Edge-Connectivity Indices in QSPR/QSAR Studies. 1. Comparison to Other Topological Indices in QSPR Studies

Ernesto Estrada*,† and Lissette Rodríguez‡

Department of Organic Chemistry, Faculty of Pharmacy, University of Santiago de Compostela, 15706 Santiago de Compostela, Spain, and Centro de Bioactivos Químicos, Universidad Central de Las Villas, Santa Clara 54830, Villa Clara, Cuba

Received April 5, 1999

The linear independence of the edge-connectivity index to other first-, second-, and third-generation topological indices is demonstrated by using principal component analysis for octane isomers. Most of the topological indices are loaded in one factor, while the edge connectivity is loaded in another independent factor. The edge-connectivity index does not produce linear correlations ($R \leq 0.7$) with any of the almost 40 topological indices studied. This index produced the best single-variable quantitative structure−property relationship (QSPR) models for five of the seven physicochemical properties of octanes studied. It is concluded that the edge connectivity is an independent index containing important structural information to be used in QSPR/QSAR (QSAR = quantitative structure−activity relationship) studies.

INTRODUCTION

The theory of molecular connectivity is almost 25 years old. The first representative of these molecular descriptors is the vertex-connectivity index, χ, introduced by Randic in 1975.[1] This has been one of the most successful single descriptors in quantitative structure−property/structure−activity relationships QSPR/QSAR) studies.[2,3] For instance, in 1993 Randic[4] determined that only a few descriptors were among the best single topological indices for correlating some 20 physicochemical properties of octane isomers. These best topological indices include, among others, the χ index.

A further development in the theory of the molecular connectivity has been the introduction of the edge-connectivity index, ϵ.[5−8] After its introduction as a single molecular descriptor in 1995,[5] several works have appeared in the literature reporting good abilities in modeling different properties of organic compounds. For instance, several physicochemical properties of alkylphenols were well-correlated by the ϵ index,[9] and the cancerogenic activity of some hormones was also well-described by this index.[10] Comparisons of the edge-connectivity index to other topological descriptors, including the vertex-connectivity one in QSPR studies, have also been reported, proving the very good abilities of the ϵ index to describe some physicochemical properties.[11−13] Consequently, Nikolic et al. have claimed that the edge-connectivity index appears to be a very promising topological index for QSPR/QSAR studies.[11]

The objective of the present work is to compare the ϵ index to a series of other topological indices in correlating some physicochemical properties of octane isomers. Among these indices we include some of the best molecular descriptors found by Randic,[4] which are among the most used topological indices, and a series of novel descriptors recently introduced in the literature.

METHODS

According to Balaban[14] the topological indices (TI) can be classified in three classes or generations. The first-generation TI's are those integer numbers obtained from graph-theoretical invariants involving integer numbers. The second-generation TI's are real numbers obtained from manipulation of local graph properties that are integer numbers. Finally, those real numbers obtained from graph invariants involving real numbers compose the third generation of topological indices. The edge-connectivity index is an example of the second generation of TI's. It is based on integer-number bond properties that are transformed in real numbers through the use of the Randic-type formula or invariant as follows:[5]

$$\epsilon = \sum_{\text{adjacent edges}} [\delta(e_i)\,\delta(e_j)]^{-1/2}$$

where $\delta(e_i)$ is the degree of edge (bond) i, defined as the number of edges adjacent to it. Two edges are adjacent if they have a common vertex. This index is identical to the vertex-connectivity index of the line graph of the molecular graph.[15] However, we have demonstrated that in the extension of this index to "higher analogous" it is preferred to use the edges in the graph instead the vertices in the line graph.[8]

To compare the edge connectivity index to other descriptors in QSPR studies, we have selected a series of 37 indices of the first, second, and third generation of topological indices. The symbols and definitions of these topological indices are given in Table 1, together with the references to the original works where they were reported. This list of descriptors is by no means complete, but we think that it is representative of the three generations of graph-theoretical descriptors reported in the literature.

The first point of the present work is the study of the existence or not of linear independence between the edge-

† University of Santiago de Compostela.
‡ Universidad Central de Las Villas.

10.1021/ci990030p CCC: $18.00
Published on Web 10/19/1999

Table 1. Description of the Topological Indices Used in the Present Study

index	description	ref
ϵ	edge-connectivity index	5
W	Wiener index	16
Z	Hosoya index	17
$\tilde{W}$	expanded Wiener index	18
MTI	Schultz's molecular topological index	19
WW	hyper-Wiener index	20
${}^2W_M - 2WW_M$, D^2-CJ_u	Wiener and hyper-Wiener numbers (walk length 2) of the distance-edge (D_e), Wiener-edge (W_e), distance-path (D_p), Wiener-path (W_p), Cluj (CJu), and distance extended unsymmetric Cluj (D^2-CJ_u) matrixes	21, 22
χ	Randic's vertex-connectivity index	1
ID	Randic's identification number	23
R_1	bond additive index based on restricted random walks	24
IBC	information content based on center	25
AIBC	average information content based on center	25
H	Harary number	26
J	Balaban's Index	27
H^V	graph vertex complexity	28
H^D	graph distance complexity	28
$\overline{I^E_D}$	distance information index	29
X(**M**), C(**M**), DM^1	indices based on layer matrixes (**D**), (**LK**), (**LM3D**)	30−33
$U - Y$	Balaban's information-based index analogous to the J index	14
VEXX, VRXX	topological indices based on graph eigenvalues or eigenvectors (XX can be A1, A3, D1, and D3; see ref 34 for definitions).	34

connectivity index and the other topological indices studied. This feature has been claimed by Randic[35] as one of the desirable attributes for novel topological indices. To conduct this analysis, we will carry out a factor analysis by using the principal component method. The theoretical aspects of this statistical technique have been extensively exposed in the literature including many chemical applications,[36] for which we will not give details of this method here. However we recall some of the main conclusions that can be drawn from a factor analysis that will be of great usefulness in the present work:[37] (i) Variables with a high loading in the same factor are interrelated and will be the more so the higher the loadings. (ii) No correlation exists between variables having nonzero loadings only in different factors. (iii) Only variables with high loadings in different factors may be combined in a regression equation in order to eliminate collinearities.

A data set of 18 octane isomers will be used in order to do the factor analysis. The use of octanes as a very suitable data set for testing topological indices has been advocated by Randic and Trinajstic.[38,39] This selection is recommended due to the interrelatedness that exists between several thermodynamic properties of alkanes (the same properties which will be studied here) with different numbers of carbon atoms. These properties are hardly interrelated when octanes are used as a data set.[35] On the other hand, all topological indices are designed to have (gradual) increments with the increments in the molecular weight. By this way, if we do the present study by using a series of compounds having different molecular weights, we will find "false" interrelations between the indices by an overestimation of the size effects inherent to these descriptors. The same is also valid when the QSPR model is to be obtained. It is not difficult to find "good" linear correlations between TIs and physicochemical properties of alkanes in data sets with great size variability. In fact, the simple use of the number of vertices in the molecular graph produced regression coefficients greater than 0.97 for most of the physicochemical properties of C_2-C_9 alkanes studied by Needham et al.[40] However, when data sets of isomeric compounds are considered, typically correlations that have high correlation coefficients when molecules of different size were considered will no longer show such good linear correlation.

Table 2. Eigenvalues, Percentage of the Variance Explained, and Cumulative Percentage of the Variance Reproduced by the Factor Analysis for the Topological Indices

	eigenvalue	% total variance	cumul %
*f*1	27.726	74.94	74.94
*f*2	3.655	9.88	84.81
*f*3	2.087	5.64	90.45

Finally, we will analyze the quality of the QSPR models obtained to describe the boiling points (BP), molar volume (MV), molar refractions (MR), heats of vaporization (HV), critical temperatures (CT), critical pressures (CP), and surface tensions (ST) of the octane isomers.

RESULTS AND DISSCUSION

The results of the factor analysis are given in Table 2. The first factor explains 75% of the variance in the topological indices studied. The addition of the second factor increases to 85% the variance explained, and the addition of the third factor allows 90% of the index variance to be accounted for.

Factor loadings from the principal component analysis, after a Quartimax rotation of the factors, are shown in Table 3. The objective of the rotational procedure is to obtain a clear pattern of loadings, that is, factors that are somehow clearly marked by high loadings for some variables and low loadings for others. The Quartimax rotation is directed to maximize the variances of (raw) factor loadings across factors for each variable. This strategy makes the structure of factors pattern as simple as possible, permitting a clearer interpretation of the factors without loss of orthogonality between them.

With only one exception of the edge-connectivity index, all the first- and second-generation TI's are strongly loaded in factor 1. Most of the third-generation TI's are also loaded in this factor, with the exceptions of Balaban's indices based

Table 3. Factor Loadings for the Topological Indices after a Quartimax Rotation of the Factors

index	*f*1	*f*2	*f*3	index	*f*1	*f*2	*f*3
ϵ	0.312	***0.858***	−0.077	*J*	***−0.970***	−0.073	0.025
W	***0.971***	0.083	0.009	H^V	***0.798***	0.015	−0.101
Z	***0.787***	−0.525	0.135	H^D	***−0.804***	−0.044	0.094
$\tilde{W}$	***0.950***	0.142	0.027	$\overline{I^E_D}$	***0.904***	−0.004	−0.096
MTI	***0.965***	0.128	−0.002	*X*(**D**)	***−0.960***	0.130	−0.040
WW	***0.962***	0.119	0.016	*X*(**LK**)	***−0.937***	0.218	−0.067
$^2W_{De}$	***0.967***	0.083	0.021	DM^1	***−0.933***	0.188	−0.001
$^2W_{We}$	***0.932***	0.282	0.011	*C*(**LM3D**)	***−0.946***	0.111	−0.004
$^2W_{Cju}$	***0.964***	0.101	0.019	*U*	***0.910***	0.311	−0.117
$^2WW_{Dp}$	***0.940***	0.127	0.046	*V*	***−0.966***	−0.069	0.037
$^2WW_{Wp}$	***0.940***	0.189	0.039	*X*	***−0.972***	−0.049	0.008
D^2-CJ_u	***0.933***	0.160	0.042	*Y*	***−0.943***	−0.058	0.085
χ	***0.839***	−0.478	0.078	VEA1	0.484	0.503	***0.677***
ID	***0.932***	−0.272	0.030	VEA3	0.481	0.502	***0.680***
R_1	***−0.920***	0.312	0.036	VED1	−0.149	***0.877***	0.184
IBC	***0.922***	−0.229	0.036	VED3	0.024	0.341	***−0.846***
AIBC	***0.891***	−0.248	−0.037	VRA1	0.632	−0.236	***−0.705***
H	***−0.968***	0.079	−0.026	VRD1	***−0.955***	−0.173	0.023

Table 4. Correlations among First-generation TI's and the Edge-Connectivity Index

index	ϵ	W	WW	$^2W_{We}$	^2D-CJu	Z	$\tilde{W}$	MTI
ϵ	1.0000							
W	0.3800	1.0000						
WW	0.4018	***0.9972***	1.0000					
$^2W_{We}$	0.5389	***0.9771***	***0.9833***	1.0000				
^2D-CJu	0.4146	***0.9778***	***0.9905***	***0.9752***	1.0000			
Z	−0.2529	***0.7607***	***0.7374***	0.6137	***0.7022***	1.0000		
$\tilde{W}$	0.4115	***0.9903***	***0.9979***	***0.9824***	***0.9973***	***0.7202***	1.0000	
MTI	0.4167	***0.9985***	***0.9988***	***0.9859***	***0.9836***	***0.7281***	***0.9941***	1.0000

Table 5. Correlations among Some Second-Generation TI's and the Edge-Connectivity Index

index	ϵ	χ	ID	R_1	*J*	AIBC	*H*	$\overline{I^E_D}$
ϵ	1.0000							
χ	−0.1723	1.0000						
ID	0.0524	***0.9690***	1.0000					
R_1	−0.0476	***−0.9586***	***−0.9857***	1.0000				
J	−0.3855	***−0.8161***	***−0.9255***	***0.9136***	1.0000			
AIBC	0.0628	***0.8830***	***0.9159***	***−0.9354***	***−0.8803***	1.0000		
H	−0.2480	***−0.9014***	***−0.9735***	***0.9596***	***0.9848***	***−0.9087***	1.0000	
$\overline{I^E_D}$	0.3471	***0.7911***	***0.8813***	***−0.8906***	***−0.9247***	***0.8047***	***−0.9189***	1.0000

on eigenvalues or eigenvectors of the adjacency and distance matrices.[34] The second factor is almost exclusively an edge-connectivity and VED1 dimension. The VED1 index is the sum of the elements of the eigenvector derived from the main eigenvalue of the distance matrix of the graph.[34] There is no apparent relationship between these two topological indices. The third factor appears to be most significant for the eigenvalue/eigenvector third-generation indices of Balaban.[34]

As previously stated, the indices with a high loading in the same factor are interrelated, while no correlation exists between indices having nonzero loadings only in different factors. Consequently, it is clear that the edge-connectivity index is orthogonal to all the first- and second-generation TI's as well as to most of the third-generation ones. The only index that appears to be related to the edge-connectivity is the VED1 index. However, the correlation coefficient between these two indices is only 0.6489, which does not overpass the cutoff value of 0.70 commonly used to consider collinearity between descriptors in QSPR/QSAR. In closing, we can say that the edge-connectivity index contains structural information not contained in any other topological index of the first, second, or third generation. In Tables 4−6 we show the correlations among the edge-connectivity index and some of the first-, second-, and third-generation topological indices studied here.

From these tables we can see that most of the topological indices are interrelated, some of them having correlations coefficients greater than 0.99. These interrelations are not only observed between indices of the same generation but also between indices of different generations, such as χ and *Z* ($R = 0.9875$), only mentioned as one example. With the exception of the ϵ index, the other indices that are not interrelated to the rest of the topological indices studied here are the Balaban's indices based on eigenvalues/eigenvectors.[34]

However, the decisive criterion of quality for any molecular descriptor is its ability to describe structure-related properties of molecules. With this objective we developed the QSPR models to describe the seven physicochemical properties of octane isomers previously mentioned. The correlation coefficients for these linear regression models are given in Table 7. All the correlation coefficients greater than or equal to 0.7 are italicized, and the best correlations for

Table 6. Correlations among Some Third-Generation TI's and the Edge-Connectivity Index

index	ϵ	X(**LK**)	DM_1	U	V	VEA1	VED1	VRD1
ϵ	1.0000							
X(**LK**)	−0.1140	1.0000						
DM_1	−0.1413	***0.9893***	1.0000					
U	0.5751	***−0.8105***	***−0.8358***	1.0000				
V	−0.3858	***0.9446***	***0.9575***	***−0.9558***	1.0000			
VEA1	0.5311	−0.4185	−0.3924	0.5417	−0.5063	1.0000		
VED1	0.6489	0.3076	0.2990	0.1057	0.0942	0.5104	1.0000	
VRD1	−0.4767	***0.9050***	***0.9196***	***−0.9794***	***0.9929***	−0.5628	0.0087	1.0000

Table 7. Correlation among First-, Second-, and Third-Generation TI's and the Physicochemical Properties of Octanes[a]

index	BP	MV	MR	HV	TC	PC	ST
ϵ	−0.4240	***0.8885*****	***0.9004*****	−0.0359	***−0.7840*****	***−0.8011*****	***−0.7318*****
χ	***0.8212***	0.0132	0.0852	***0.9545*****	0.4991	−0.1412	0.5590
Z	***0.8884*****	−0.1123	−0.0361	***0.9581*****	0.5659	−0.0994	0.6636
ID	***0.7305***	0.2874	0.3639	***0.9210***	0.3117	−0.3472	0.3298
R_1	−0.6993	−0.3033	−0.3802	***−0.9012***	−0.3157	0.3106	−0.2956
IBC	***0.7313***	0.2862	0.3817	***0.9028***	0.3553	−0.2899	0.3220
AIBC	0.6966	0.2750	0.3820	***0.8751***	0.3925	−0.2087	0.2999
H	−0.5982	−0.5017	−0.5708	***−0.8321***	−0.1554	0.4623	−0.1046
X(**LK**)	−0.6752	−0.3481	−0.4157	***−0.8932***	−0.2774	0.3481	−0.2567
X(**D**)	−0.6247	−0.4527	−0.5210	***−0.8548***	−0.1980	0.4223	−0.1547
DM^1	−0.5907	−0.4566	−0.5150	***−0.8357***	−0.2385	0.3281	−0.1362
C(**LM3D**)	−0.6297	−0.4204	−0.4905	***−0.8578***	−0.2023	0.4155	−0.1719
W	0.5388	0.5916	0.6692	***0.7729***	0.0419	−0.5792	0.0079
J	−0.4845	−0.6468	***−0.7128***	***−0.7310***	−0.0245	0.5522	0.0677
$\tilde{W}$	0.5399	0.5764	0.6621	***0.7644***	0.0234	−0.6084	0.0212
MTI	0.5086	0.6207	0.6981	***0.7461***	0.0038	−0.6070	−0.0294
WW	0.5364	0.5888	0.6706	***0.7663***	0.0271	−0.5994	0.0096
$^2W_{De}$	0.5601	0.5658	0.6477	***0.7875***	0.0530	−0.5828	0.0395
$^2W_{We}$	0.3963	***0.7055***	***0.7771***	0.6458	−0.1107	−0.6700	−0.1550
$^2W_{CJu}$	0.5493	0.5762	0.6583	***0.7776***	0.0405	−0.5915	0.0260
$^2WW_{Dp}$	0.5697	0.5360	0.6259	***0.7837***	0.0481	−0.5992	0.0676
$^2WW_{Wp}$	0.5118	0.5975	0.6818	***0.7388***	−0.0074	−0.6275	−0.0091
D^2-CJ_u	0.5469	0.5550	0.6448	***0.7635***	0.0260	−0.6107	0.0412
H^V	0.3415	0.5475	0.6109	0.5457	−0.0075	−0.4285	−0.1335
$\overline{H^D}$	−0.3955	−0.5289	−0.5702	−0.6105	0.0607	0.5518	0.0528
$\overline{I^E_D}$	0.4602	0.5675	0.6235	0.6982	−0.0002	−0.5379	−0.0414
U	0.2863	***0.7969***	***0.8577***	0.5421	−0.1750	−0.6580	−0.2969
V	−0.4665	−0.6626	***−0.7251***	***−0.7155***	−0.0207	0.5402	0.0917
X	−0.5099	−0.6193	−0.6864	***−0.7536***	−0.0480	0.5411	0.0321
Y	−0.4028	***−0.7160***	***−0.7687***	−0.6543	−0.0088	0.4957	0.1778
VEA1	0.1390	0.4598	0.4787	0.3253	−0.2320	−0.5498	−0.1390
VEA3	0.1391	0.4572	0.4759	0.3251	−0.2315	−0.5489	−0.1373
VED1	−0.4948	0.4864	0.4714	−0.3157	−0.5902	−0.4108	−0.5424
VED3	−0.3195	0.3542	0.3533	−0.2125	−0.3204	−0.1717	−0.4033
VRA1	0.3551	0.3199	0.3808	0.4520	0.1416	−0.2030	−0.0056
VRD1	−0.4046	***−0.7170***	***−0.7784***	−0.6576	0.0609	0.6027	0.1682

[a] Correlation coefficients greater than 0.70 are bold and italicized, and the best correlations for each property are double-asterisked.

each physicochemical property are double-asterisked.

As can be seen in this table, there are two physicochemical properties that are well-accounted for the topological indices with high loadings of factor 1. These properties are the boiling point and the heat of vaporization. Most of these *f*1-indices produce good linear relationships with the heat of vaporization, and only a small group of them correlate well with the boiling points. The molar volume and molar refraction are well-described for the edge-connectivity index as well as for the Diudea's $^2W_{We}$ index[22] and some of the Balaban's third-generation TI's.[14] The other physicochemical properties, that is, critical pressure and volume, and the surface tension, are well-described exclusively by the edge-connectivity index. The eigenvalue-/eigenvector-based indices do not produce any significant correlation with the physicochemical properties studied, with only the exception of the index VRD1 that has correlation coefficients greater than 0.70 for molar volume and molar refraction.

The Hosoya's Z index[17] produces the best single regression models for describing the boiling points and the heat of vaporization of octanes. The Randic's χ index[1] produces a regression coefficient for the heat of vaporization very similar to that obtained by the Z index. However, for the other five physicochemical properties the best models are obtained by using the edge-connectivity index. The differences in the correlation coefficients of the edge-connectivity index and the rest of the topological indices for these five properties are very significant. From the physicochemical point of view, the edge connectivity index appears to be more appropriated for describing molecular volume properties; typical examples are the MV and MR, but other properties such as TC, PC, and ST can also be considered in this category. However, typical molecular "surface" dependent properties, such as BP and HV, are not well-described by this index. Conse-

quently, the edge-connectivity index can be considered as a "molecular volume" descriptor complementing the rest of the family of topological indices that are mainly "molecular surface" descriptors.

CONCLUSIONS

After analysis of the results previously discussed, the conclusions of the present work are almost straightforward. They can be restated as follows:

(i) The edge-connectivity index is an independent topological index not linearly related to any of the first-, second-, or third-generation TI's studied here.

(ii) The edge-connectivity index contains structural information not accounted for by the other topological indices previously defined in the literature. This structural information is quite important for the QSPR studies of several physicochemical properties.

(iii) The edge-connectivity index produces the best single regression models for the study of molar volume, molar refraction, critical pressure, critical temperature, and surface tension of octane isomers. The best models for boiling points and heat of vaporization are obtained by using the Hosoya or Randic indices.

(iv) The edge-connectivity index may be combined in one regression equation with any other of the topological indices studied in order to describe a given property of octane isomers without risk of collinearities between the descriptors.

At this point there is no doubt that the edge-connectivity index is a very promising molecular descriptor to be use in QSPR/QSAR studies. In the forthcoming paper of this series (the following paper in this issue) we will study the inclusion of long-range contributions in the edge-connectivity index. We will show there how the inclusion of these contributions increases the quality of the QSPR models developed. The third work of this series will be devoted to the study of molecules containing heteroatoms in their structures.

REFERENCES AND NOTES

(1) Randic, M. On characterization of molecular branching. *J. Am. Chem. Soc.* **1975**, *97*, 6609−6615.

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

(3) Kier, L. B.; Hall, L. H. *Molecular Connectivity in Structure−Activity Analysis*; Research Studies Press: Letchworth, 1986.

(4) Randic, M. Comparative regression analysis. Regressions based on a single descriptor. *Croat. Chem. Acta* **1993**, *66*, 289−312.

(5) Estrada, E. Edge adjacency relationships and a novel topological index related to molecular volume. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 31−33.

(6) 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.

(7) Estrada, E.; Ramirez, A. Edge adjacency relationships and molecular topographic descriptors. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 837−843.

(8) Estrada, E.; Guevara, N.; Gutman, I. Extension of edge connectivity index. Relationships to line graph indices and QSPR applications *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 428−431.

(9) Casch, G. G. Correlation of physicochemical properties of alkylphenols with their graph-theoretical ϵ parameter. *Chemosphere* **1995**, *31*, 4307−4315.

(10) Lekishvili, T.; Lekishvili, G.; Aleksidze, N. Mathematical investigation of the cancerogenic activity of some hormones. *Bull. Georgian Acad. Sci.* **1997**, *155*, 441−443.

(11) Nikolic, S.; Trinajstic, N. Comparison between the vertex- and edge-connectivity indices for benzenoid hydrocarbons. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 42−46.

(12) Plavsic, D.; Trinajstic, N.; Amic, D.; Solskic, M. Comparison between the structure-boiling points relationships with different descriptors for condensed benzenoids. *New J. Chem.* **1998**, *22*, 1075−1078.

(13) Nikolic, S.; Trinajstic, N. Modeling the aqueous solubility of aliphatic alcohols. *SAR QSAR Environ. Res.* **1998**, *9*, 117−126.

(14) Balaban, A. T. Using real numbers as vertex invariants for third-generation topological indexes. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 23−28.

(15) Gutman, I.; Estrada, E. Topological indices based on the line graph of the molecular graph. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 541−543.

(16) Wiener, H. Structural determination of paraffin boiling points. *J. Am. Chem. Soc.* **1947**, 69, 17−20.

(17) Hosoya, H. Topological index. A newly proposed quantity characterizing topological nature of structural isomers of saturated hydrocarbons. *Bull. Chem. Soc. Jpn.* **1971**, *44*, 2332−2339.

(18) Tratch, S. S.; Stankevich, M. I.; Zefirov, N. S. Combinatorial models and algorithms in chemistry. The expanded Wiener number-A novel topological index. *J. Comput. Chem.* **1990**, *8*, 899−908.

(19) Schultz, H. P. Topological organic chemistry. 1. Graph theory and topological indices of alkanes. *J. Chem. Inf. Comput. Sci.* **1989**, *29*, 277−288.

(20) Randic, M. Novel molecular description for structure−property studies. *Chem. Phys. Lett.* **1993**, 211, 478−483.

(21) Diudea, M. V. Cluj matrix, **Cj**u: source of various graph descriptors. *Commun. Math. Comput. Chem.* (*MATCH*) **1997**, *35*, 169−183.

(22) Diudea, M. V. Walk numbers $^{e}W_{M}$: Wiener-type numbers of higher rank. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 535−540.

(23) Randic, M. On molecular identification numbers. *J. Chem. Inf. Comput. Sci.* **1984**, *24*, 164−175.

(24) Randic, M. Restricted random walks on graphs. *Theor. Chim. Acta* **1995**, *92*, 97−106.

(25) Balaban, A. T.; Bertelsen, S.; Basak, S. C. New centric indexes for acyclic molecules (trees) and substituents (rooted trees), and coding of rooted trees. *Commun. Math. Comput. Chem.* (*MATCH*) **1994**, *30*, 55−72.

(26) (a) Mihalic, Z.; Trinajstic, N. A graph-theoretical approach to structure−property relationships. *J. Chem. Educ.* **1992**, *69*, 701−712. (b) Plavsic, D.; Nikolic, S.; Trinajstic, N.; Mihalic, Z. On the Harary index for the characterization of chemical graphs. *J. Math. Chem.* **1993**, *12*, 235−250.

(27) Balaban, A. T. Highly discriminating distance-based topological index. *Chem. Phys. Lett.* **1982**, *89*, 399−404.

(28) Raychadhury, C.; Ray, S. K.; Ghosh, J. J.; Roy, A. B.; Basak, S. C. Discrimination of isomeric structures using information theoretic topological indices. *J. Comput. Chem.* **1984**, *5*, 581−488.

(29) Bonchev, D.; Trinajstic, N. Information theory, distance matrix and molecular branching. *J. Chem. Phys.* **1977**, *67*, 4517−4533.

(30) Balaban, A. T.; Ciubotariu, D.; Ivanciuc, O. Design of topological indices. Part 2. Distance measure connectivity indices. *Commun. Math. Comput. Chem.* (*MATCH*) **1990**, *25*, 41−70.

(31) Balaban, A. T. Diudea, M. V. Real number vertex invariants: Regressive distance sums and related topological indices. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 421−428.

(32) Diudea, M. V. Layer matrices in molecular graphs. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1064−1071.

(33) Diudea, M. V.; Horvath, D.; Graovac, A. Molecular topology. 15. 3D distance matrices and related topological indices. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 129−135.

(34) Balaban, A. T.; Ciubotariu, D.; Medeleanu, M. Topological indices and real number vertex invariants based on graph eigenvalues and eigenvectors. *J. Chem. Inf. Comput. Sci.* **1991**, *31*, 517−523.

(35) Randic, M. Generalized molecular descriptors. *J. Math. Chem.* **1991**, *7*, 155−168.

(36) Malinowski, E. R.; Howery, D. G. *Factor Analysis in Chemistry*; Wiley: New York, 1980.

(37) Franke, R. *Theoretical Drug Design Methods*; Elsevier: Amsterdam, 1984, pp 197−188.

(38) Randic, M.; Trinajstic, N. Viewpoint 4-Comparative structure-property studies: The connectivity basis. *J. Mol. Struct. (THEOCHEM)* **1993**, *284*, 209−221.

(39) Randic, M.; Trinajstic, N. In search for graph invariants of chemical interest. *J. Mol. Struct. (THEOCHEM)* **1993**, *300*, 551−572.

(40) Needham, D. E.; Wei, I.-C.; Seybold, P. G. Molecular modeling of the physical properties of alkanes. *J. Am. Chem. Soc.* **1988**, *110*, 4186−4194.

CI990030P