CROATICA CHEMICA ACTA
CCACAA **77** (1–2) 251–257 (2004)

ISSN-0011-1643
CCA-2923
Original Scientific Paper

Novel Graphical Matrix and Distance-based Molecular Descriptors*

Milan Randić,[a], Nabamita Basak,[b],*** and Dejan Plavšić[c],****

[a]*Natural Resources Research Institute, University of Minnesota Duluth, 5013 Miller Trunk Highway, Duluth, Minnesota 55811, USA*

[b]*Department of Chemistry, University of Minnesota Duluth, 10 University Drive, Duluth, Minnesota 55812, USA*

[c]*The Ruđer Bošković Institute, P.O.Box 180, HR-10002 Zagreb, Croatia*

RECEIVED APRIL 1, 2003; REVISED SEPTEMBER 8, 2003; ACCEPTED SEPTEMBER 11, 2003

Key words
molecular descriptor
»double« invariant
Wiener index
Wiener-Wiener number
graphical matrix
total steric energy
octanes

We draw attention to graphical matrices as a source of numerous structural invariants that could be used as predictor variables in QSPR and QSAR studies of molecules. In particular, we put forward a novel graphical matrix $\boldsymbol{G}$ associated with a molecule whose off-diagonal element $[G]_{ij}$ is the subgraph of the corresponding molecular graph obtained from it by deleting vertices i and j. Several molecular descriptors have been extracted from the $\boldsymbol{G}$ matrix and its numerical realization, the $\boldsymbol{G_w}$ matrix, based on the Wiener index. The usability of the extracted »double« invariants as predictor variables in QSPR studies of molecules has been tested on the total steric energies of octane isomers.

INTRODUCTION

In this contribution we will report on the construction of several novel molecular descriptors (topological indices). In view of the fact that there are hundreds of molecular descriptors available, why should anyone need additional topological indices? We believe that molecular descriptors will continue to emerge for the following reasons: (i) There is no universal molecular descriptor; (ii) There are still a number of physicochemical properties of a variety of compounds, alkanes included, for which there are no good regression models based on a single or a few molecular descriptors; (iii) It is a challenge to design relatively simple descriptors to replace the descriptors that give good regressions but are conceptually (or computationally) too involved; (iv) There is a need to find a set of structurally related descriptors that can form a satisfactory basis for the structure-property/activity space. A review of the current literature on structure-property/activity studies shows that many molecular properties require *ad hoc* combinations of several molecular descriptors, often combinations of descriptors being structurally unrelated. This not only makes comparative QSPR (quantitative structure-property relationship) and QSAR (quantitative structure-activity relationship) studies difficult to formulate, but also makes the interpretation of the corresponding regression models very complicated. All the above-mentioned facts stimu-

---

* Dedicated to Professor Nenad Trinajstić, one of the pioneers of chemical graph theory.

** Authors to whom correspondence should be addressed. (M.R., Permanent address: 3225 Kingman Rd. Ames, IA 50014, USA; E-mail: mrandic@msn.com and dplavsic@rudjer.irb.hr).

*** Undergraduate student.

late interest in the design of novel molecular descriptors that would improve the graph-theoretical characterization of molecular structure and make it possible to develop new and better QSPR and QSAR models. The search for novel descriptors has been undertaken in several directions.[1] One direction of investigations is the creation of novel matrices associated with a molecular graph and the extraction of their invariants that could be used as molecular descriptors. Here we take this approach and put forward a novel graphical matrix and several novel molecular descriptors based on it.

GRAPHICAL MATRICES

The notion of graphical matrices representing molecular structures is relatively new. A matrix associated with a (molecular) graph whose elements are subgraphs of the graph rather than numbers (as is the case with traditional graph matrices) is called a graphical matrix. In the first such matrix $\boldsymbol{R}$, introduced by Randić, Razinger and Plavšić,[2] the element $[R]_{ij}$ was defined as the subgraph made up of all the shortest paths joining vertices i and j. The subgraphs that appear as matrix elements may be paths and cyclic fragments, including the molecule as a whole. The virtue of a graphical matrix lies in the fact that it allows a multitude of numerical realizations. In order to arrive at a numerical form of a graphical matrix, we need to select a graph invariant and replace all the graphical elements by the corresponding numerical values of the selected invariant. Once we obtain the numerical matrix, we have a second opportunity to freely select an invariant of choice, this time an invariant of the numerical matrix. Clearly, graphical matrices have considerable potential for generating novel molecular descriptors (»double« invariants) in view of the fact that they call twice for the selection of an invariant to be considered in the construction of molecular descriptors.

THE NOVEL GRAPHICAL MATRIX

We introduce now the novel graphical matrix $\boldsymbol{G}$ associated with a molecular graph M whose off-diagonal element $[G]_{ij}$ is defined as the subgraph of M obtained from M by deleting vertices i and j. By definition, all the diagonal elements are an empty subgraph. The vertex set and the edge set of the empty (sub)graph, Ø, contain no elements.[3] Since it does not matter whether we first erase vertex i or vertex j, it has to be $[G]_{ij} = [G]_{ji}$. The construction of elements of $\boldsymbol{G}$ is illustrated with the graph representing 3-methylheptane in Figure 1. As one can see, the matrix entries may be disconnected fragments, including isolated vertices. It is important to note that, in contrast to the $\boldsymbol{R}$ matrix,[2] here all the subgraphs appearing as matrix off-diagonal elements have the same number of vertices. We may mention that subgraphs appearing in the matrix as off-diagonal entries can be used to verify the legitimacy of the set of Ulam subgraphs. Ulam subgraphs that appear in the famous problem of Ulam's reconstruction of graphs[4] are obtained from a given graph when one vertex is deleted at a time.

The transformation of $\boldsymbol{G}$ matrix into a numerical matrix is carried out by means of a graph invariant. By definition, all the elements in the resulting numerical matrix that correspond to the entries in $\boldsymbol{G}$ being the empty subgraph are equal to zero. The $\boldsymbol{G}$ matrix numerically transformed by means of the Wiener index, W,[5,6] is designated $\boldsymbol{G}_w$. The $\boldsymbol{G}_w$ matrix associated with the molecular graph of 3-methylheptane (Figure 1) is shown in

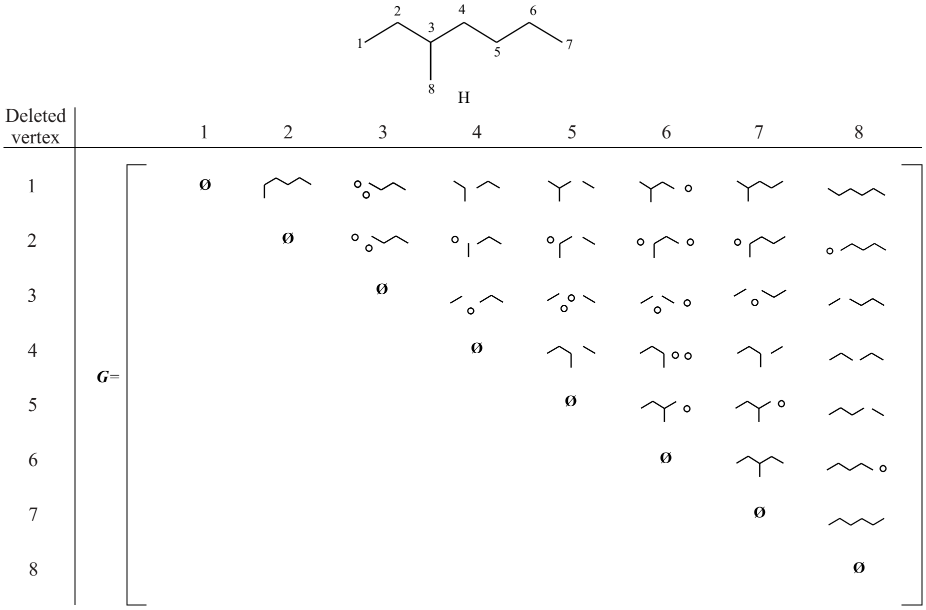

Figure 1. Construction of the **G** matrix for graph H representing the carbon skeleton of 3-methylheptane. The off-diagonal entry $[G]_{ij}$ is the subgraph of H obtained from H by deleting vertices i and j. As the **G** matrix is symmetrical, only its upper triangle is shown.

Table I. One should notice that we calculate W of a disconnected (sub)graph S with components S_i ($i = 1,\cdots, n$), that is:

$$\mathrm{S} = \bigcup_{i=1}^{n} \mathrm{S}_i \tag{1}$$

by means of the expression:

$$W = W(\mathrm{S}) = \sum_{i=1}^{n} W(\mathrm{S}_i) \tag{2}$$

where $W(\mathrm{S}_i)$ denotes the Wiener number of component S_i and the summation goes over all components of S. As one can see from Table I, the values of W for various entries of the $\boldsymbol{G}$ matrix may considerably vary in spite of the fact that all of them have six vertices. The entries corresponding to the terminal vertices have large values while the entries corresponding to the inner vertices have smaller numerical values. This situation may be contrasted with the one in the corresponding Wiener matrix, $\boldsymbol{W}$,[7, 8] where the opposite is the case. The off-diagonal entry $[W]_{ij}$ of the Wiener matrix of an acyclic molecular graph is defined as the product of the number of vertices on each side of the path connecting vertices i and j, and the entries on the main diagonal are by definition zero.

TABLE I. The upper triangle of the $\boldsymbol{G_W}$ matrix associated with the molecular graph of 3-methylheptane (Figure 1)

	1	2	3	4	5	6	7	8
1	0	35	10	8	10	18	32	35
2		0	10	5	5	10	20	20
3			0	5	2	2	5	11
4				0	11	10	11	8
5					0	18	18	11
6						0	31	20
7							0	35
8								0

NOVEL DISTANCE-BASED MOLECULAR DESCRIPTORS

We will now extract some novel structural invariants from the $\boldsymbol{G}$ matrix and its numerical form based on the Wiener index. An invariant of $\boldsymbol{G}_w$ matrix suggesting itself as a molecular descriptor is the sum of all the entries of the upper triangle of $\boldsymbol{G}_w$. This invariant is analogous to the Wiener number, which, as Hosoya has shown,[9] can be obtained by adding up all the entries above the main diagonal of the distance matrix.[10] Hence, the novel »double« invariant will shortly be referred to as the Wiener-Wiener number and designated $W(W)$. In Table II we list the values of the $W(W)$ number for all the 18 octane isomers.

TABLE II. Values of the $W(W)$ number, leading eigenvalues $\lambda(\boldsymbol{G_W})$ of the $\boldsymbol{G_W}$ matrices, leading eigenvalues $\lambda(\boldsymbol{W})$ of the $\boldsymbol{W}$ matrices, and the hyper-Wiener number WW for all the 18 octane isomers

Isomer[(a)]	$W(W)$	$\lambda(\boldsymbol{G_w})$	$\lambda(\boldsymbol{W})$	WW
n-octane	378	104.4359	57.1698	210
2-M-heptane	398	112.6326	52.6122	184
3-M-heptane	416	117.9955	48.4059	170
4-M-heptane	423	119.5012	46.6606	165
3-E-hexane	441	124.8102	42.2041	150
2,2-MM-hexane	438	126.8001	44.4713	149
2,3-MM-hexane	449	128.6530	42.0589	143
2,4-MM-hexane	443	127.9006	43.4185	147
2,5-MM-hexane	423	122.6675	47.7238	161
3,3-MM-hexane	464	132.7865	38.5332	131
3,4-MM-hexane	462	132.0751	39.2901	134
3,2-EM-pentane	469	133.6979	37.4277	129
3,3-EM-pentane	483	137.2475	34.1415	118
2,2,3-MMM-pentane	485	139.8462	34.9935	115
2,2,4-MMM-pentane	467	136.7979	39.1411	127
2,3,3-MMM-pentane	491	140.8804	33.4679	111
2,3,4-MMM-pentane	477	137.8224	37.0246	122
2,2,3,3-MMMM-butane	507	146.4616	30.3305	97

(a) Abbreviations M and E denote methyl and ethyl, respectively.

The $W(W)$ number is a global molecular descriptor that has no apparent edge contributions. In order to gain a deeper insight into the relationship between $W(W)$ and molecular structure, one needs to decompose $W(W)$ into edge contributions. Randić and Zupan[11] have recently shown that the partition of a molecular descriptor, defined as the sum of all the matrix elements of the upper triangle of a graph matrix in the case of connected acyclic graphs, can be accomplished by considering contributions from paths of different length. We will apply this approach to decompose the $W(W)$ number into edge contributions. The procedure is illustrated with the molecular graph of 3-methylheptane in Figure 2. The numbers assigned to each edge in the diagrams in the left column of Figure 2 represent the contributions arising formally from paths of different length and whose weights are given by the corresponding entries in the $\boldsymbol{G}_w$ matrix (Table I). For example, consider the diagram in the second row that shows the contributions of paths of length two. The edge 1–2 is involved only in a single path of length two: 1–2–3 and has been assigned a weight of 10, being the value of the matrix element $[G_w]_{13}$. The edge 2–3 is involved in three paths of length two: 1–2–3, 2–3–4, and 2–3–8, which have weights 10, 5, and 20, respectively, corresponding to matrix elements $[G_w]_{13}$, $[G_w]_{24}$ and $[G_w]_{28}$, respectively, totalling 35. Other contributions shown on the left side of Figure 2 are similarly derived. In order to obtain the partitioning of $W(W)$, we have to divide the

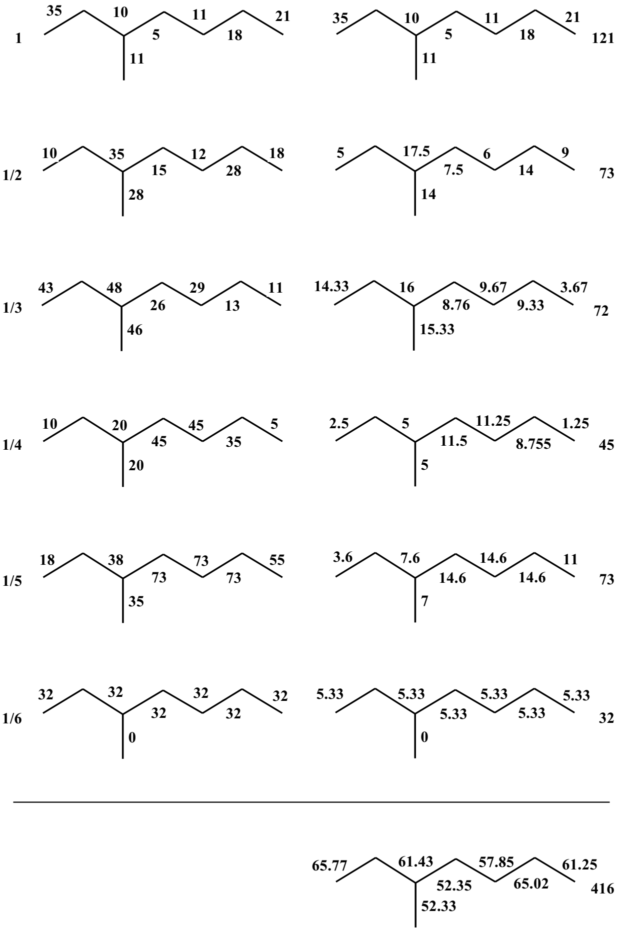

Figure 2. Partitioning of the $W(W)$ number into edge and path contributions illustrated with graph H of Figure 1.

TABLE III. The partitioning of $W(W)$ into path contributions, P_i, for all the 18 octane isomers

Isomer[(a)]	Path contribution						
	P_1	P_2	P_3	P_4	P_5	P_6	P_7
n-octane	140	70	35	26	32	40	35
2-M-heptane	118	93	33	34	50	70	
3-M-heptane	121	73	72	45	73	32	
4-M-heptane	122	73	60	101	36	31	
3-E-hexane	134	52	85	108	62		
2,2-MM-hexane	88	155	41	58	96		
2,3-MM-hexane	104	86	119	78	62		
2,4-MM-hexane	102	94	64	119	64		
2,5-MM-hexane	98	112	33	40	140		
3,3-MM-hexane	94	125	121	96	28		
3,4-MM-hexane	106	70	157	100	29		
3,2-EM-pentane	116	64	136	153			
3,3-EM-pentane	102	96	201	84			
2,2,3-MMM-pentane	78	138	182	87			
2,2,4-MMM-pentane	72	168	35	192			
2,3,3-MMM-pentane	80	128	227	56			
2,3,4-MMM-pentane	88	97	168	124			
2,2,3,3-MMMM-butane	54	192	261				

[(a)] Abbreviations M and E denote methyl and ethyl, respectively.

contributions arising from paths of length two by two, because each path of length two in a structure has been encountered twice. Similarly, contributions from paths of length three have to be divided by three, and so on, in order for all the contributions to add to $W(W)$. The edge contributions normalized in this manner are shown at the edges of the diagrams on the right hand side of Figure 2. By adding up all the contributions for a particular edge (*e.g.*, by adding 35 + 5 + 14.33 + 2.5 + 3.6 + 5.33, one gets 65.77 for edge 1–2), one obtains the contribution of the edge to $W(W)$, shown at the bottom of Figure 2. On the other hand, if one sums up all the edge contributions corresponding to paths of the same length, one obtains the partitioning of $W(W)$ according to paths of different length. In the case of paths of length two, we have: 5 + 17.5 + 7.5 + 6 + 14 + 9 + 14 = 73, the value shown on the right hand side of Figure 2. In Table III we show the partitioning of $W(W)$ into path contributions for all the 18 octane isomers, while Figure 3 shows the contributions of symmetry non-equivalent edges to $W(W)$ for all the 18 octane isomers. As can be seen from Figure 3, the edge partitioning of $W(W)$ shows variation in the relative contributions of the peripheral and inner edges. In less branched isomers, the terminal edges have a greater weight, but the opposite is the case in more branched isomers. In this respect, $W(W)$ behaves differently from several other topological indices, which either assign peripheral edges greater weight (*e.g.*, the connectivity index[12] and the Hosoya Z index[9]) or lesser weight (*e.g.*, the Wiener index[5] and Balaban's J index[13]).

The row sums of $\boldsymbol{G}_w$ matrix can be viewed as local vertex (atomic) descriptors. As one can see from Table I, the row sums associated with carbon atoms of 3-methylheptane show considerable variations. However, there are a number of cases of degeneracy of row sums within the set of 18 octane isomers, that is, the symmetry non-equivalent vertices in a molecular graph may have the same row sums. This is particularly apparent in the case of terminal non-equivalent vertices. The smallest row sum belongs to the most branched inner vertices. Local structural invariants such as the vertex, edge, and path contributions to $W(W)$ may be of interest in QSAR studies and, at the same time, they may help in the interpretation of $W(W)$ by pointing out the relative importance of individual vertices (atoms), edges (bonds), or subgraphs (molecular fragments).

The Wiener-Wiener number is but one of the structural invariants that can be extracted from the graphical matrix $\boldsymbol{G}$ and its numerical form $\boldsymbol{G}_w$. Therefore, we will continue to explore the construction of additional inva-

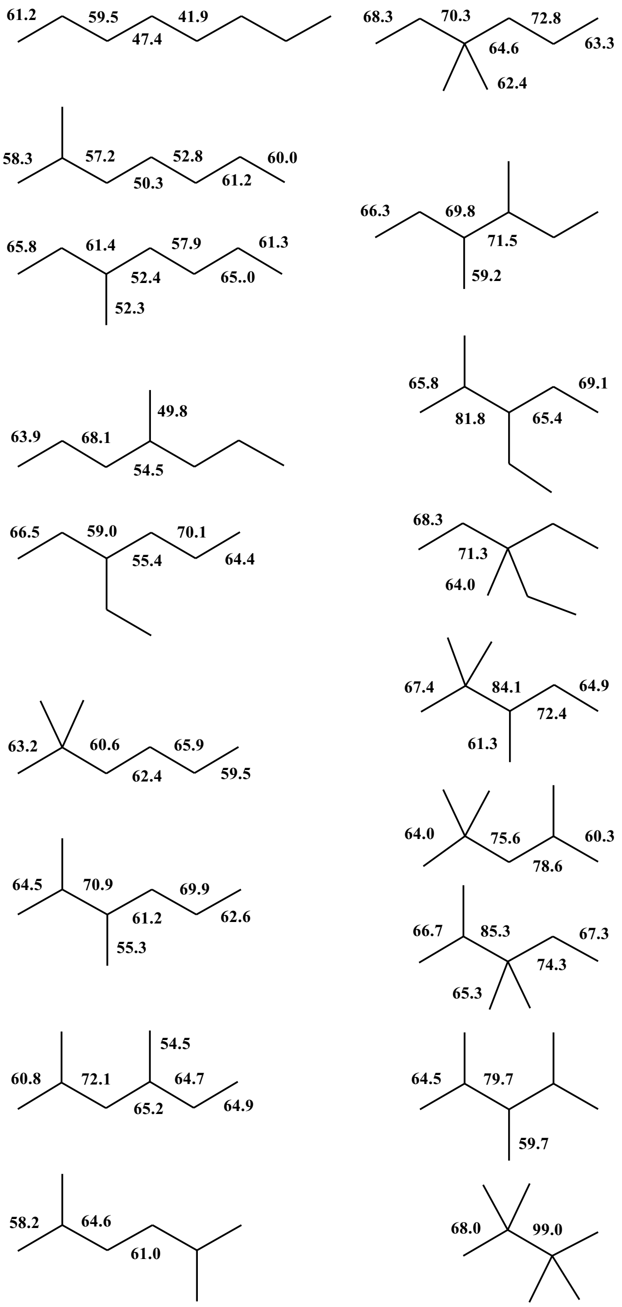

Figure 3. Molecular graphs of 18 octane isomers. Contributions to W(W) arising from all the symmetry non-equivalent edges in a graph are shown.

riants based on the $\boldsymbol{G}_w$ matrix. In particular, we will consider the leading eigenvalue, $\lambda(\boldsymbol{G}_w)$, of $\boldsymbol{G}_w$ matrix, in parallel with the leading eigenvalue, $\lambda(\boldsymbol{W})$, of the Wiener matrix $\boldsymbol{W}$ and the hyper-Wiener index WW[8,14–16] for comparison. In Table II, we list the values of these molecular descriptors for all the 18 octane isomers. The simple linear regression between $W(W)$ and $\lambda(\boldsymbol{G}_w)$ for the 18 octane isomers has a high correlation coefficient ($r = 0.9926$), yet there is sufficient variation between the two descriptors to produce different results in different regressions. The $W(W)$ number shows degeneracy for 4-methylheptane and 2,5-dimethylhexane, which is not the case with their corresponding $\lambda(\boldsymbol{G}_w)$ values. Similarly, 3-ethylhexane and 2,4-dimethylhexane have similar values of $W(W)$ but their corresponding $\lambda(\boldsymbol{G}_w)$ values show greater discrimination, and the same is again true of 3,2-ethylmethylpentane and 2,2,4-trimethylpentane.

MODELING OF TOTAL STERIC ENERGIES OF OCTANES

We have tested the usability of the novel molecular descriptors as predictor variables in QSPR studies on the total steric energies, ε, of octane isomers estimated from molecular fragments.[17] Octane isomers are a suitable set for testing a molecular descriptor, because they show sufficient structural variations. All the possible CC bond types present in alkanes appear in octanes, except for the isolated CC bond that only occurs in ethane. Moreover, by restricting attention to isomers, we have practically eliminated the dominant influence of the molecular size, which tends to obscure minor variations of properties caused by variations in the shape of molecules.

It has already been demonstrated that 2-D molecular descriptors are capable of capturing certain aspects of 3-D molecular structure, although they do not explicitly encode the spatial geometry of the molecule. For instance, when the connectivity indices $^1\chi$ and $^2\chi$ are used as descriptors of octane isomers, then the resulting multiple linear regression model of the enthalpies of formation of octanes in the gas phase[18] has the following statistical parameters: the multiple correlation coefficient $R = 0.9323$ and the standard error of estimate $s = 1.9929$ kJ mol^{-1}. As the residuals of this regression show poor correlation with the total steric energies of octanes ($r = 0.149$), one has to deduce that these two 2-D molecular descriptors account for the majority of this 3-D aspect of octane structures. Clearly, it is useful to test the novel molecular descriptors against the total steric energies of octanes, because this test indicates the ability of a descriptor to capture not only molecular connectivity (a 2-D molecular feature) but also some 3-D molecular characteristics.

The values of the statistical parameters R, s, and the Fisher ratio F for the quadratic regression models of the total steric energies of octane isomers based on different distance-based molecular descriptors are given in Table IV. In order to facilitate comparison, we show the values of these molecular descriptors for n-octane in the second column. The last four rows of the table contain the results obtained with descriptors that are not based on distance. As one can see, all the distance-based descriptors show much stronger correlations with the total steric energies of octane isomers in comparison with the molecular descriptors that are not based on distance. The best

TABLE IV. Values of various molecular descriptors for n-octane and the values of the multiple correlation coefficient (*R*), the standard error of estimate (*s*), and the Fisher ratio (*F*) for the quadratic regression models of total steric energies of octane isomers based on these molecular descriptors

Descriptor	n-octane	R	$\frac{s}{\text{kJ mol}^{-1}}$	$F^{2,15}$	Ref.
$W(W)$	378	0.9865	1.2971	272	(a)
WW	210	0.9839	1.4184	227	7
$\lambda(\boldsymbol{G_w})$	104.4359	0.9832	1.4467	217	(a)
$1/J$	0.1967	0.9813	1.5230	195	19
W	84	0.9806	1.5523	188	5
J	2.53	0.9766	1.7029	155	13
$\lambda(\boldsymbol{W})$	57.1698	0.9753	1.7531	146	8
H	13.743	0.9639	2.1087	98	20,21
$1/JJ$	0.3353	0.9497	2.4811	69	19
RRW_2	3.6667	0.9008	3.4434	32	22
ID	14.831	0.8998	3.4602	32	23
p_2/w_2	0.4583	0.8953	3.5313	30	24
χ	3.9142	0.7824	4.9371	12	12
Z	34	0.7036	5.6317	7	9

(a) This work.

TABLE V. Values of the total steric energies of octanes, ε, based on the work of Scott[17], the values of the total steric energies, $\varepsilon_{\text{calc}}$, calculated by means of the quadratic regression model based on *W*(*W*), and the corresponding residuals

Isomer(a)	$\frac{\varepsilon}{\text{kJ mol}^{-1}}$	$\frac{\varepsilon_{\text{calc}}}{\text{kJ mol}^{-1}}$	Residuals
n-octane	0	0.4683	–0.4683
2-M-heptane	2.5439	1.8924	0.6515
3-M-heptane	5.0877	4.0255	1.0622
4-M-heptane	5.0877	5.0729	0.0148
3-E-hexane	7.6316	8.3263	–0.6947
2,2-MM-hexane	5.8450	7.7280	–1.8830
2,3-MM-hexane	11.7905	10.0311	1.7594
2,4-MM-hexane	7.6316	8.7375	–1.1059
2,5-MM-hexane	5.0877	5.0729	0.0148
3,3-MM-hexane	11.6901	13.6571	–1.9670
3,4-MM-hexane	14.6984	13.1412	1.5572
3,2-EM-pentane	14.6984	14.9902	–0.2918
3,3-EM-pentane	17.5351	19.0540	–1.5189
2,2,3-MMM-pentane	18.7987	19.6744	–0.8757
2,2,4-MMM-pentane	15.6607	14.4495	1.2112
2,3,3-MMM-pentane	22.0999	21.5953	0.5046
2,3,4-MMM-pentane	19.2380	17.2526	1.9854
2,2,3,3-MMMM-butane	27.2002	27.1558	0.0444

(a) Abbreviations M and E denote methyl and ethyl, respectively.

result is obtained with the *W*(*W*) number introduced in this paper. In Table V we list the values of total steric energies of octanes based on the work of Scott,[17] the calculated values of the total steric energies by means of the quadratic regression model based on *W*(*W*), and the corresponding residuals. The plot of ε *versus* *W*(*W*) and the least-squares parabola are shown in Figure 4. Although the results appear satisfactory, we may notice that the computed values for the steric energies of 2,2-dimethylhexane and 3,3-dimethylhexane are much higher than the empirical values. On the other hand, the computed steric energies of 2,3-dimethylhexane and 2,3,4-trimethylpentane are somewhat lower than the corresponding empirical values. This may suggest that the relative values of Scott's empirical parameters, which are based on alkanes of different sizes, may be some-

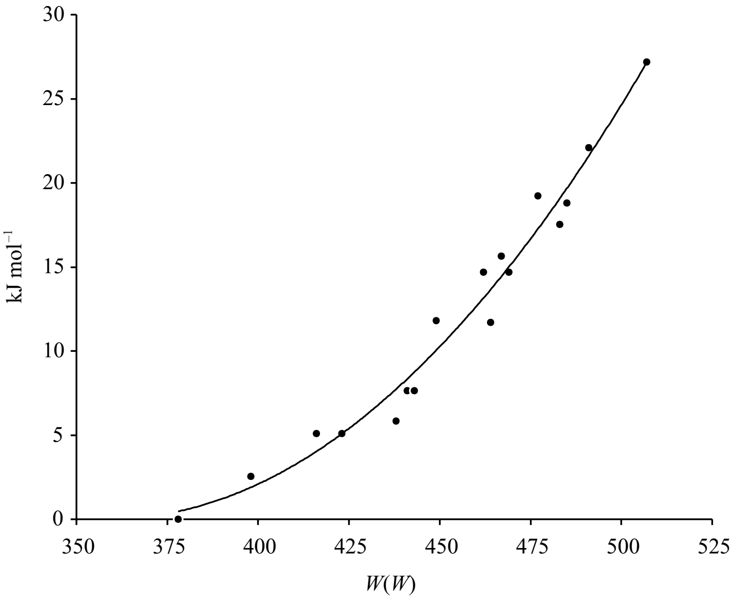

Figure 4. The plot of the values of the total steric energy based on the work of Scott[17] against *W*(*W*) for all the 18 octane isomers. The regression equation and the statistical parameters are ε = 0.0012 (± 0.0002) [*W*(*W*)]^2 – 0.8947 (± 0.0002)*W*(*W*) + 160.8084 (± 47.4667); R^2 = 0.9732; *s* = 1.2971; $F^{2,15}$ = 272.58.

what influenced by the size of the molecules. These somewhat larger residuals may indicate the inherent limitations of molecular descriptors to mimic experimental data beyond some critical precision value. On the other hand, they may also call for some revision of the empirical parameters. Comparison of regression models may facilitate interpretation of more involved molecular descriptors. For example, Table IV shows that the value of *R* for the model based on *ID* approaches those of distance-based descriptors and hence we may conclude that *ID* must contain a substantial input from distance-related structural factors. This is indeed the case because *ID* contains contributions of all paths, including longer paths, which indirectly encode some information on the distances between atoms. All the findings presented here support the claim that molecular descriptors, simple and elegant in their origin, are capable of characterizing important structural features of molecules.

Acknowledgement. – This work was supported in part by the Ministry of Science and Technology of the Republic of Croatia.

REFERENCES

1. M. Randić and N. Trinajstić, *J. Mol. Struct.* **300** (1993) 551–571.
2. M. Randić, M. Razinger, and D. Plavšić, *MATCH – Commun. Math. Comput. Chem.* **35** (1997) 243–259.
3. R. Diestel, *Graph Theory*, 2nd edition, Springer, New York, 2000.
4. P. V. O'Niel, *Am. Math. Monthly* **77** (1970) 35.
5. H. Wiener, *J. Am. Chem. Soc.* **69** (1947) 17–20.
6. M. Randić, *Acta Chim. Slov.* **49** (2002) 483–496.
7. M. Randić, *Chem. Phys. Lett.* **211** (1993) 478–483.
8. M. Randić, X. Guo, T. Oxley, H. Krishnapriyan, and R. Naylor, *J. Chem. Inf. Comput. Sci.* **34** (1994) 361–367.
9. H. Hosoya, *Bull. Chem. Soc. Jpn.* **44** (1971) 2332–2339.
10. Z. Mihalić, D. Veljan, D. Amić, S. Nikolić, D. Plavšić, and N. Trinajstić, *J. Math. Chem.* **11** (1992) 223–258.
11. M. Randić and J. Zupan, *J. Chem. Inf. Comput. Sci.* **41** (2001) 550–560.
12. M. Randić, *J. Am. Chem. Soc.* **97** (1975) 6609–6615.
13. A. T. Balaban, *Chem. Phys. Lett.* **89** (1982) 399–404.
14. D. J. Klein, I. Lukovits, and I. Gutman, *J. Chem. Inf. Comput. Sci.* **35** (1995) 50–52.
15. D. Plavšić, *Chem. Phys. Lett.* **304** (1999) 111–116.
16. D. Plavšić, N. Lerš, and K. Sertić-Bionda, *J. Chem. Inf. Comput. Sci.* **40** (2000) 516–519.
17. D. W. Scott, *J. Chem. Phys*. **60** (1974) 3144–3165.
18. M. Randić, *J. Mol. Struct. (Theochem)* **233** (1991) 45–59.
19. M. Randić and M. Pompe, *J. Chem. Inf. Comput. Sci.* **41** (2001) 575–581.
20. O. Ivanciuc, T. S. Balaban, and A. T. Balaban, *J. Math. Chem.* **12** (1993) 309–318.
21. D. Plavšić, S. Nikolić, N. Trinajstić, and Z. Mihalić, *J. Math. Chem.* **12** (1993) 235–250.
22. M. Randić, *Theor. Chim. Acta* **92** (1984) 164–175.
23. M. Randić, *J. Chem. Inf. Comput. Sci.* **24** (1984) 164–175.
24. M. Randić, *J. Chem. Inf. Comput. Sci.* **41** (2001) 607–613.

SAŽETAK

Nova grafička matrica i molekularni opisivači temeljeni na udaljenosti

Milan Randić, Nabamita Basak i Dejan Plavšić

Prikazana je važnost grafičkih matrica kao izvora numeričkih strukturnih invarijanata koje mogu biti uporabljene kao prediktorske varijable u QSPR i QSAR istraživanjima svojstava molekula. Predložena je nova grafička matrica ***G***, koja se pridružuje molekuli. Nedijagonalni element $[G]_{ij}$ matrice ***G*** podgraf je molekularnoga grafa razmatrane molekule dobiven iz njega uklanjanjem vrhova *i* i *j*. Nekoliko je molekularnih opisivača (deskriptora) izvedeno iz matrice ***G*** i njenoga numeričkoga oblika – matrice G_w, temeljenih na Wienerovome indeksu. Uporabivost izvedenih »dvostrukih« invarijanata kao prediktorskih varijabli u QSPR istraživanjima svojstava molekula testirana je na ukupnim steričkim energijama izomera oktana.