius of the electronic orbits.[55] However, the relation between this

effective radius of the atoms forming a bond and its bond distance has permitted us to find an excellent relation between spectral moments and diamagnetic susceptibility of alkyl halides.[56]

Finally, we want to state that graph theory brings many opportunities to build theoretical models of chemical structure which permits us the combination of a rigorous mathematical theory and the simple structural concepts used by organic chemists for so long.

REFERENCES AND NOTES

(1) Rouvray, D. Taking a Short Cut to Drug Design. *New Scientist* **1993**, May 29, 35−38.

(2) Balaban, A. T. Applications of Graph Theory to Chemistry. *J. Chem. Inf. Comput. Sci*. **1985**, *25*, 334−343.

(3) Hansen, P. J.; Jurs, P. C. Chemical Applications of Graph Theory. Part I. Fundamentals and Topological Indices. *J. Chem. Educ*. **1988**, *65*, 574−580.

(4) Mihalic, Z.; Trinajstić, N. A Graph-Theoretical Approach to Structure−Property Relationships. *J. Chem. Educ*. **1992**, *69*, 701−712.

(5) Sabljic, A.; Trinajstić, N. Quantitative Structure−Activity Relationships: The Role of Topological Indices. *Acta Pharm. Jugosl*. **1981**, *31*, 189−214.

(6) Albert, A. *Selective Toxicity: The Physico-chemical Basis of Therapy*; Chapman and Hall: London, 1984.

(7) Cvetkovic, D. M.; Doob, M.; Sach, H. *Spectra of Graphs*; VEB: Berlin, 1982.

(8) Randić, M. Characterization of Atoms, Molecules and Classes of Molecules Based on Path Enumeration. *MATCH* **1979**, *7*, 5−64. Randić, M.; Wilkins, C. L. Graph Theoretical Study of Structural Similarity in Benzomorphans. *Int. J. Quantum Chem.: Quantum Biol. Symp*. **1979**, *6*, 55−71. Wilkins, C. L.; Randić, M. A Graph Theoretical Approach to Structure−Property and Structure−Activity Correlations. *Theor. Chim. Acta* **1980**, *58*, 45−68.

(9) Randić, M.; Brissey, G. M.; Spencer, R. B.; Wilkins, C. L. Search for Self-avoiding Path for Molecular Graphs. *Comput. Chem*. **1979**, *3*, 5−13. Randić, M.; Brissey, G. M.; Spencer, R. B.; Wilkins, C. L. Use of Self-avoiding Path for Characterization of Molecular Graphs with Multiple Bonds. *Comput. Chem*. **1980**, *4*, 27−32.

(10) Randić, M. Random Walks and Their Diagnostic Value for Characterization of Atomic Environment. *J. Comput. Chem*. **1980**, *1*, 386−399.

(11) Barysz, M.; Knop, J. V.; Pejakovic, S.; Trinajstić, N. Characterization of Branching. *Polish J. Chem*. **1985**, *59*, 405−432.

(12) Randić, M. Representation of Molecular Graphs by Basic Graphs. *J. Chem. Inf. Comput. Sci*. **1992**, *32*, 57−69.

(13) Estrada, E. Spectral Moments of Edge Adjacency Matrix in Molecular Graphs. 1. Definition and Applications to the Prediction of Physical Properties of Alkanes. *J. Chem. Inf. Comput. Sci*. **1996**, *36*, 844−849.

(14) Randić, M. Novel Graph Theoretical Approach to Heteroatoms in Quantitative Structure-Activity Relationships. *Chemom. Intell. Lab. Syst*. **1991**, *10*, 213−217.

(15) Kier, L. B.; Hall, L. H. *Molecular Connectivity in Chemistry and Drug Research*; Academic Press: New York, 1976.

(16) Kupchik, E. J. General Treatment of Heteroatoms with the Randic Molecular Connectivity Index. *Quant. Struct.-Act. Relat*. **1989**, *8*, 98−103.

(17) Estrada, E.; Montero, L. A. Bond Order Weighted Graphs as Structure-Property Indices. *Mol. Eng*. **1993**, *2*, 363−373.

(18) Estrada, E.; Ramirez, A. Edge Adjacency Relationships and Molecular Topographic Descriptors. Definition and QSAR Applications. *J. Chem. Inf. Comput. Sci*. **1996**, *36*, 837−845.

(19) Estrada, Three-Dimensional Molecular Descriptors Based on Electron Charge Density Weighted Graphs. *J. Chem. Inf. Comput. Sci*. **1995**, *35*, 708−713.

(20) Grossman, S. C.; Jerman-Blazic Dzonova, B.; Randić, M. A Graph Theoretical Approach to Quantitative Structure-Activity Relationship. *Int. J. Quantum Chem.: Quantum Biol. Symp*. **1986**, *12*, 123−139.

(21) Estrada, E. Edge Adjacency Relationships in Molecular Graphs Containing Heteroatoms: A New Topological Index Related to Molar Volume. *J. Chem. Inf. Comput. Sci*. **1995**, *35*, 701−707.

(22) Estrada, E. Edge Adjacency Relationships and A New Topological Index Related to Molecular Volume. *J. Chem. Inf. Comput. Sci*. **1995**, *35*, 31−33.

(23) Rouvray, D. H. In *Chemical Applications of Graph Theory*. *Balaban*, A. T., Ed.; Academic Press: London, 1976; pp 180−181. Trinajstić, N. *Chemical Graph Theory*; CRC Press: Boca Raton, FL, 1983; 2nd ed. 1992; pp 32−33.

(24) Gutman, I.; Estrada, E. Topological Indices Based on Line Graph of the Molecular Graphs. *J. Chem. Inf. Comput. Sci*. **1996**, *36*, 541−543.

(25) Jiang, Y.; Tang, A.; Hoffmann, R. Evaluation of Moments and their Application in Hückel Molecular Orbital Theory. *Theoret. Chim. Acta* **1984**, *66*, 183−192. Cioslowski, J. Decomposition of the Total pi-electron Energy of Polycyclic Hydrocarbons into Benzene Ring Increments. *Chem. Phys. Lett*. **1985**, *122*, 234−236. Cioslowski, J. The Generalized Padé Approximants and the Chemical Graph Theory. *Int. J. Quantum Chem*. **1988**, *34*, 217−224.

(26) Burkert, U.; Allinger, N. L. *Molecular Mechanics*; American Chemical Society: Washington, DC, 1982.

(27) Schlick, T. Optimization Methods in Computational Chemistry. In *Reviews in Computational Chemistry*; Lipkowitz, B., Boyd, B., Eds.; VCH Publishers: New York, 1992; Vol. 3, Chapter 1.

(28) Potapov, V. M. *Stereochemistry*. MIR Publ.: Moscow, 1978. Montero, L. A. *HAVPAC: User Manual*; Universidad de La Habana: Havana, 1991.

(29) Randić, M. Generalized Molecular Descriptors. *J. Math. Chem*. **1991**, *7*, 155−168.

(30) Belsey, D. A.; Kuh, E.; Welsch, R. E. *Regression Diagnostics*; Wiley: New York, 1980.

(31) Randić, M. Search for Optimal Molecular Descriptors. *Croat. Chem. Acta* **1991**, *64*, 43−54. Miyashita, Y.; Ohsako, H.; Okuyama, T.; Sasaki, S.; Randic, M. Computer-Assisted Studies of Structure−Property Relationships Using Graph Invariants. *Magn. Res. Chem*. **1991**, *29*, 362−365.

(32) Smolenskii, E. A. Utilization of the Theory of Graphs for the Calculation of Structure-Additive Properties of Hydrocarbons. *Zh. Fiz. Khim*. **1964**, *38*, 1288−1291.

(33) Klein, D. J. Chemical Graph-Theoretic Cluster Expansions. *Int. J. Quantum Chem.: Quantum Chem. Symp*. **1986**, *20*, 153−171. Poshusta, R. D.; McHughes, M. C. Embedding Frequencies of Trees. *J. Math. Chem*. **1989**, *3*, 193−215. McHughes, M. C.; Poshusta, R. D. Graph-theoretic Cluster Expansions. Thermodynamic Properties of Alkanes. *J. Math. Chem*. **1990**, *4*, 227−249. Schmalz, T. G.; Klein, D. J.; Sandleback, B. L. Chemical Graph-Theoretical Cluster Expansion and Diamagnetic Susceptibility. *J. Chem. Inf. Comput. Sci*. **1992**, *32*, 54−57. Kvasnicka, V.; Pospíchal, J. Simple Construction of Embedding Frequencies of Trees and Rooted Trees. *J. Chem. Inf. Comput. Sci*. **1995**, *35*, 121−128.

(34) Needham, D. E.; Wei, I-C.; Seybold, P. G. Molecular Modeling of the Physical Properties of the Alkanes. *J. Am. Chem. Soc*. **1988**, *110*, 4186−4194.

(35) Franke, R. *Theoretical Drug Design Methods*; Elsevier: Amsterdam, 1984.

(36) Hansch, C. A Quantitative Approach to Biochemical Structure−Activity Relationships. *Acc. Chem. Res*. **1969**, *2*, 232−239.

(37) See ref 35 Chapter 11, pp 254−274.

(38) Stuper, A. J.; Brugger, W. E.; Jurs, P. C. *Computer Assisted Studies of Chemical Structure and Biological Function*; Wiley-Interscience: New York, 1979.

(39) Cramer III, R. D.; Patterson, D. G.; Bunce, J. D. Comparative Molecular Field Analysis (CoMFA). 1. Effect of Shape on Binding of Steroids to Carrier Proteins. *J. Am. Chem. Soc*. **1988**, *110*, 5959−5967.

(40) Kier, L. B. *Molecular Orbital Studies in Chemical Pharmacology*; Springer-Verlag: Berlin, 1970. Sadlej, J. *Semi-empirical Methods of Quantum Chemistry*. Ellis Horwood: Chichester, 1985; Chapter 10.

(41) Randić, M. Nonempirical Approach to Structure−Activity Studies. *Int. J. Quantum Chem.: Quantum Biol. Symp*. **1984**, *11*, 137−153.

(42) Kier, L. B.; Hall, L. H. *Molecular Connectivity in Structure−Activity Analysis*; Wiley: New York, 1986.

(43) Trinajstić, N.; Randic, M.; Klein, D. J. On the Quantitative Structure−Activity Relationships in Drug Research. *Acta Pharm. Jugosl*. **1986**, *36*, 267−279.

(44) Carter, D. V.; Carlton, P. T.; Fenton, A. H.; Housley, J. R.; Lessel, B. The Preparation and the Antibacterial and Antifungal Properties of Some Substituted Benzyl Alcohols. *J. Pharm. Pharmacol*. **1958**, 10, Suppl. 147T−159T.

(45) Kier, L. B. Structural Information from Molecular Connectivity $^4\chi_{PC}$ Index. *J. Pharm. Sci*. **1980**, *69*, 1034−1039.

(46) A detailed discussion about the use of each substituent parameters is given in ref 35. Original works defining each of these parameters are as follows: Fujita, T.; Iwasa, J.; Hansch, C. A New Substituent Constant, π, Derived from Partition Coefficients. *J. Am. Chem. Soc*. **1964**, *36*, 5175−5180. Swain, C. G.; Lupton, E. C. Field and Resonance Components of Substituent Effects. *J. Am. Chem. Soc*. **1968**, *90*, 4328−4337. Verloop, A.; Hoogenstraaten, W.; Tipker, J. In *Drug Design, Vol VII*; Ariens, E. J., Ed.; Academic Press: New York, 1976; p 165.

(47) Kubinyi, H. *QSAR: Hansch Analysis and Related Approaches* (Methods and Principles in Medicinal Chemistry; Mannhold, R., Krogsgaard-Larsen, P., Timmerman, H., Eds.; VCH: Weinheim, 1993; Vol. I.

(48) Trinajstić, N.; Klein, D. J.; Randic, M. On Some Solved and Unsolved Problems of Chemical Graph Theory. *Int. J. Quantum Chem.: Quantum Chem. Symp.* **1986**, *20*, 699−742.
(49) Dewar, M. J. S. Quantum Mechanical Molecular Models. *J. Phys. Chem.* **1985**, *89*, 2145−2150.
(50) Dewar, M. J. S. Chemical Implications of σ Conjuation. *J. Am. Chem. Soc.* **1984**, *106*, 669−682.
(51) See, for instances: Laurencelle, N.; Pacey, P. D. Simple Electrostatic Model for Enthalpies of Formation of Singly Substituted Alkanes. *J. Am. Chem. Soc.* **1993**, *115*, 625−631.
(52) Bykov, G. V. Dipole Moments of Bonds in Organic Compounds. *Russ. J. Phys. Chem.* **1969**, *43*, 310−313.
(53) Pauling, L. *The Nature of the Chemical Bond*; Cornell University Press: New York, 1960.
(54) Kier, L. B.; Hall, L. H. An Electrotopological State Index for Atoms in Molecules. *Pharm. Res.* **1990**, *7*, 801−807. Hall, L. H.; Mohney, B.; Kier, L. B. The Electrotopological State: Structure Information at the Atomic Level for Molecular Graphs. *J. Chem. Inf. Comput. Sci.* **1991**, *31*, 76−82. Hall, L. H.; Kier, L. B. An Index of Electrotopological State for Atoms in Molecules. *J. Math. Chem.* **1991**, *7*, 229−241. Hall, L. H.; Mohney, B.; Kier, L. B. The Electrotopological State: An Atom Index for QSAR. *Quant. Struct-Act. Relat.* **1991**, *10*, 43−51.
(55) Feguson, L. N. *The Modern Structural Theory of Organic Chemistry*; Prentice-Hall: NJ, 1963; pp 53−55.
(56) Estrada, E. Spectral Moments of Edge Adjacency Matrix in Molecular Graphs. 3. Applications to the Prediction of Diamagnetic Susceptibilities. Manuscript in preparation.

CI960113V