Internet Electronic Journal of
Molecular Design

June 2002, Volume 1, Number 6, Pages 319–331

Editor: Ovidiu Ivanciuc

Special issue dedicated to Professor Milan Randić on the occasion of the 70th birthday
Part 2

Guest Editor: Mircea V. Diudea

Optimum Structural Descriptors Derived from the Ivanciuc–Balaban Operator

Ovidiu Ivanciuc,[1] Teodora Ivanciuc,[2] Daniel Cabrol–Bass,[3] and Alexandru T. Balaban[4]

[1] Sealy Center for Structural Biology, Department of Human Biological Chemistry & Genetics, University of Texas Medical Branch, Galveston, Texas 77555–1157
[2] University "Politehnica" of Bucharest, Department of Organic Chemistry, Bucharest, Romania
[3] Aromes, Synthèses et Interactions, Equipe Chimiométrie et Modélisation, University of Nice–Sophia Antipolis, Parc Valrose, 06108 Nice Cedex 2, France
[4] Department of Marine Sciences, Texas A & M University at Galveston, Fort Crockett Campus, 5007 Avenue U, Galveston, TX 77551, USA

Received: April 15, 2002; Accepted: May 27, 2002; Published: June 30, 2002

Citation of the article:
O. Ivanciuc, T. Ivanciuc, D. Cabrol–Bass, and A. T. Balaban, Optimum Structural Descriptors Derived from the Ivanciuc–Balaban Operator, *Internet Electron. J. Mol. Des.* **2002**, *1*, 319–331, http://www.biochempress.com.

Inter*net* Electronic Journal of Molecular Design

BioChem Press
http://www.biochempress.com

Optimum Structural Descriptors Derived from the Ivanciuc–Balaban Operator[#]

Ovidiu Ivanciuc,[1,*] Teodora Ivanciuc,[2] Daniel Cabrol–Bass,[3] and Alexandru T. Balaban[4]

[1] Sealy Center for Structural Biology, Department of Human Biological Chemistry & Genetics, University of Texas Medical Branch, Galveston, Texas 77555–1157
[2] University "Politehnica" of Bucharest, Department of Organic Chemistry, Bucharest, Romania
[3] Aromes, Synthèses et Interactions, Equipe Chimiométrie et Modélisation, University of Nice–Sophia Antipolis, Parc Valrose, 06108 Nice Cedex 2, France
[4] Department of Marine Sciences, Texas A & M University at Galveston, Fort Crockett Campus, 5007 Avenue U, Galveston, TX 77551, USA

Received: April 15, 2002; Accepted: May 27, 2002; Published: June 30, 2002

Internet Electron. J. Mol. Des. **2002**, *1* (6), 319–331

Abstract

Topological indices represent an important class of structural descriptors, widely used in developing structure–property models, in drug design, virtual synthesis, similarity and diversity assessment. Because there is a considerable interest in deriving additional topological indices for QSPR and QSAR models, we introduce in this paper an extended Ivanciuc–Balaban operator. We consider that the exponent –1/2 used in the standard formula of this operator may not be the optimum for modeling some molecular properties, and we now define a new formula containing an adjustable exponent that is optimized for each investigated property. The Ivanciuc–Balaban structural descriptors were computed for 134 alkanes C_6–C_{10} from seven molecular matrices, namely adjacency **A**, distance **D**, reciprocal distance **RD**, distance–path $\mathbf{D_p}$, reciprocal distance–path $\mathbf{RD_p}$, path Szeged $\mathbf{Sz_p}$, and reciprocal path Szeged $\mathbf{RSz_p}$ matrix. Structural descriptors derived from the new operator were tested in QSPR studies for several alkane properties and we found that in most cases the optimum value of the exponent is different from –1/2.

Keywords. Structural descriptors; molecular graph operators; topological indices; molecular matrices; QSAR.

1 INTRODUCTION

Numerous attempts have been made in theoretical chemistry to develop structural descriptors that express in a numerical form the chemical structure. Such structural descriptors are widely used in modeling physical, chemical, or biological properties, in similarity and diversity assessment, database mining, and in the screening of virtual combinatorial libraries. In a first approximation the

---

[#] Dedicated to Professor Milan Randić on the occasion of the 70th birthday.
[*] Correspondence author; E–mail: ivanciuc@netscape.net.

chemical structure of a molecule can be represented as a molecular graph. Molecular graphs are non–directed chemical graphs that represent molecules, using different conventions. In the graph representation of molecules, their geometrical features such as bond lengths or bond angles, are not taken into account and the chemical bonding of atoms is regarded as being their most important characteristic. In molecular graphs vertices correspond to atoms and edges represent covalent bonds between atoms. Numerous reviews and research papers present the computation of structural descriptors from molecular graphs, and their use in drug design, toxicology, property prediction, and design of chemical libraries [1–14].

Using molecular graphs the chemical structure of a compound can be expressed by means of various graph matrices, polynomials, spectra, spectral moments, sequences counting distances, paths, and walks, or topological indices. A topological index (TI) is a numerical descriptor of the molecular structure based on certain topological features of the molecular graph. When compared with other classes of structural descriptors, such as geometric or quantum descriptors, topological indices have some important advantages because they can be easily computed from the molecular graph and they offer a simple way of measuring molecular branching, shape, size, and molecular similarity. Although it is possible to perform semiempirical quantum computations for molecules with more than one hundred atoms, the similarity and diversity of chemical databases is characterized almost exclusively by descriptors derived from the molecular graph.

Topological indices are mainly used as structural descriptors in quantitative structure–property relationship (QSPR) and quantitative structure–activity relationship (QSAR) models. Although there are several hundreds TIs defined in the literature, only a few have been extensively used in QSPR and QSAR studies. One of the most successful graph descriptors is the Randić connectivity index χ defined with the following equation [15]:

$$\chi = \sum_{e_{ij} \in E(G)} (\mathbf{deg}_i \times \mathbf{deg}_j)^{-1/2} \tag{1}$$

where $\mathbf{deg}_i$ and $\mathbf{deg}_j$ denote the degrees of the two adjacent vertices v_i and v_j that are incident with an edge e_{ij} in the molecular graph G, and the summation is extended over all edges from the edge set $E(G)$. Altenburg [16] suggested that the exponent –1/2 used in the standard formula for computing the connectivity index may not be the optimum for modeling some molecular properties and proposed a generalized χ index with the formula:

$$\chi(p) = \sum_{e_{ij} \in E(G)} (\mathbf{deg}_i \times \mathbf{deg}_j)^p \tag{2}$$

where the exponent p is optimized for each property investigated in a QSPR or QSAR study. This proposal had almost no echo in the scientific community that continued to use the basic equation (1) to compute the connectivity index. A study dedicated to the optimization of TIs by simple modifications of the original equations found that in formula (2) the exponent –1/3 gives the best correlation with the boiling temperatures of a set of 21 alkanes [17].

The connectivity index χ has a fairly high degeneracy, i.e. two or more non–isomorphic molecular graphs have the same value for the χ index. Estrada studied the degeneracy of connectivity indices in regular graphs and graphs representing alkanes and cycloalkanes [18]. He found that the exponent –1/2 is the main cause of the degeneracy of these indices and proposed to use the exponent –1/3 that reduces the level of degeneracy for the studied graphs. In a recent study, Altenburg's modification was used in an attempt to improve several QSPR models [19], demonstrating that the optimum value of the exponent p may be different from –1/2, and that it depends on the investigated property and on the set of chemical compounds. For example, for the boiling temperatures of 18 octanes, $p = -1.15$; for the boiling temperatures of 35 nonanes, $p = -1.3$; for the boiling temperatures of 20 cycloalkanes, $p = -0.5$; for the solubility of 54 aliphatic alcohols, $p = -0.4$; for the retention times of phenolic acids and anthocyanidins, $p = -1.6$.

The conclusion of the above studies is that the exponent –1/2 used in the original equation of the connectivity index may not be the optimum for modeling some molecular properties and equation (2), with a variable exponent, should be routinely employed in QSPR and QSAR models. This finding has relevance for all structural descriptors computed from vertex invariants with a similar formula. Such an index is the Balaban index J defined by the formula [20,21]:

$$J(G) = \frac{M}{\mu + 1} \sum_{e_{ij} \in E(G)} \left[\mathbf{DS}_i(G) \times \mathbf{DS}_j(G) \right]^{-1/2} \tag{3}$$

where $\mathbf{DS}_i$ and $\mathbf{DS}_j$ denote the distance sums of the two adjacent vertices v_i and v_j that are incident with an edge e_{ij} in the molecular graph G, M is the number of edges in the molecular graph, μ is the cyclomatic number (the number of cycles in the graph, $\mu = M - N + 1$, where N is the number of atoms the molecular graph), and the summation goes over all edges from the edge set $E(G)$. The distance sum is a vertex invariant computed from the distance matrix of the molecular graph. In a graph G with N vertices, the distance sum of the vertex v_i, $\mathbf{DS}_i$, is defined as the sum of topological distances between vertex v_i and every vertex in the molecular graph, *i.e.* the sum over row i or column i in the distance matrix $\mathbf{D}$ [20,21]:

$$\mathbf{DS}_i(G) = \sum_{j=1}^{N} \mathbf{D}_{ij}(G) = \sum_{j=1}^{N} \mathbf{D}_{ji}(G) \tag{4}$$

The factor $M/(\mu + 1)$ from formula (3) of index J ensures that the values of J do not necessarily increase with increasing number of vertices and cycles in the molecular graph. By design, the index J has a very low degeneracy, as was proved analytically [22] and tested with a computer program [23]. It was demonstrated that for many infinite graphs, J attains a finite limit, *e.g.* for an infinite linear alkane, J tends towards $\pi = 3.14159$ [24]. The definition of J was extended for molecular graphs containing heteroatoms and/or multiple bonds [25,26], and used successfully in structure–

property and structure–activity studies [27–30].

The Ivanciuc–Balaban operator [31] **IB** represents an extension of the index *J*, that can be computed with vertex invariants derived from any symmetric molecular matrix [32]. Because the **IB** operator is computed with a formula similar to equation (3), it is possible that the exponent −1/2 used in the standard formula may not be the optimum for modeling some molecular properties. In this paper we define an extended Ivanciuc–Balaban operator with a formula that contains an adjustable exponent to be optimized for each investigated property.

2 THE EXTENDED IVANCIUC–BALABAN OPERATOR

For any graph matrix **M** we introduce here the vertex sum operator, representing a generalization of the distance sum. The vertex sum operator for the vertex v_i in a graph *G*, $\mathbf{VS}(\mathbf{M},G)_i$, is defined as the sum of the elements in the column *i*, or row *i*, of the molecular matrix **M** [8,9,31,32]:

$$\mathbf{VS}(\mathbf{M},G)_i = \sum_{j=1}^{N} \mathbf{M}_{ij}(G) = \sum_{j=1}^{N} \mathbf{M}_{ji}(G) \tag{5}$$

The vertex sum operator was used to define new vertex and graph descriptors. If **M** is the adjacency matrix, the operator **VS** is identical with the degree vector **Deg**, if **M** is the distance matrix, the operator is identical with the distance sum **DS**, while if **M** is the reciprocal distance matrix **RD** [33], this operator gives the reciprocal distance sum **RDS** [34].

Using the vertex sum local invariant, the Ivanciuc–Balaban operator of a graph *G*, **IB**(**M**) = **IB**(**M**,*G*) was defined by analogy with the *J* index [31,32]:

$$\mathbf{IB}(\mathbf{M},G) = \frac{M}{\mu+1} \sum_{e_{ij}\in E(G)} \left[\mathbf{VS}_i(\mathbf{M},G)\times \mathbf{VS}_j(\mathbf{M},G)\right]^{-1/2} \tag{6}$$

where the summation goes over all edges from the edge set *E*(*G*). The **IB** operator gives the *J* index when **M** is the distance matrix, **VS** represents the distance sum **DS**, and *G* is the molecular graph of a hydrocarbon. The Ivanciuc–Balaban operator can be applied to any symmetric molecular matrix [32], such as the adjacency **A**, distance **D**, reciprocal distance **RD** [33–37], resistance distance **Ω** [38], path Szeged $\mathbf{Sz_p}$ [39–42], detour Δ [43], distance–valency **Dval** [44,45], distance–delta $\mathbf{D}_\Delta$ [46], or distance–path $\mathbf{D_p}$ [46] matrices.

In this paper we introduce the extended Ivanciuc–Balaban operator which contains an adjustable exponent that is optimized for each investigated property:

$$\mathbf{IB}(\mathbf{M},p,G) = \frac{M}{\mu+1} \sum_{e_{ij}\in E(G)} \left[\mathbf{VS}_i(\mathbf{M},G)\times \mathbf{VS}_j(\mathbf{M},G)\right]^{p} \tag{7}$$

In the following sections we present the influence of the parameter *p* on the correlational abilities of the **IB** indices in several QSPR models involving six physical properties of alkanes.

3 MATERIALS AND METHODS

Chemical Data. The QSPR models were developed for a data set consisting of 134 alkanes between C_6 and C_{10}, for the following six physical properties [47]: t_b, boiling temperature at normal pressure (°C); C_p, molar heat capacity at 300 K (J K^{-1} mol^{-1}); $\Delta_f G°_{300}$ (g), standard Gibbs energy of formation in the gas phase at 300 K (kJ mol^{-1}); $\Delta_{vap}H_{300}$, vaporization enthalpy at 300 K (kJ mol^{-1}); n_D^{25}, refractive index at 25 °C; ρ, density at 25 °C (kg m^{-3}). As it is known, there are 142 constitutional isomers for these alkanes, but data for all six properties are missing for the following eight of them: *n*–hexane, *n*–nonane, *n*–decane, 2–methylnonane, 3–methylnonane, 4–methylnonane, 5–methylnonane, 3–ethyl–2,4–dimethylhexane. The value of the refractive index of 2,2,3,3–tetramethylbutane is missing, while the reported density of this compound, 821.70 kg m^{-3} [47], is too high when compared with the density of similar alkanes and it was not considered in the computation of the density QSPR models. This group of 134 alkanes together with their six physical properties forms a database used as a standard test for structural descriptors and QSPR models [48–55].

Structural Descriptors. All Ivanciuc–Balaban structural descriptors **IB**(**M**,*p*) were computed with the formula (7) for the range of the exponent *p* from –2 to 2 with a step equal to 0.1. The vertex sum local invariants **VS**(**M**) and the **IB** descriptors were computed from seven molecular matrices: adjacency **A**, distance **D**, reciprocal distance **RD**, distance–path $\mathbf{D_p}$, reciprocal distance–path $\mathbf{RD_p}$, path Szeged $\mathbf{Sz_p}$, and reciprocal path Szeged $\mathbf{RSz_p}$. Because the Ivanciuc–Balaban descriptors encode mainly the molecular shape and branching and not the size, the number of carbon atoms *N* was added as a molecular size descriptor in some equations.

QSPR Model. The QSPR models were developed using mono– and bi–parametric regression equations. All biparametric correlations contain *N* and an **IB** index as variables. During the optimization of the exponent *p* the statistical quality of the QSPR models was monitored by computing the correlation coefficient *r*, standard deviation *s*, and Fisher test F.

4 RESULTS AND DISCUSSION

In Table 1 we present the statistical indices for selected monoparametric QSPR models obtained for the six alkane properties. For each property and molecular matrix **M** we present in Table 1 the results obtained with the Ivanciuc–Balaban indices **IB**(**M**,*p*) for the optimum *p* value and for *p* equal to –0.5. For comparison, we give also the statistics of the models that use as structural descriptor the number of carbon atoms *N*. We have also investigated the influence of the variable exponent *p* in biparametric QSPR models that contain as structural descriptors the number of carbon atoms N and an index **IB**(**M**,p). For each property and molecular matrix **M** we present in the right–hand half of Table 1 the results obtained in biparametric QSPR models for the optimum p value and for p equal to –0.5.

(1) Normal boiling temperature. Although we have investigated a large number of **IB** descriptors, the best correlation is obtained for the number of carbon atoms N:

$$t_b = -78.906(\pm 2.698) + 23.859(\pm 0.816)N$$
$$r = 0.973 \qquad s = 6.19 \qquad F = 2315$$

Very close statistical parameters has the equation that uses **IB**(**A**,–0.5):

$$t_b = 28.543(\pm 0.996) + 3.212(\pm 0.112)\mathbf{IB}(\mathbf{A},-0.5)$$
$$r = 0.972 \qquad s = 6.31 \qquad F = 2222$$

We have to mention that for alkanes, **IB**(**A**,–0.5) is identical with the Randić connectivity index χ. The results obtained indicate that the **IB** descriptors do not model satisfactorily the alkane boiling temperature in monoparametric equations. This results does not rule out the use of these descriptors in multiparametric models, together with other structural descriptors.

Although in the best monoparametric equation the **IB** index has p = –0.5, for biparametric models involving N and an **IB** index, the results from Table 1 indicate that **IB**(**A**,–2.0) gives the best results:

$$t_b = -63.987(\pm 4.120) + 20.068(\pm 1.292)N + 3.170(\pm 0.204)\mathbf{IB}(\mathbf{A},-2.0)$$
$$r = 0.983 \qquad s = 4.94 \qquad F = 1851$$

The decrease of the standard deviation from 6.19 to 4.94 is significant, and shows that a combination of a size descriptor, N, and a shape descriptor, **IB**, can give effective correlations for the alkane boiling temperature.

(2) Molar heat capacity. The result obtained in modeling the alkane molar heat capacities show again that a simple descriptor like N outperforms in monoparametric equations all **IB** indices tested here. From Table 1 it is clear that the best descriptor is N:

$$C_p = 3.689(\pm 0.086) + 23.161(\pm 0.540)N$$
$$r = 0.987 \qquad s = 4.10 \qquad F = 4978$$

The second place is occupied by the Ivanciuc–Balaban index **IB**(**A**,–0.1):

$$C_p = 110.965(\pm 2.903) + 1.779(\pm 0.047)\mathbf{IB}(\mathbf{A},-0.1)$$
$$r = 0.984 \qquad s = 4.58 \qquad F = 3953$$

The biparametric equations show barely a minor improvement in the standard deviation for **IB**(**A**,–2.0) when compared to the equation with N only. From our limited investigation, it appears that the **IB** descriptors are not particularly good descriptors for the alkane molar heat capacities which, being additive properties, depend of composition rather than on constitution.

Table 1. Statistical indices of the monoparametric and biparametric QSPR models for the six alkane properties. The biparametric QSPR models have as structural descriptors the number of carbon atoms *N* and an index **IB**(**M**,*p*). For each property and matrix **M** we present the results obtained with the indices **IB**(**M**,*p*) for the optimum *p* value and for *p* equal to –0.5; for comparison, we give also the statistics of the monoparametric models with *N* as structural descriptor.

Descriptor	*r*	*s*	F	Descriptor	*r*	*s*	F
Monoparametric correlations				Biparametric correlations			
(1) boiling temperature at normal pressure, t_b (°C)							
N	0.973	6.19	2315				
IB(**A**,–0.5)	0.972	6.31	2222	**IB**(**A**,–2.0)	0.983	4.94	1851
IB(**A**,–0.5)	0.972	6.31	2222	**IB**(**A**,–0.5)	0.976	5.84	1310
IB(**D**,–0.1)	0.964	7.07	1742	**IB**(**D**,–2.0)	0.975	5.97	1251
IB(**D**,–0.5)	0.507	22.96	46	**IB**(**D**,–0.5)	0.973	6.21	1149
IB(**RD**,–0.1)	0.966	6.86	1857	**IB**(**RD**,–2.0)	0.975	5.95	1260
IB(**RD**,–0.5)	0.942	8.93	1042	**IB**(**RD**,–0.5)	0.973	6.18	1163
IB($\mathbf{D_p}$,–0.1)	0.952	8.12	1287	**IB**($\mathbf{D_p}$,–2.0)	0.975	5.97	1248
IB($\mathbf{D_p}$,–0.5)	0.002	26.64	<1	**IB**($\mathbf{D_p}$,–0.5)	0.973	6.21	1151
IB($\mathbf{RD_p}$,–0.1)	0.966	6.86	1857	**IB**($\mathbf{RD_p}$,–2.0)	0.976	5.81	1323
IB($\mathbf{RD_p}$,–0.5)	0.939	9.19	977	**IB**($\mathbf{RD_p}$,–0.5)	0.973	6.14	1178
IB($\mathbf{Sz_p}$,–0.1)	0.969	6.54	2057	**IB**($\mathbf{Sz_p}$,0.7)	0.977	5.71	1369
IB($\mathbf{Sz_p}$,–0.5)	0.662	19.98	103	**IB**($\mathbf{Sz_p}$,–0.5)	0.973	6.17	1163
IB($\mathbf{RSz_p}$,–0.1)	0.955	7.87	1380	**IB**($\mathbf{RSz_p}$,0.1)	0.973	6.12	1185
IB($\mathbf{RSz_p}$,–0.5)	0.800	15.97	235	**IB**($\mathbf{RSz_p}$,–0.5)	0.973	6.21	1151
(2) molar heat capacity at 300 K, C_p ($J\ K^{-1}\ mol^{-1}$)							
N	0.987	4.10	4978				
IB(**A**,–0.1)	0.984	4.58	3953	**IB**(**A**,–2.0)	0.987	4.04	2567
IB(**A**,–0.5)	0.976	5.59	2613	**IB**(**A**,–0.5)	0.987	4.07	2519
IB(**D**,–0.1)	0.980	5.10	3165	**IB**(**D**,–2.0)	0.987	4.11	2478
IB(**D**,–0.5)	0.507	21.96	46	**IB**(**D**,–0.5)	0.987	4.11	2472
IB(**RD**,–0.1)	0.984	4.59	3933	**IB**(**RD**,2.0)	0.987	4.09	2502
IB(**RD**,–0.5)	0.953	7.74	1301	**IB**(**RD**,–0.5)	0.987	4.10	2488
IB($\mathbf{D_p}$,0.1)	0.970	6.16	2131	**IB**($\mathbf{D_p}$,–2.0)	0.987	4.10	2479
IB($\mathbf{D_p}$,–0.5)	0.015	25.48	<1	**IB**($\mathbf{D_p}$,–0.5)	0.987	4.11	2471
IB($\mathbf{RD_p}$,–0.1)	0.983	4.64	3850	**IB**($\mathbf{RD_p}$,–2.0)	0.987	4.08	2515
IB($\mathbf{RD_p}$,–0.5)	0.947	8.20	1144	**IB**($\mathbf{RD_p}$,–0.5)	0.987	4.09	2501
IB($\mathbf{Sz_p}$,–0.1)	0.978	5.30	2918	**IB**($\mathbf{Sz_p}$,1.0)	0.987	4.10	2480
IB($\mathbf{Sz_p}$,–0.5)	0.690	18.45	120	**IB**($\mathbf{Sz_p}$,–0.5)	0.987	4.11	2470
IB($\mathbf{RSz_p}$,–0.1)	0.973	5.86	2363	**IB**($\mathbf{RSz_p}$,2.0)	0.987	4.06	2532
IB($\mathbf{RSz_p}$,–0.5)	0.813	14.84	258	**IB**($\mathbf{RSz_p}$,–0.5)	0.987	4.10	2482
(3) standard Gibbs energy of formation in the gas phase at 300 K, $\Delta_f G°_{300}$ (g) ($kJ\ mol^{-1}$)							
N	0.801	8.85	236				
IB(**A**,1.5)	0.932	5.36	872	**IB**(**A**,2.0)	0.939	5.10	489
IB(**A**,–0.5)	0.758	9.63	179	**IB**(**A**,–0.5)	0.819	8.51	134
IB(**D**,–0.3)	0.925	5.62	780	**IB**(**D**,–0.2)	0.933	5.34	440
IB(**D**,–0.5)	0.813	8.60	258	**IB**(**D**,–0.5)	0.927	5.58	398
IB(**RD**,2.0)	0.928	5.51	819	**IB**(**RD**,2.0)	0.928	5.53	406
IB(**RD**,–0.5)	0.657	11.14	100	**IB**(**RD**,–0.5)	0.901	6.43	283
IB($\mathbf{D_p}$,–0.2)	0.936	5.19	938	**IB**($\mathbf{D_p}$,–0.2)	0.939	5.10	488
IB($\mathbf{D_p}$,–0.5)	0.461	13.11	36	**IB**($\mathbf{D_p}$,–0.5)	0.929	5.50	411
IB($\mathbf{RD_p}$,1.9)	0.935	5.23	920	**IB**($\mathbf{RD_p}$,2.0)	0.941	5.02	507
IB($\mathbf{RD_p}$,–0.5)	0.650	11.23	97	**IB**($\mathbf{RD_p}$,–0.5)	0.887	6.86	241
IB($\mathbf{Sz_p}$,–0.2)	0.927	5.55	803	**IB**($\mathbf{Sz_p}$,–0.2)	0.942	4.99	514
IB($\mathbf{Sz_p}$,–0.5)	0.250	14.31	9	**IB**($\mathbf{Sz_p}$,–0.5)	0.910	6.15	316
IB($\mathbf{RSz_p}$,0.3)	0.906	6.25	607	**IB**($\mathbf{RSz_p}$,0.3)	0.907	6.26	303
IB($\mathbf{RSz_p}$,–0.5)	0.411	13.47	27	**IB**($\mathbf{RSz_p}$,–0.5)	0.905	6.30	297

Table 1. (Continued)

Descriptor	*r*	*s*	F	Descriptor	*r*	*s*	F
Monoparametric correlations				Biparametric correlations			
(4) vaporization enthalpy at 300 K, $\Delta_{vap}H_{300}$ (kJ mol^{-1})							
N	0.930	1.58	849				
IB(**A**,–0.8)	0.969	1.08	1999	**IB**(**A**,0.7)	0.975	0.96	1282
IB(**A**,–0.5)	0.963	1.17	1667	**IB**(**A**,–0.5)	0.968	1.09	966
IB(**D**,0.3)	0.952	1.32	1289	**IB**(**D**,–0.3)	0.970	1.05	1060
IB(**D**,–0.5)	0.246	4.19	8	**IB**(**D**,–0.5)	0.970	1.05	1052
IB(**RD**,–0.6)	0.981	0.84	3367	**IB**(**RD**,–2.0)	0.986	0.72	2304
IB(**RD**,–0.5)	0.979	0.88	3062	**IB**(**RD**,–0.5)	0.980	0.86	1586
IB(**D**$_p$,0.2)	0.956	1.27	1394	**IB**(**D**$_p$,–0.2)	0.968	1.09	980
IB(**D**$_p$,–0.5)	0.278	4.15	11	**IB**(**D**$_p$,–0.5)	0.968	1.09	968
IB(**RD**$_p$,–0.6)	0.982	0.81	3588	**IB**(**RD**$_p$,–2.0)	0.989	0.64	2933
IB(**RD**$_p$,–0.5)	0.981	0.83	3428	**IB**(**RD**$_p$,–0.5)	0.981	0.83	1719
IB(**Sz**$_p$,0.1)	0.934	1.55	897	**IB**(**Sz**$_p$,–0.4)	0.952	1.32	636
IB(**Sz**$_p$,–0.5)	0.795	2.62	227	**IB**(**Sz**$_p$,–0.5)	0.952	1.33	635
IB(**RSz**$_p$,–0.2)	0.974	0.98	2443	**IB**(**RSz**$_p$,0.3)	0.978	0.90	1450
IB(**RSz**$_p$,–0.5)	0.933	1.55	890	**IB**(**RSz**$_p$,–0.5)	0.977	0.91	1407
(5) refractive index at 25 °C, n_D^{25}							
N	0.875	0.0066	428				
IB(**A**,1.2)	0.944	0.0045	1080	**IB**(**A**,2.0)	0.960	0.0038	766
IB(**A**,–0.5)	0.838	0.0075	309	**IB**(**A**,–0.5)	0.886	0.0064	238
IB(**D**,–0.3)	0.950	0.0043	1225	**IB**(**D**,–0.4)	0.951	0.0043	614
IB(**D**,–0.5)	0.780	0.0086	204	**IB**(**D**,–0.5)	0.950	0.0043	601
IB(**RD**,1.3)	0.925	0.0052	776	**IB**(**RD**,–0.3)	0.937	0.0048	464
IB(**RD**,–0.5)	0.753	0.0090	171	**IB**(**RD**,–0.5)	0.935	0.0049	449
IB(**D**$_p$,–0.2)	0.951	0.0042	1233	**IB**(**D**$_p$,–0.3)	0.954	0.0041	654
IB(**D**$_p$,–0.5)	0.379	0.0127	22	**IB**(**D**$_p$,–0.5)	0.951	0.0043	613
IB(**RD**$_p$,1.1)	0.931	0.0050	853	**IB**(**RD**$_p$,2.0)	0.946	0.0045	550
IB(**RD**$_p$,–0.5)	0.745	0.0091	163	**IB**(**RD**$_p$,–0.5)	0.926	0.0052	392
IB(**Sz**$_p$,–0.2)	0.957	0.0039	1443	**IB**(**Sz**$_p$,–0.2)	0.959	0.0039	738
IB(**Sz**$_p$,–0.5)	0.366	0.0127	20	**IB**(**Sz**$_p$,–0.5)	0.939	0.0047	485
IB(**RSz**$_p$,0.2)	0.931	0.0050	856	**IB**(**RSz**$_p$,–0.3)	0.942	0.0046	507
IB(**RSz**$_p$,–0.5)	0.514	0.0117	47	**IB**(**RSz**$_p$,–0.5)	0.941	0.0046	503
(6) density at 25 °C, ρ (kg m^{-3})							
N	0.854	13.75	354				
IB(**A**,1.3)	0.938	9.21	951	**IB**(**A**,2.0)	0.953	8.06	642
IB(**A**,–0.5)	0.819	15.18	268	**IB**(**A**,–0.5)	0.864	13.37	192
IB(**D**,–0.3)	0.946	8.60	1111	**IB**(**D**,–0.4)	0.948	8.49	571
IB(**D**,–0.5)	0.801	15.84	235	**IB**(**D**,–0.5)	0.947	8.57	560
IB(**RD**,1.5)	0.913	10.79	657	**IB**(**RD**,–0.3)	0.927	9.96	398
IB(**RD**,–0.5)	0.726	18.19	146	**IB**(**RD**,–0.5)	0.924	10.17	378
IB(**D**$_p$,–0.2)	0.950	8.30	1201	**IB**(**D**$_p$,–0.3)	0.951	8.18	621
IB(**D**$_p$,–0.5)	0.418	24.04	28	**IB**(**D**$_p$,–0.5)	0.948	8.44	579
IB(**RD**$_p$,1.2)	0.921	10.30	734	**IB**(**RD**$_p$,2.0)	0.936	9.36	459
IB(**RD**$_p$,–0.5)	0.721	18.35	142	**IB**(**RD**$_p$,–0.5)	0.912	10.89	322
IB(**Sz**$_p$,–0.2)	0.956	7.76	1394	**IB**(**Sz**$_p$,–0.2)	0.961	7.30	795
IB(**Sz**$_p$,–0.5)	0.320	25.07	15	**IB**(**Sz**$_p$,–0.5)	0.937	9.30	466
IB(**RSz**$_p$,0.2)	0.917	10.54	695	**IB**(**RSz**$_p$,–0.3)	0.932	9.62	431
IB(**RSz**$_p$,–0.5)	0.483	23.18	40	**IB**(**RSz**$_p$,–0.5)	0.932	9.64	429

(3) Standard Gibbs energy of formation. In the case of the alkane standard Gibbs energy of formation, the **IB** indices give much better results than N:

$$\Delta_f G^\circ_{300} = -49.224(\pm 2.643) + 5.032(\pm 0.270)\mathbf{IB}(\mathbf{D_p}, -0.2)$$
$$r = 0.936 \qquad s = 5.19 \qquad F = 938$$

Interestingly, the results from Table 1 show for this index a marked improvement over those obtained with **IB**(**D**$_\mathbf{p}$,–0.5). The size parameter N gives worse results, indicating that the shape information contained in the **IB** indices has a great importance in modeling the alkane standard Gibbs energy of formation:

$$\Delta_f G^\circ_{300} = -69.896(\pm 7.485) + 10.896(\pm 1.167)N$$
$$r = 0.801 \qquad s = 8.85 \qquad F = 236$$

A slightly better model is obtained in the biparametric models, as can be seen also in Table 1. In this case, the **IB** index derived from the path Szeged matrix gives the best equation:

$$\Delta_f G^\circ_{300} = -66.940(\pm 8.176) - 6.234(\pm 0.761)N + 14.791(\pm 1.806)\mathbf{IB}(\mathbf{Sz_p}, -0.2)$$
$$r = 0.942 \qquad s = 4.99 \qquad F = 514$$

Overall, our results indicate that **IB** descriptors have a good modeling ability for the alkane standard Gibbs energy of formation. This finding deserves further investigations for other classes of chemical compounds.

(4) Vaporization enthalpy. Interesting results were obtained for vaporization enthalpy of alkanes, with best equation offered by **IB**(**RD**$_\mathbf{p}$,–0.6):

$$\Delta_{vap} H_{300} = 17.842(\pm 0.490) + 1.252(\pm 0.034)\mathbf{IB}(\mathbf{RD_p}, -0.6)$$
$$r = 0.982 \qquad s = 0.81 \qquad F = 3588$$

Again, the lower statistical results obtained with N show that the shape information contained in the **IB**(**RD**$_\mathbf{p}$) index has a great importance in modeling the alkane vaporization enthalpies:

$$\Delta_{vap} H_{300} = 4.195(\pm 0.237) + 3.700(\pm 0.209)N$$
$$r = 0.930 \qquad s = 1.58 \qquad F = 849$$

For biparametric models, the same reciprocal distance–path **RD**$_\mathbf{p}$ matrix gives the best model, but with $p = -2.0$:

$$\Delta_{vap} H_{300} = 2.683(\pm 0.137) + 3.246(\pm 0.166)N + 8.054(\pm 0.412)\mathbf{IB}(\mathbf{RD_p}, -2.0)$$
$$r = 0.989 \qquad s = 0.64 \qquad F = 2933$$

It appears that **IB** indices give good models for alkane vaporization enthalpies, especially with descriptors derived from the **RD**$_\mathbf{p}$ matrix.

(5) Refractive index. The number of carbon atoms N is not a good descriptor for the alkane refractive indexes showing that this property depends also on the shape of alkanes:

$$n_D^{25} = 1.30795(\pm 0.10396) + 0.01104(\pm 0.00088)N$$
$$r = 0.875 \qquad s = 0.0066 \qquad F = 428$$

The results from Table 1 indicate that **IB**(**Sz$_\mathbf{p}$**,–0.2) gives the best monoparametric model:

$$n_D^{25} = 1.30852(\pm 0.05666) + 0.00966(\pm 0.00042)\mathbf{IB}(\mathbf{Sz_p}, -0.2)$$
$$r = 0.957 \qquad s = 0.0039 \qquad F = 1443$$

The improvement over the corresponding index with $p = -0.5$ is significant, adding new evidence that the exponent –0.5 is not always the optimum value for the Ivanciuc–Balaban descriptors. With biparametric equations, a slight improvement is obtained by using **IB**(**A**,2.0):

$$n_D^{25} = 1.335428(\pm 0.133625) + 0.006162(\pm 0.000617)N + 0.000009(\pm 0.000001)\mathbf{IB}(\mathbf{A}, 2.0)$$
$$r = 0.960 \qquad s = 0.0038 \qquad F = 766$$

The results reported in Table 1 demonstrate that **IB** descriptors are capable of modeling with reasonable results the alkane refractive indexes.

(6) Density. A size descriptor is not able to model the alkane densities, as demonstrated by the low quality statistical parameters obtained with N:

$$\rho = 536.027(\pm 46.842) + 20.835(\pm 1.821)N$$
$$r = 0.854 \qquad s = 13.75 \qquad F = 354$$

A significant improvement is offered by the **IB** descriptor derived from the path Szeged matrix, **IB**(**Sz$_\mathbf{p}$**,–0.2), showing that the shape effect on the alkane densities is well reflected by this index:

$$\rho = 532.799(\pm 23.472) + 18.657(\pm 0.822)\mathbf{IB}(\mathbf{Sz_p}, -0.2)$$
$$r = 0.956 \qquad s = 7.76 \qquad F = 1394$$

A biparametric model that uses the same index **IB**(**Sz$_\mathbf{p}$**,–0.2) and N brings a small improvement:

$$\rho = 542.389(\pm 53.275) - 6.846(\pm 0.672)N + 23.764(\pm 2.334)\mathbf{IB}(\mathbf{Sz_p}, -0.2)$$
$$r = 0.961 \qquad s = 7.30 \qquad F = 795$$

These models developed with the **IB**(**Sz$_\mathbf{p}$**) indices demonstrate that this descriptor reflects in an adequate way the shape effect on alkane densities. The results from Tables 1 indicate that the use of a variable exponent in the computation of the **IB** descriptors greatly improves the correlation for this property.

5 CONCLUSIONS

In summary, in monoparametric correlations the best results were provided by N for t_b and C_p, by **IB**(**Sz$_\mathbf{p}$**,–0.2) for n_D^{25} and ρ, by **IB**(**D$_\mathbf{p}$**,–0.2) for $\Delta_f G°_{300}$ and by **IB**(**RD$_\mathbf{p}$**,–0.6) for $\Delta_{vap}H_{300}$. In biparametric correlations including N and a topological index, the best results were provided by **IB**(**A**,–2.0) for t_b and C_p, by **IB**(**Sz$_\mathbf{p}$**,–0.2) for $\Delta_f G°_{300}$ and ρ, by **IB**(**RD$_\mathbf{p}$**,–2.0) for $\Delta_{vap}H_{300}$ and by **IB**(**A**,2.0) for n_D^{25}. As one can observe, the molecular matrix that gives the best **IB** index is conserved (but with a different exponent) only for $\Delta_{vap}H_{300}$ and ρ.

Obviously, in biparametric regression equations, the intercorrelation coefficient between the size descriptor (which here was the number N of carbon atoms, but in other cases may be the molecular weight, etc.) and the shape (structural) descriptors $\mathbf{IB}(\mathbf{M},p)$ will vary with the value of the exponent p. This behavior can be exploited in generating quasi–orthogonal basis sets of structural descriptors for compound clustering, screening virtual libraries, and for measuring the molecular similarity or diversity. Whereas in QSPR and QSAR models the exponent p is modified to maximize the correlation coefficient, in devising sets of quasi–orthogonal structural descriptors the parameter p would be set to a value that offers the minimum intercorrelation coefficient between a pair of descriptors.

6 REFERENCES

[1] L. B. Kier and L. H. Hall, *Molecular Connectivity in Chemistry and Drug Research*, Academic Press, New York, 1976.
[2] L. B. Kier and L. H. Hall, *Molecular Connectivity in Structure–Activity Analysis*, Research Studies Press, Letchworth, 1986.
[3] N. Trinajstić, *Chemical Graph Theory*, 2nd ed., CRC Press, Boca Raton, 1992.
[4] A. T. Balaban (Ed.), *From Chemical Topology to Three–Dimensional Geometry*, Plenum, New York, 1997.
[5] M. V. Diudea (Ed.), *QSPR/QSAR Studies by Molecular Descriptors*, Nova Science, Huntington, N.Y., 2001.
[6] O. Ivanciuc and A. T. Balaban, Graph Theory in Chemistry; in: *The Encyclopedia of Computational Chemistry*, Eds. P. v. R. Schleyer, N. L. Allinger, T. Clark, J. Gasteiger, P. A. Kollman, H. F. Schaefer III, and P. R. Schreiner, John Wiley & Sons, Chichester, 1998, pp. 1169–1190.
[7] O. Ivanciuc and A. T. Balaban, The Graph Description of Chemical Structures; in: *Topological Indices and Related Descriptors in QSAR and QSPR*, Eds. J. Devillers and A. T. Balaban, Gordon and Breach Science Publishers, Amsterdam, 1999, pp. 59–167.
[8] O. Ivanciuc, T. Ivanciuc, and A. T. Balaban, Vertex– and Edge–Weighted Molecular Graphs and Derived Structural Descriptors; in: *Topological Indices and Related Descriptors in QSAR and QSPR*, Eds. J. Devillers and A. T. Balaban, Gordon and Breach Science Publishers, Amsterdam, 1999, pp. 169–220.
[9] O. Ivanciuc and T. Ivanciuc, Matrices and Structural Descriptors Computed from Molecular Graph Distances; in: *Topological Indices and Related Descriptors in QSAR and QSPR*, Eds. J. Devillers and A. T. Balaban, Gordon and Breach Science Publishers, Amsterdam, 1999, pp. 221–277.
[10] O. Ivanciuc, Building–Block Computation of the Ivanciuc–Balaban Indices for the Virtual Screening of Combinatorial Libraries, *Internet Electron. J. Mol. Des.* **2002**, *1*, 1–9, http://www.biochempress.com.
[11] I. Rios–Santamarina, R. García–Domenech, J. Cortijo, P. Santamaria, E. J. Morcillo, and J. Gálvez, Natural Compounds with Bronchodilator Activity Selected by Molecular Topology, *Internet Electron. J. Mol. Des.* **2002**, *1*, 70–79, http://www.biochempress.com.
[12] O. Ivanciuc, QSAR Comparative Study of Wiener Descriptors for Weighted Molecular Graphs, *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 1412–1422.
[13] D. J. G. Marino, P. J. Peruzzo, E. A. Castro, and A. A. Toropov, QSAR Carcinogenic Study of Methylated Polycyclic Aromatic Hydrocarbons Based on Topological Descriptors Derived from Distance Matrices and Correlation Weights of Local Graph Invariants, *Internet Electron. J. Mol. Des.* **2002**, *1*, 115–133, http://www.biochempress.com.
[14] A. A. Toropov and A. P. Toropova, QSAR Modeling of Mutagenicity Based on Graphs of Atomic Orbitals, *Internet Electron. J. Mol. Des.* **2002**, *1*, 108–114, http://www.biochempress.com.
[15] M. Randić, On Characterization of Molecular Branching, *J. Am. Chem. Soc.* **1975**, *97*, 6609–6615.
[16] K. Altenburg, Eine Bemerkung zu dem Randicschen “Molekularen Bindungs–Index (Molecular Connectivity Index)”, *Z. Phys. Chem. (Leipzig)* **1980**, *261*, 389–393.
[17] M. Randić, P. J. Hansen, and P. C. Jurs, Search for Useful Graph Theoretical Invariants of Molecular Structure, *J. Chem. Inf. Comput. Sci.* **1988**, *28*, 60–68.
[18] E. Estrada, Graph Theoretical Invariant of Randić Revisited, *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 1022–1025.
[19] D. Amić, D. Bešlo, B. Lucić, S. Nikolić, and N. Trinajstić, The Vertex–Connectivity Index Revisited, *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 819–822.
[20] A. T. Balaban, Highly Discriminating Distance–Based Topological Index, *Chem. Phys. Lett.* **1982**, *89*, 399–404.
[21] A. T. Balaban, Topological Indices Based on Topological Distances in Molecular Graphs, *Pure Appl. Chem.*

1983, *55*, 199–206.

[22] A. T. Balaban and L. V. Quintas, The Smallest Graphs, Trees, and 4–Trees with Degenerate Topological Index J, *MATCH* (*Commun. Math. Chem.*) **1983**, *14*, 213–233.

[23] A. T. Balaban and P. Filip, Computer Program for Topological Index J (Average Distance Sum Connectivity), *MATCH* (*Commun. Math. Chem.*) **1984**, *16*, 163–190.

[24] A. T. Balaban, N. Ionescu–Pallas, and T.–S. Balaban, Asymptotic Values of Topological Indices J and J' (Average Distance Sum Connectivities) for Infinite Acyclic and Cyclic Graphs, *MATCH* (*Commun. Math. Chem.*) **1985**, *17*, 121–146.

[25] A. T. Balaban, Chemical Graphs. Part 48. Topological Index J for Heteroatom–Containing Molecules Taking Into Account Periodicities of Element Properties, *MATCH* (*Commun. Math. Chem.*) **1986**, *21*, 115–122.

[26] A. T. Balaban and O. Ivanciuc, FORTRAN 77 Computer Program for Calculating the Topological Index J for Molecules Containing Heteroatoms; in: *MATH/CHEM/COMP 1988*, Ed. A. Graovac, Elsevier, Amsterdam, 1989, pp. 193–211.

[27] D. Bonchev, C. F. Mountain, W. A. Seitz, A. T. Balaban, Modeling the Anticancer Action of Some Retinoid Compounds by Making Use of the OASIS Method, *J. Med. Chem.* **1993**, *36*, 1562–1569.

[28] G. Grassy, B. Calas, A. Yasri, R. Lahana, J. Woo, S. Iyer, M. Kaczorek, R. Floc'h, and R. Buelow, Computer–Assisted Rational Design of Immunosuppressive Compounds, *Nature Biotechnol.* **1998**, *16*, 748–752.

[29] A. T. Balaban, L. B. Kier, and N. Joshi, Correlations between Chemical Structure and Normal Boiling Points of Acyclic Ethers, Peroxides, Acetals, and Their Sulfur Analogues, *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 237–244.

[30] A. T. Balaban, C. Catana, M. Dawson, I. Niculescu–Duvaz, Applications of Weighted Topological Index J for QSAR of Carcinogenesis Inhibitors (Retinoic Acid Derivatives), *Rev. Roum. Chim.* **1990**, *35*, 997–1003.

[31] O. Ivanciuc, T. Ivanciuc, and A. T. Balaban, Design of Topological Indices. Part 10. Parameters Based on Electronegativity and Covalent Radius for the Computation of Molecular Graph Descriptors for Heteroatom–Containing Molecules, *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 395–401.

[32] O. Ivanciuc, T. Ivanciuc, and M. V. Diudea, Molecular Graph Matrices and Derived Structural Descriptors, *SAR QSAR Environ. Res.* **1997**, *7*, 63–87.

[33] T. S. Balaban, P. A. Filip, and O. Ivanciuc, Computer Generation of Acyclic Graphs Based on Local Vertex Invariants and Topological Indices. Derived Canonical Labelling and Coding of Trees and Alkanes, *J. Math. Chem.* **1992**, *11*, 79–105.

[34] O. Ivanciuc, T.–S. Balaban, and A. T. Balaban, Design of Topological Indices. Part 4. Reciprocal Distance Matrix, Related Local Vertex Invariants and Topological Indices, *J. Math. Chem.* **1993**, *12*, 309–318.

[35] O. Ivanciuc, Design on Topological Indices. Part 1. Definition of a Vertex Topological Index in the Case of 4–Trees, *Rev. Roum. Chim.* **1989**, *34*, 1361–1368.

[36] D. Plavšić, S. Nikolić, N. Trinajstić, and Z. Mihalić, On the Harary Index for the Characterization of Chemical Graphs, *J. Math. Chem.* **1993**, *12*, 235–250.

[37] M. V. Diudea, O. Ivanciuc, S. Nikolić, and N. Trinajstić, Matrices of Reciprocal Distance, Polynomials and Derived Numbers, *MATCH* (*Commun. Math. Comput. Chem.*) **1997**, *35*, 41–64.

[38] D. J. Klein and M. Randić, Resistance Distance, *J. Math. Chem.* **1993**, *12*, 81–95.

[39] M. V. Diudea, O. Minailiuc, G. Katona, and I. Gutman, Szeged Matrices and Related Numbers, *MATCH* (*Commun. Math. Comput. Chem.*) **1997**, *35*, 129–143.

[40] M. V. Diudea, Cluj Matrix Invariants, *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 300–305.

[41] M. V. Diudea, B. Pârv, and M. I. Topan, Derived Szeged and Cluj Indices, *J. Serb. Chem. Soc.* **1997**, *62*, 267–276.

[42] O. Ivanciuc, Design of Topological Indices. Part 13. Structural Descriptors Computed from the Szeged Molecular Matrices, *Rev. Roum. Chim.* **2000**, *45*, 475–493.

[43] O. Ivanciuc and A. T. Balaban, Design of Topological Indices. Part 8. Path Matrices and Derived Molecular Graph Invariants, *MATCH* (*Commun. Math. Chem.*) **1994**, *30*, 141–152.

[44] O. Ivanciuc, Design of Topological Indices. Part 11. Distance–Valency Matrices and Derived Molecular Graph Descriptors, *Rev. Roum. Chim.* **1999**, *44*, 519–528.

[45] O. Ivanciuc, Design of Topological Indices. Part 14. Distance–Valency Matrices and Structural Descriptors for Vertex– and Edge–Weighted Molecular Graphs, *Rev. Roum. Chim.* **2000**, *45*, 587–596.

[46] M. V. Diudea, Wiener and Hyper–Wiener Numbers in a Single Matrix, *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 833–836.

[47] A. A. Gakh, E. G. Gakh, B. G. Sumpter, and D. W. Noid, Neural Network–Graph Theory Approach to the Prediction of the Physical Properties of Organic Compounds, *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 832–839.

[48] E. Estrada, O. Ivanciuc, I. Gutman, A. Gutierrez, and L. Rodríguez, Extended Wiener Indices. A New Set of Descriptors for Quantitative Structure–Property Studies, *New J. Chem.* **1998**, *22*, 819–822.

[49] O. Ivanciuc, The Neural Network MolNet Prediction of Alkane Enthalpies, *Anal. Chim. Acta* **1999**, *384*, 271–284.

[50] O. Ivanciuc, Artificial Neural Networks Applications. Part 12. The Prediction of Alkane Heat Capacities with the

MolNet Neural Network, *Rev. Roum. Chim*. **2000**, *45*, 391–403.

[51] O. Ivanciuc, T. Ivanciuc, D. Cabrol–Bass, and A. T. Balaban, Evaluation in Quantitative Structure–Property Relationship Models of Structural Descriptors Derived from Information–Theory Operators, *J. Chem. Inf. Comput. Sci*. **2000**, *40*, 631–643.

[52] A. T. Balaban, D. Mills, O. Ivanciuc, and S. C. Basak, Reverse Wiener Indices, *Croat. Chem. Acta* **2000**, *73*, 923–941.

[53] O. Ivanciuc, T. Ivanciuc, and A. T. Balaban, The Complementary Distance Matrix, a New Molecular Graph Metric, *ACH – Models Chem*. **2000**, *137*, 57–82.

[54] O. Ivanciuc, T. Ivanciuc, D. J. Klein, W. A. Seitz, and A. T. Balaban, Wiener Index Extension by Counting Even/Odd Graph Distances, *J. Chem. Inf. Comput. Sci*. **2001**, *41*, 536–549.

[55] O. Ivanciuc, T. Ivanciuc, and D. J. Klein, Quantitative Structure–Property Relationships Generated with Optimizable Even/Odd Wiener Polynomial Descriptors, *SAR QSAR Environ. Res*. **2001**, *12*, 1–16.