

## **INTRINSIC GRAPH DISTANCES COMPARED TO EUCLIDEAN DISTANCES FOR CORRESPONDENT GRAPH EMBEDDING<sup>#</sup>**

Ovidiu Ivanciuc,<sup>\*</sup> Teodora Ivanciuc, and Douglas J. Klein<sup>\*</sup>

Department of Marine Sciences, Texas A & M University at Galveston,  
Fort Crockett Campus, 5007 Avenue U, Galveston, TX 77551, USA  
E-mail: ivanciuc@netscape.net and kleind@tamug.tamu.edu, respectively

**Abstract.** Chemical structures of organic compounds are characterized numerically by a variety of structural descriptors computed either from the molecular graph or from the three-dimensional (3D) molecular geometry. Extensive use of such structural descriptors or topological indices has been made in drug design, screening of chemical databases, similarity and diversity assessment, and quantitative structure-activity relationships. In recent years a large variety of topological indices were derived from different sorts of graph distance functions which have been considered to characterize the molecular shape and structure. These include not only the shortest-path distance but also the resistance distance and the quasi-Euclidean distance. A comparison is made between five intrinsic graph distance functions and the geometric distance for a set of benzenoid hydrocarbons. Overall, a very good correlation is obtained for all graph distances, indicating that the graph descriptors derived from them capture some part of the 3D information of the molecular structure.

### **INTRODUCTION**

Professor Alexandru T. Balaban (Sandy) is one of the early developers and proponents of the field of topological indices (TIs) which are chemical descriptors derived from the

---

<sup>#</sup> Dedicated on the occasion of the 70<sup>th</sup> birthday to Professor Alexandru T. Balaban, who has long pursued the use of topological indices as QSAR and QSPR descriptors

molecular graphs, particularly of organic compounds, and which characterize numerically the molecular structure. Several recent reviews present the main uses of TIs in quantitative structure-property relationships (QSPR) and quantitative structure-activity relationships (QSAR), similarity and diversity assessment, database mining and in the virtual screening of combinatorial libraries [1-7]. One of the earliest and widely used topological index is the Wiener index  $W$  [8,9], derived from the shortest-path distances in a molecular graph. Its successful application in QSPR and QSAR stimulated the research in the domain of descriptors based on weighted graph distances [10], related novel molecular matrices [11] and other Wiener-like indices [12-16]. Elements of the distance matrix were used to define degree-distance  $VTI$  indices [17]; such vector-matrix invariants can generate TIs with a low degeneracy. The idea to use reciprocal distances in computing  $VTI$  indices was adopted in the definition of the reciprocal distance matrix **RD** [18-22]. Another distance measure was defined as the resistance distance matrix  $\Omega$  [23]; this metric is identical with that induced by the distance matrix only for acyclic compounds, while for cyclic compounds the resistance distance matrix offers the possibility to compute  $W'$ , an index related but not identical to the Wiener index. This original graph metric [24,25] was used to characterize the molecular cyclicity and centrality of polycyclic graphs [26] and to compute graph invariants for fullerenes [27]. Other recently defined distance-related matrices are the detour  $\Delta$  [28], detour-distance  $\Delta$ -**D** [28], distance-valency **Dval** [29], complementary distance **CD** [30], reciprocal complementary distance **RCD** [30], and reverse Wiener **RW** [31] matrices.

The possibility to compute the geometrical structure of (almost) any chemical compound with various molecular mechanics or quantum mechanics software and the growth of chemical databases containing the three-dimensional structure of a large number of compounds, it becomes possible to generate structural descriptors directly from the molecular geometry. Such structural invariants, computed from three-dimensional molecular geometry with equations initially defined for graph descriptors, are called topographic indices [32-38]. The structural descriptors computed from the molecular geometry offer a simple and efficient way for treating molecules with heteroatoms and multiple bonds. Using only the constitutional (connectivity) information contained in the molecular graph it is not possible to discriminate between *cis/trans* (*E/Z*) or other types of stereoisomers; however, topographic indices have different values for stereoisomers, and their use in QSAR/QSPR equations can improve their

predictive ability. In this paper we study the statistical relationship between five intrinsic graph distance functions and the geometric distance for a set of benzenoid hydrocarbons. This comparison can indicate the extent of overlap between the structural information contained in graph descriptors and geometric descriptors computed with the same formula.

## COMPUTATION OF VARIOUS DISTANCE FUNCTIONS

In this section we briefly present the five graph distance functions used in the comparison with the geometric distance for a set of molecular graphs. Each distance function encodes some aspect of the molecular characteristics (perhaps the “shape” of a molecular structure).

### **The Shortest-Path Distance $d$**

The shortest-path distance  $d$  is the standard distance function used in computing topological indices from molecular graphs. In simple (nonweighted) molecular graphs, in which all vertices represent carbon atoms and all edges represent carbon-carbon single bonds,  $d(i,j)$  between vertices  $v_i$  and  $v_j$  equals the number of edges on the shortest path between  $v_i$  and  $v_j$ . For weighted molecular graphs, representing organic molecules containing heteroatoms and multiple bonds various weighting schemes have been developed for computing the shortest-path distance  $d$  [3].

### **The Square-Root Shortest Path Distance $d_{1/2}$**

The square-root shortest path distance  $d_{1/2}$  was defined by Zhu and Klein [24] on the basis of the graph distance  $d$ :

$$d_{1/2}(i,j) = \{d(i,j)\}^{1/2} \quad (1)$$

That is, this function  $d_{1/2}$  turns out also to satisfy the conditions (positivity, symmetry, and triangle conditions) which make it a distance function.

## The Resistance Distance Distance $\Omega$

Klein and Randić introduced a new distance function on graphs named resistance distance, applying some results from the electrical network theory [23]; this novel graph distance was utilized to define the resistance distance matrix  $\Omega$ , proposed as an alternative to the distance matrix  $\mathbf{D}$ . For the computation of the molecular matrix  $\Omega$ , Klein and Randić superposed onto the molecular graph  $G$  an electrical network of resistors, in such a way that carbon atoms become nodes in the network and carbon-carbon single bonds are represented as 1 ohm resistors; the matrix element  $\Omega(i,j)$  is equal to the effective electrical resistance between the vertices  $v_i$  and  $v_j$ . From the theory of electrical networks it is easy to determine that in the case of acyclic compounds, the resistance distance matrix  $\Omega$  is identical with the distance matrix  $\mathbf{D}$ , while in the case of cyclic compounds the two matrices are different. From the several algorithms proposed for the computation of the resistance distance matrix  $\Omega$  [23-25] we have used the one that uses the eigenvalues and eigenvectors of the Laplacian matrix  $\mathbf{L}(G)$  of the molecular graph  $G$  with  $N$  vertices:

$$\mathbf{L} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^t \quad (2)$$

where  $\mathbf{U}$  is an  $N \times N$  column matrix of eigenvectors of the Laplacian matrix  $\mathbf{L}$ ,  $\mathbf{U}^t$  is the transpose matrix, and  $\mathbf{\Lambda}$  is an  $N \times N$  diagonal matrix containing on the main diagonal the eigenvalues of  $\mathbf{L}$ ; the eigenvalue  $[\mathbf{\Lambda}]_{ii}$  corresponds to the eigenvector from the  $i$ -th column of matrix  $\mathbf{U}$ . The  $N \times N$  diagonal matrix  $\mathbf{V}$  is computed from the eigenvalues of the Laplacian matrix  $\mathbf{L}$ :

$$[\mathbf{V}]_{ij} = \begin{cases} 0 & \text{if } [\mathbf{\Lambda}]_{ij} = 0 \\ [\mathbf{\Lambda}]_{ij}^{-1} & \text{if } [\mathbf{\Lambda}]_{ij} \neq 0 \end{cases} \quad (3)$$

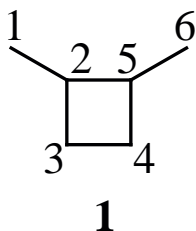
For any connected molecular graph the Laplacian matrix  $\mathbf{L}$  has all eigenvalues positive except for one which is 0; the generalized inverse of  $\mathbf{L}$  is the matrix  $\mathbf{\Gamma}$  which is 0 on this null eigenspace and the “true” inverse on the subspace orthogonal to this null space:

$$\mathbf{\Gamma} = \mathbf{U}\mathbf{V}\mathbf{U}^t \quad (4)$$

Then the resistance distance between vertices  $v_i$  and  $v_j$  is

$$\Omega(i, j) = [\Gamma]_{ii} - 2[\Gamma]_{ij} + [\Gamma]_{jj} \quad (5)$$

The computation of the resistance distance matrix  $\Omega$  of a simple (non-weighted) graph is presented for the molecular graph **1**, representing 1,2-dimethylcyclobutane.



The Laplacian matrix of 1,2-dimethylcyclobutane **1** is:

		<b>L(1)</b>					
		1	2	3	4	5	6
1		1	-1	0	0	0	0
2		-1	3	-1	0	-1	0
3		0	-1	2	-1	0	0
4		0	0	-1	2	-1	0
5		0	-1	0	-1	3	-1
6		0	0	0	0	-1	1

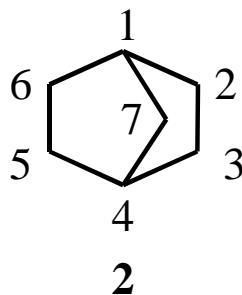
The generalized inverse of **L(1)** is the matrix  **$\Gamma(1)$** :

		<b><math>\Gamma(1)</math></b>					
		1	2	3	4	5	6
1		0.965	0.132	-0.160	-0.285	-0.243	-0.410
2		0.132	0.299	0.007	-0.118	-0.076	-0.243
3		-0.160	0.007	0.465	0.090	-0.118	-0.285
4		-0.285	-0.118	0.090	0.465	0.007	-0.160
5		-0.243	-0.076	-0.118	0.007	0.299	0.132
6		-0.410	-0.243	-0.285	-0.160	0.132	0.965

From Eq. (5) one obtains the corresponding resistance-distance matrix  **$\Omega(1)$** :

		<b><math>\Omega(1)</math></b>					
		1	2	3	4	5	6
1		0.000	1.000	1.750	2.000	1.750	2.750
2		1.000	0.000	0.750	1.000	0.750	1.750
3		1.750	0.750	0.000	0.750	1.000	2.000
4		2.000	1.000	0.750	0.000	0.750	1.750
5		1.750	0.750	1.000	0.750	0.000	1.000
6		2.750	1.750	2.000	1.750	1.000	0.000

An identical result can be obtained by applying the “series” and “parallel” transformation rules for an electric circuit of 1 ohm resistors corresponding to the molecular graph **1**. The second example for the computation of the resistance-distance matrix  $\mathbf{\Omega}$  considers a bicyclic compound, bicyclo[2.2.1]heptane (norbornane) **2**:



The Laplacian matrix of norbornane **2** is:

$\mathbf{L(2)}$							
	1	2	3	4	5	6	7
1	3	-1	0	0	0	-1	-1
2	-1	2	-1	0	0	0	0
3	0	-1	2	-1	0	0	0
4	0	0	-1	3	-1	0	-1
5	0	0	0	-1	2	-1	0
6	-1	0	0	0	-1	2	0
7	-1	0	0	-1	0	0	2

Using Eq. (4) one obtains the generalized inverse of  $\mathbf{L(2)}$ , the matrix  $\mathbf{\Gamma(2)}$ :

$\mathbf{\Gamma(2)}$							
	1	2	3	4	5	6	7
1	0.306	0.020	-0.122	-0.122	-0.122	0.020	0.020
2	0.020	0.497	0.116	-0.122	-0.218	-0.170	-0.122
3	-0.122	0.116	0.497	0.020	-0.170	-0.218	-0.122
4	-0.122	-0.122	0.020	0.306	0.020	-0.122	0.020
5	-0.122	-0.218	-0.170	0.020	0.497	0.116	-0.122
6	0.020	-0.170	-0.218	-0.122	0.116	0.497	-0.122
7	0.020	-0.122	-0.122	0.020	-0.122	-0.122	0.449

The generalized inverse matrix  $\mathbf{\Gamma(2)}$  gives the corresponding resistance-distance matrix  $\mathbf{\Omega(2)}$ :

		$\Omega(2)$						
	1	2	3	4	5	6	7	
1	0.000	0.762	1.048	0.857	1.048	0.762	0.714	
2	0.762	0.000	0.762	1.048	1.429	1.333	1.190	
3	1.048	0.762	0.000	0.762	1.333	1.429	1.190	
4	0.857	1.048	0.762	0.000	0.762	1.048	0.714	
5	1.048	1.429	1.333	0.762	0.000	0.762	1.190	
6	0.762	1.333	1.429	1.048	0.762	0.000	1.190	
7	0.714	1.190	1.190	0.714	1.190	1.190	0.000	

The hand application of the “series” and “parallel” transformation rules for electric circuits is often more difficult for polycyclic molecular graph and in some cases (as for many “cage” molecules) the technique is insufficient; in such cases it is preferable to compute  $\Omega$  via Eqs. (2-5).

### The Square-Root Resistance Distance $\Omega_{1/2}$

Analogously with the square-root shortest path distance  $d_{1/2}$ , Zhu and Klein [24] defined the square-root resistance distance  $\Omega_{1/2}$ :

$$\Omega_{1/2}(i,j) = \{\Omega(i,j)\}^{1/2} \quad (6)$$

The square-rooted distances of Eqs. (1) and (6) share with the usual Euclidean distances some extra properties [25] beyond that of being a distance function.

### The Quasi-Euclidean Distance $\rho_{qe}$

Another fundamental graph metric, based on the Laplacian matrix  $\mathbf{L}(G)$  and its generalized inverse  $\Gamma(G)$ , is the quasi-Euclidean metric  $\rho_{qe}$  [24]:

$$\rho_{qe}(i,j) = \{[\Gamma^2]_{ii} - 2[\Gamma^2]_{ij} + [\Gamma^2]_{jj}\}^{1/2} \quad (7)$$

The generalized inverse matrix  $\Gamma(\mathbf{1})$  of 1,2-dimethylcyclobutane gives the corresponding quasi-Euclidean matrix  $\rho_{qe}(\mathbf{1})$ :

		$\rho_{qe}(1)$					
		1	2	3	4	5	6
1		0.000	0.913	1.358	1.541	1.475	2.023
2		0.913	0.000	0.654	0.842	0.771	1.475
3		1.358	0.654	0.000	0.586	0.842	1.541
4		1.541	0.842	0.586	0.000	0.654	1.358
5		1.475	0.771	0.842	0.654	0.000	0.913
6		2.023	1.475	1.541	1.358	0.913	0.000

Similarly, for norbornane **2** the  $\Gamma(2)$  matrix gives the quasi-Euclidean matrix  $\rho_{qe}(2)$ :

		$\rho_{qe}(2)$						
		1	2	3	4	5	6	7
1		0.000	0.656	0.822	0.670	0.822	0.656	0.571
2		0.656	0.000	0.579	0.822	1.107	1.054	0.893
3		0.822	0.579	0.000	0.656	1.054	1.107	0.893
4		0.670	0.822	0.656	0.000	0.656	0.822	0.571
5		0.822	1.107	1.054	0.656	0.000	0.579	0.893
6		0.656	1.054	1.107	0.822	0.579	0.000	0.893
7		0.571	0.893	0.893	0.571	0.893	0.893	0.000

## DISSIMILARITY FUNCTIONS BETWEEN MOLECULAR DISTANCES

In order to compute the similarity/dissimilarity of graph and Euclidean (3D) distances, we use three distance functions, i.e. the arc-cosine coefficient  $AC$ , the power-covering distance measure  $\Delta_1$  [39], and mean-power distance measure  $\Delta_2$ . Because the scope of our investigation is to measure the similarity/dissimilarity of graph and Euclidean distances for the same molecule, each compound  $A$  with  $N$  non-hydrogen atoms is represented by a vector of interatomic distances  $\mathbf{X} = \mathbf{X}(A) = (x_1, x_2, x_3, \dots, x_n)$ , where  $n = N(N-1)/2$ ; as is usual in molecular graph theory, only non-hydrogen atoms are included in the graph. The comparisons between such  $n$ -sequences  $\mathbf{X}(A)$  are to be made so that the comparisons do not depend on rescalings, i.e.,  $\mathbf{X}(A)$  and the  $n$ -sequences  $s\mathbf{X}(A)$  with  $i$ -th element  $sx_i$  are to be treated as equivalent, for any positive scale factor  $s$ . That is, we focus on aspects of “shape” rather than scale (as especially is appropriate when the different distance functions to be compared might measure distances in different “units”).



### Arc-Cosine Metric $AC$

The cosine coefficient  $C$  for the similarity between two distance functions  $\mathbf{X}$  and  $\mathbf{Y}$  for molecule  $A$  is given by:

$$C(\mathbf{X}, \mathbf{Y}, A) = \frac{\sum_{i=1}^n \mathbf{X}(A)_i \mathbf{Y}(A)_i}{\left[ \sum_{i=1}^n \mathbf{X}(A)_i^2 \right]^{1/2} \times \left[ \sum_{i=1}^n \mathbf{Y}(A)_i^2 \right]^{1/2}} \quad (8)$$

with the property  $-1 \leq C \leq 1$ . The cosine coefficient measures the deviation of two datasets from proportionality. The arc-cosine distance measure is

$$AC(\mathbf{X}, \mathbf{Y}, A) = \arccos\{C(\mathbf{X}, \mathbf{Y}, A)\}$$

This is itself a metric on the scale equivalence classes of  $n$ -sequences (as  $\mathbf{X}_1(A)$ ,  $\mathbf{X}_2(A)$ , ...,  $\mathbf{X}_n(A)$ ).

### Power-Covering Metric $\Delta_1$

The power-covering metric  $\Delta_1$  [39] computes the dissimilarity between two interatomic distance vectors  $\mathbf{X}$  and  $\mathbf{Y}$  computed for structure  $A$  as:

$$\Delta_1(\mathbf{X}, \mathbf{Y}, A) = \log \left\{ \max_{i=1 \rightarrow n} \left( \frac{\mathbf{X}(A)_i}{\mathbf{Y}(A)_i} \right) \times \max_{i=1 \rightarrow n} \left( \frac{\mathbf{Y}(A)_i}{\mathbf{X}(A)_i} \right) \right\} \quad (9)$$

This also is a metric on the set of scale equivalence classes of  $n$ -sequences.

### Mean-Power Metric $\Delta_2$

A similar distance measure is obtained by considering all ratios of the interatomic distances from the vectors  $\mathbf{X}(A)$  and  $\mathbf{Y}(A)$ :

$$\Delta_2(\mathbf{X}, \mathbf{Y}, A) = \log \left\{ \left[ \frac{1}{n} \sum_{i=1}^n \frac{\mathbf{X}(A)_i}{\mathbf{Y}(A)_i} \right] \times \left[ \frac{1}{n} \sum_{i=1}^n \frac{\mathbf{Y}(A)_i}{\mathbf{X}(A)_i} \right] \right\} \quad (10)$$

This is a semi-metric on the set of scale equivalence classes of  $n$ -sequences, but it pays more attention to all the members of a sequence than does  $\Delta_1$  (which attends only to extreme ratios of members being compared).

### DISTANCE COMPARISONS FOR SINGLE CYCLES

As a first test we compare each of the five intrinsic graph distance functions to a geometric distance for a cyclic graph inscribed as a regular polygon into a circle. Such a polygon with  $N$  vertices represents a cycloalkane with  $N$  carbon atoms. However, for cycloalkanes the comparison of graph distances with geometric distances is complicated by the large flexibility of higher alkanes that results in the existence of many conformations. Since we are mainly interested in a preliminary comparison between graph and geometric distances, we have considered the much easier problem of the embedding of a polygon into a circle. All computations were done for polygons having between 3 and 40 vertices. The plots of the AC dissimilarity indices between the five graph distances ( $d$ ,  $d_{1/2}$ ,  $\Omega$ ,  $\Omega_{1/2}$ , and  $\rho_{qe}$ ) and the geometric distances between the vertices of a polygon embedded into a circle are presented in Figure 1. For the shortest path distance  $d$  one can observe an even-odd dependency between  $AC(d,3D)$  and  $N$ . For all five graph distances, their similarity with the geometric embedding decreases when  $N$  increases. The highest similarity with the geometric embedding is exhibited by the resistance distance  $\Omega$  followed by the quasi-Euclidean distance  $\rho_{qe}$ . If we neglect the AC values obtained for small  $N$ , the similarity with the geometric embedding decreases in the following order:  $\Omega$ ,  $\rho_{qe}$ ,  $d_{1/2}$ ,  $d$ ,  $\Omega_{1/2}$ . This ordering is apparent also from the average values for the AC index:  $AC(d,3D)_{av} = 0.1194$ ,  $AC(d_{1/2},3D)_{av} = 0.1035$ ,  $AC(\Omega,3D)_{av} = 0.0366$ ,  $AC(\Omega_{1/2},3D)_{av} = 0.1750$ ,  $AC(\rho_{qe},3D)_{av} = 0.0434$ . The AC values indicate that there is a significant similarity between graph and geometric embedding distances (especially for the cases of resistance distance and quasi-Euclidean distances), suggesting that the graph distance functions are able to capture a significant amount from the structural information present in the

geometric embedding of a polygon into a circle. All the AC-differences seem to approach finite asymptotic values.

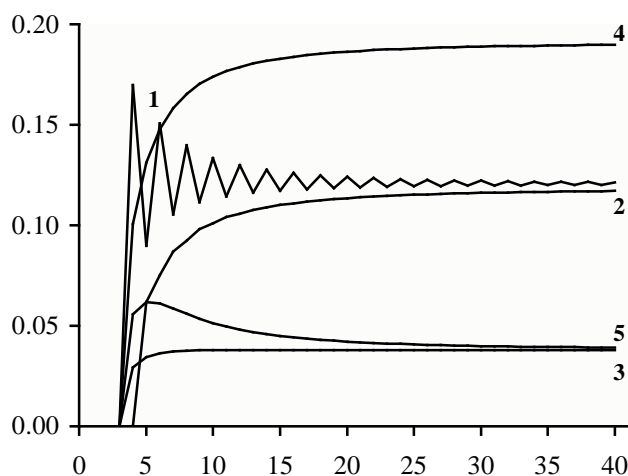


FIGURE 1. The AC index of dissimilarity between five graph distance measures ( $d$ , 1;  $d_{1/2}$ , 2;  $\Omega$ , 3;  $\Omega_{1/2}$ , 4;  $\rho_{qe}$ , 5) and the distance in a regular polygon, computed for 38 cycloalkanes  $C_N$  with  $N$  between 3 and 40.

In Figure 2 we present the plot of the power-covering distance indices  $\Delta_1$  computed for the same regular polygon embedding into a circle.

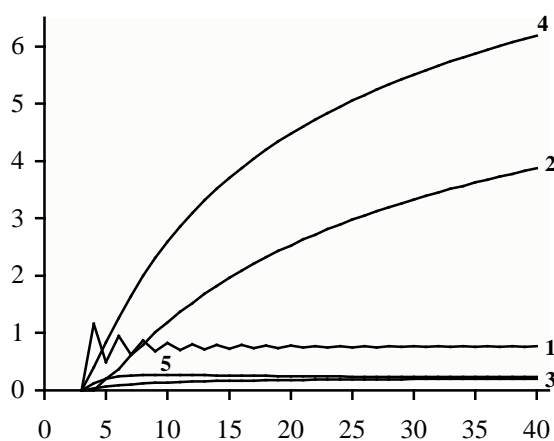


FIGURE 2. The  $\Delta_1$  index of dissimilarity between five graph distance measures ( $d$ , 1;  $d_{1/2}$ , 2;  $\Omega$ , 3;  $\Omega_{1/2}$ , 4;  $\rho_{qe}$ , 5) and the distance in a regular polygon, computed for 38 cycloalkanes  $C_N$  with  $N$  between 3 and 40.

Similarly with the situation encountered in the previous computations, there is an even-odd oscillation between the  $\Delta_1(d,3D)$  and  $N$ . The overall similarity with the geometric embedding decreases in the following order:  $\Omega$ ,  $\rho_{qe}$ ,  $d$ ,  $d_{1/2}$ ,  $\Omega_{1/2}$ . Compared with AC, an inversion occurs between  $d$  and  $d_{1/2}$ : with the  $\Delta_1$  dissimilarity indices,  $d$  is more similar to the geometric embedding than  $d_{1/2}$ . The average values for the  $\Delta_1$  index are:  $\Delta_1(d,3D)_{av} = 0.1738$ ,  $\Delta_1(d_{1/2},3D)_{av} = 0.3136$ ,  $\Delta_1(\Omega,3D)_{av} = 0.0772$ ,  $\Delta_1(\Omega_{1/2},3D)_{av} = 0.4184$ ,  $\Delta_1(\rho_{qe},3D)_{av} = 0.0944$ . The differences  $\Delta_1(d,3D)$ ,  $\Delta_1(\Omega,3D)$ , and  $\Delta_1(\rho_{qe},3D)$  all seem to approach finite asymptotic values (namely  $\log \pi/2$ ,  $\log 4/\pi$ , and  $\approx \log 4/\pi$ , respectively). On the other hand the differences  $\Delta_1(d_{1/2},3D)$  and  $\Delta_1(\Omega_{1/2},3D)$  both seem to diverge ( $\sim \log N$ ).

The plots of the mean-power distance indices  $\Delta_2$  (multiplied with 100) are presented in Figure 3 for the same five graph distances. The even-odd oscillation for  $\Delta_2(d,3D)$  is present but less severe than for the AC and  $\Delta_1$  indices. The ordering of the five graph distances by  $\Delta_2$  is identical with that induced by  $\Delta_1$ , with the following average values:  $\Delta_2(d,3D)_{av} = 0.7455$ ,  $\Delta_2(d_{1/2},3D)_{av} = 2.3803$ ,  $\Delta_2(\Omega,3D)_{av} = 0.1615$ ,  $\Delta_2(\Omega_{1/2},3D)_{av} = 4.1645$ ,  $\Delta_2(\rho_{qe},3D)_{av} = 0.2325$ . The convergence vs. divergence for the five different distances seems to be much similar as for  $\Delta_1$ .

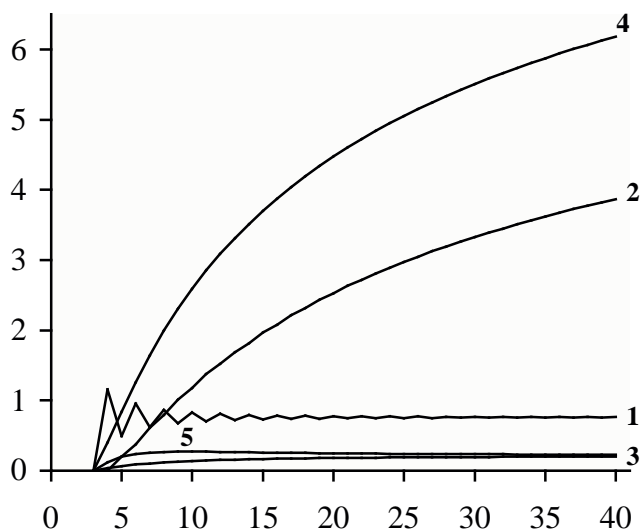


FIGURE 3. The  $\Delta_2$  ( $\times 100$ ) index of dissimilarity between five graph distance measures ( $d$ , 1;  $d_{1/2}$ , 2;  $\Omega$ , 3;  $\Omega_{1/2}$ , 4;  $\rho_{qe}$ , 5) and the distance in a regular polygon, computed for 38 cycloalkanes  $C_N$  with  $N$  between 3 and 40.

The conclusion of the plots from Figures 1-3 is that all five graph distances are highly correlated with the geometric distances between vertices of a polygon embedded into a circle, with the highest similarity exhibited by the resistance distance  $\Omega$  followed by the quasi-Euclidean distance  $\rho_{qe}$ . The square-rooted distance functions ( $d_{1/2}$  and  $\Omega_{1/2}$ ) usually compare the least favorably. These results suggest that in this particular case, the novel graph distances  $\Omega$  and  $\rho_{qe}$  encapsulate a large portion of the variation of the Euclidean distance. Much of the results for our three different comparators AC,  $\Delta_1$ , and  $\Delta_2$  are somewhat similar, so that in the following we look only at one of these – namely the AC comparator.

### DISTANCE COMPARISONS FOR POLYACENES AND POLYPHENACENES

The second comparison of the graph and Euclidean distances was performed for a set of polyacenes (linear benzenoid hydrocarbons) and polyphenacenes (zigzag benzenoid hydrocarbons). The geometries of the first 20 compounds from the two series of benzenoid hydrocarbons were optimized with the MM+ force field from HyperChem 5 [40]. For our purpose, the molecular mechanics results offer a good trade-off between reasonable molecular geometries and a short computation time.

The plots of the AC dissimilarity indices between the five graph distances and the Euclidean distances for the first 20 linear benzenoid hydrocarbons are presented in Figure 4.

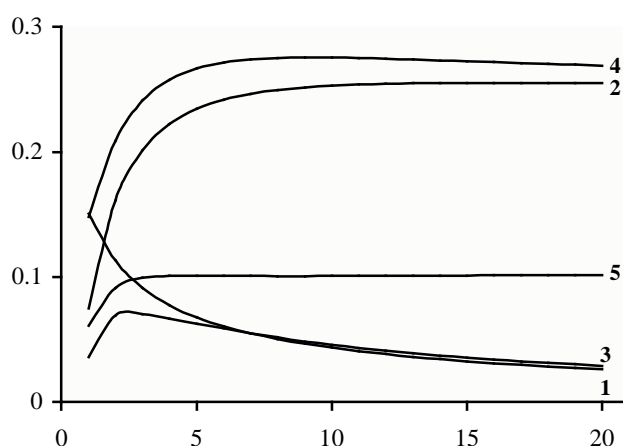


FIGURE 4. The AC index of dissimilarity between five graph distance measures ( $d$ , 1;  $d_{1/2}$ , 2;  $\Omega$ , 3;  $\Omega_{1/2}$ , 4;  $\rho_{qe}$ , 5) and the geometric distance, computed for the first 20 linear benzenoid hydrocarbons.

The trend exhibited by the higher benzenoid hydrocarbons indicates that the similarity between graph and Euclidean distances decreases in the following order:  $d$ ,  $\Omega$ ,  $\rho_{qe}$ ,  $d_{1/2}$ ,  $\Omega_{1/2}$ . However, for the first benzenoid hydrocarbons in this series, i.e. benzene and naphthalene,  $\Omega$  and  $\rho_{qe}$  have a higher similarity with the Euclidean distance than  $d$ ; as the number of benzenoid rings increases,  $d$  becomes more similar with the Euclidean distance than  $\Omega$  and  $\rho_{qe}$ . The average AC indices indicate that  $\Omega$  and  $d$  are very similar with the Euclidean distance:  $AC(d,3D)_{av} = 0.0540$ ,  $AC(d_{1/2},3D)_{av} = 0.2343$ ,  $AC(\Omega,3D)_{av} = 0.0459$ ,  $AC(\Omega_{1/2},3D)_{av} = 0.2608$ ,  $AC(\rho_{qe},3D)_{av} = 0.0985$ . All five comparisons to the Euclidean distances appear to approach finite asymptotic values. Apparently the shortest-path and resistance distances both approach ever more closely the Euclidean distance, the differences seemingly approaching zero as the number of benzenoid rings becomes ever larger.

The same statistical analysis was performed for the first 20 zigzag benzenoid hydrocarbons, and the results are presented in Figure 5. Because the first two molecules (benzene and naphthalene) are identical in the two series of benzenoid hydrocarbons, the first parts of Figures 4 and 5 are identical. The plots from Figure 5 indicate that the similarity between the Euclidean distance and the graph distances  $d$ ,  $\Omega$ , and  $\rho_{qe}$  increases when the number of benzenoid rings increases. The shortest-path distance  $d$  has the greatest correlation with the Euclidean distance, as indicated also by the average values of AC:  $AC(d,3D)_{av} = 0.0499$ ,  $AC(d_{1/2},3D)_{av} = 0.2196$ ,  $AC(\Omega,3D)_{av} = 0.0655$ ,  $AC(\Omega_{1/2},3D)_{av} = 0.2614$ ,  $AC(\rho_{qe},3D)_{av} = 0.1053$ . The asymptotic behaviors seem to mirror those for the polyacenes.

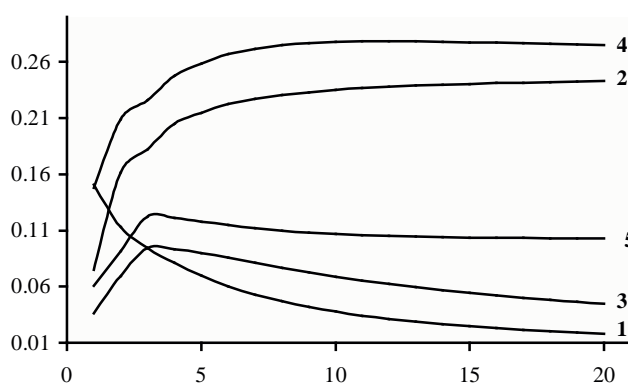


FIGURE 5. The AC index of dissimilarity between five graph distance measures ( $d$ , 1;  $d_{1/2}$ , 2;  $\Omega$ , 3;  $\Omega_{1/2}$ , 4;  $\rho_{qe}$ , 5) and the geometric distance, computed for the first 20 zigzag benzenoid hydrocarbons.

The results presented in Figures 4 and 5 indicate that, at least for benzenoid hydrocarbons, topological indices computed with the shortest-path distance  $d$  are remarkable similar with topographical descriptors obtained from the Euclidean distance. Especially the shortest-path and resistance distances seem ultimately to compare quite favorably with the Euclidean distances.

### INVERSE DISTANCE COMPARISONS FOR POLYACENES AND POLYPHENACENES

Reciprocal graph distances were used with success in computing a large number of topological indices that proved to be very useful in QSAR and QSPR models [2,4,17-22]. Therefore, it seems of interest to compare reciprocal graph distances with the reciprocal Euclidean distance. The AC dissimilarity indices were computed for the same sets of 20 linear and zigzag benzenoid hydrocarbons, giving the plots presented in Figures 6 and 7.

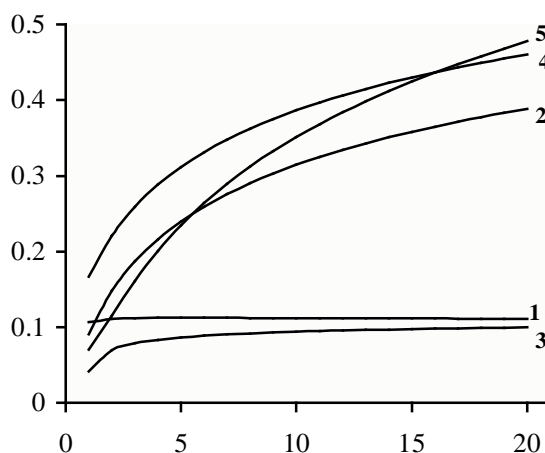


FIGURE 6. The AC index of dissimilarity between five reciprocal graph distance measures ( $d^{-1}$ , 1;  $d_{1/2}^{-1}$ , 2;  $\Omega^{-1}$ , 3;  $\Omega_{1/2}^{-1}$ , 4;  $\rho_{qe}^{-1}$ , 5) and the reciprocal geometric distance, computed for the first 20 linear benzenoid hydrocarbons.

For linear benzenoids, the similarity between the reciprocal shortest-path and Euclidean distances is almost constant irrespective of the number of benzenoid rings. The average values of AC show that the highest similarity is obtained between  $\Omega^{-1}$  and  $3D^{-1}$ :  $AC(d^{-1}, 3D^{-1})_{av} = 0.1116$ ,  $AC(d_{1/2}^{-1}, 3D^{-1})_{av} = 0.2961$ ,  $AC(\Omega^{-1}, 3D^{-1})_{av} = 0.0897$ ,  $AC(\Omega_{1/2}^{-1}, 3D^{-1})_{av} = 0.3684$ ,

$AC(\rho_{qe}^{-1}, 3D^{-1})_{av} = 0.3305$ . The results presented in Figure 6 show also that a high correlation exists between  $d^{-1}$  and  $3D^{-1}$ , indicating that structural descriptors from these two molecular distances are also highly intercorrelated.

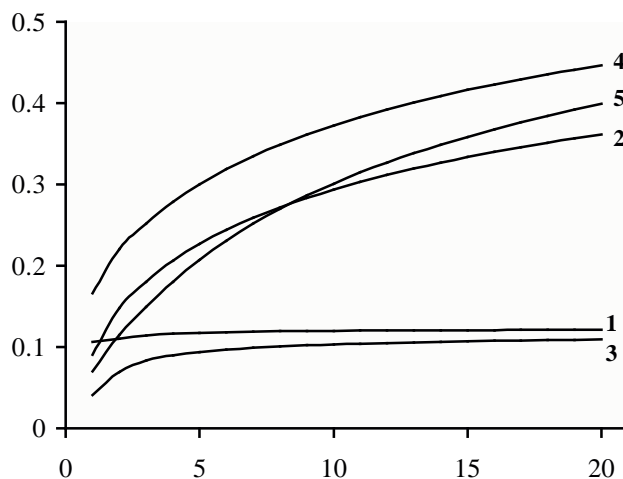


FIGURE 7. The AC index of dissimilarity between five reciprocal graph distance measures ( $d^{-1}$ , 1;  $d_{1/2}^{-1}$ , 2;  $\Omega^{-1}$ , 3;  $\Omega_{1/2}^{-1}$ , 4;  $\rho_{qe}^{-1}$ , 5) and the reciprocal geometric distance, computed for the first 20 zigzag benzenoid hydrocarbons.

A rather similar behavior is again obtained for the polyphenacenes, as one can see from the plots from Figure 7. Especially the reciprocal shortest-path and resistance distances compare well with reciprocal Euclidean distances, the differences (as measured by AC) evidently approaching a finite value as the number of benzenoid rings increases, while it seems there is a divergence for  $d_{1/2}^{-1}$ ,  $\Omega_{1/2}^{-1}$ , and  $\rho_{qe}^{-1}$ . The highest similarity with the reciprocal Euclidean distance is obtained with the reciprocal resistance distance  $\Omega^{-1}$  and reciprocal shortest-path distance  $d^{-1}$ , respectively, while the remaining three reciprocal distance functions have a much lower correlation with  $3D^{-1}$ . The average AC values for the reciprocal distances of zigzag benzenoids are:  $AC(d^{-1}, 3D^{-1})_{av} = 0.1186$ ,  $AC(d_{1/2}^{-1}, 3D^{-1})_{av} = 0.2779$ ,  $AC(\Omega^{-1}, 3D^{-1})_{av} = 0.0977$ ,  $AC(\Omega_{1/2}^{-1}, 3D^{-1})_{av} = 0.3568$ ,  $AC(\rho_{qe}^{-1}, 3D^{-1})_{av} = 0.2837$ . While in comparing graph distances to geometric distances  $d$  proved to be more similar to  $3D$  than did  $\Omega$ , the reverse situation is found when comparing reciprocal distances.



## SQUARED INVERSE DISTANCE COMPARISONS FOR POLYACENES AND POLYPHENACENES

An overview of the molecular graph descriptors indicates that distance functions are the most largely used, with reciprocal distances gaining a wider acceptance in recent years. Other modifications of the elements of the shortest-path distances were proposed in the squared reciprocal distance matrix [41] or the distance-valency **Dval** matrix [29]. Since these variants of graph distances deserve more consideration as sources of graph descriptors, we have investigated the similarity between the five squared reciprocal graph distance measures and the squared reciprocal Euclidean distance, using as benchmark the sets of 20 linear and zigzag benzenoid hydrocarbons.

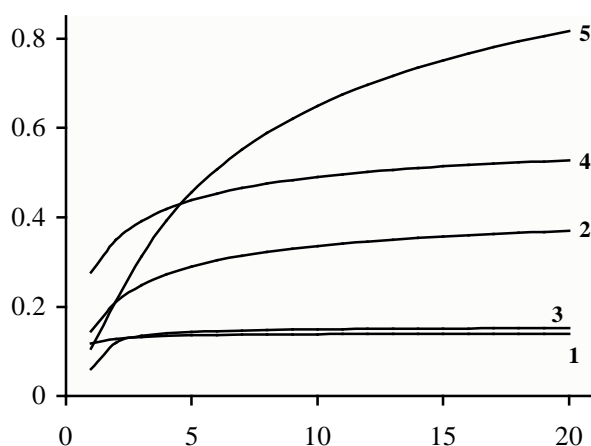


FIGURE 8. The AC index of dissimilarity between five squared reciprocal graph distance measures ( $d^{-2}$ , 1;  $d_{1/2}^{-2}$ , 2;  $\Omega^{-2}$ , 3;  $\Omega_{1/2}^{-2}$ , 4;  $\rho_{qe}^{-2}$ , 5) and the squared reciprocal geometric distance, computed for the first 20 linear benzenoid hydrocarbons.

The plots for linear benzenoids from Figure 8 show that, similarly with the results obtained for reciprocal distances, the correlation between  $d^{-2}$  and  $3D^{-2}$  is almost constant irrespective of the number of benzenoid rings; the same observation can be made about the correlation between  $\Omega^{-2}$  and  $3D^{-2}$ . The similarity between the squared reciprocal Euclidean distance and the graph invariants  $d^{-2}$  and  $\Omega^{-2}$  is remarkable high, and decreases considerably for  $d_{1/2}^{-2}$ ,  $\Omega_{1/2}^{-2}$ , and  $\rho_{qe}^{-2}$ . The average AC values display the same trend:  $AC(d^{-2}, 3D^{-2})_{av} =$

0.1366,  $AC(d_{1/2}^{-2}, 3D^{-2})_{av} = 0.3172$ ,  $AC(\Omega^{-2}, 3D^{-2})_{av} = 0.1425$ ,  $AC(\Omega_{1/2}^{-2}, 3D^{-2})_{av} = 0.4692$ ,  $AC(\rho_{qe}^{-2}, 3D^{-2})_{av} = 0.5966$ .

For the set of 20 zigzag benzenoids the AC dissimilarity indices between squared reciprocal graph and Euclidean distances the plots are very similar with those obtained for linear benzenoids, as can be seen from Figure 9.

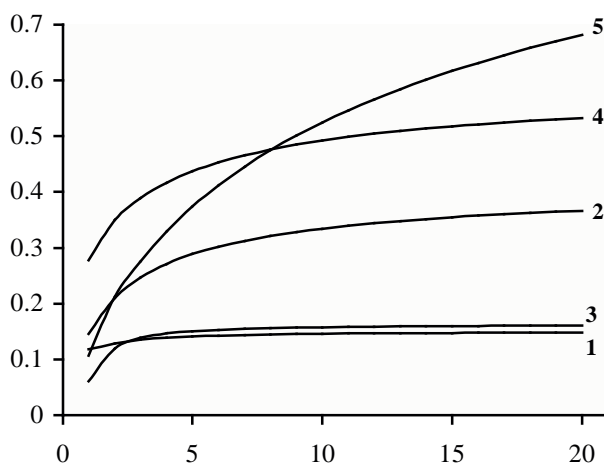


FIGURE 9. The AC index of dissimilarity between five squared reciprocal graph distance measures ( $d^{-2}$ , 1;  $d_{1/2}^{-2}$ , 2;  $\Omega^{-2}$ , 3;  $\Omega_{1/2}^{-2}$ , 4;  $\rho_{qe}^{-2}$ , 5) and the squared reciprocal geometric distance, computed for the first 20 zigzag benzenoid hydrocarbons.

For higher zigzag benzenoids the similarity with the squared reciprocal Euclidean distance decreases in the following order:  $d$ ,  $\Omega$ ,  $d_{1/2}$ ,  $\Omega_{1/2}$ ,  $\rho_{qe}$ . This ordering is apparent also from the average values for the AC index:  $AC(d^{-2}, 3D^{-2})_{av} = 0.1431$ ,  $AC(d_{1/2}^{-2}, 3D^{-2})_{av} = 0.3156$ ,  $AC(\Omega^{-2}, 3D^{-2})_{av} = 0.1497$ ,  $AC(\Omega_{1/2}^{-2}, 3D^{-2})_{av} = 0.4711$ ,  $AC(\rho_{qe}^{-2}, 3D^{-2})_{av} = 0.4928$ . For the same set of 20 zigzag benzenoids we present in Figure 10 the plots of the mean-power distance indices  $\Delta_2$  (multiplied with 100), which exhibit a slightly different pattern compared with AC, as can be seen also from the average values for the  $\Delta_2$  index:  $\Delta_2(d^{-2}, 3D^{-2})_{av} = 1.5772$ ,  $\Delta_2(d_{1/2}^{-2}, 3D^{-2})_{av} = 17.8780$ ,  $\Delta_2(\Omega^{-2}, 3D^{-2})_{av} = 1.8573$ ,  $\Delta_2(\Omega_{1/2}^{-2}, 3D^{-2})_{av} = 24.9879$ ,  $\Delta_2(\rho_{qe}^{-2}, 3D^{-2})_{av} = 7.5222$ .

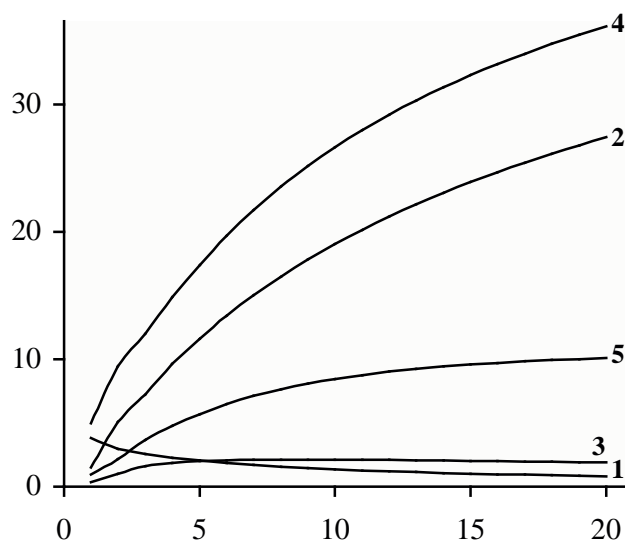


FIGURE 10. The  $\Delta_2$  ( $\times 100$ ) index of dissimilarity between five squared reciprocal graph distance measures ( $d^{-2}$ , 1;  $d_{1/2}^{-2}$ , 2;  $\Omega^{-2}$ , 3;  $\Omega_{1/2}^{-2}$ , 4;  $\rho_{qe}^{-2}$ , 5) and the squared reciprocal geometric distance, computed for the first 20 zigzag benzenoid hydrocarbons.

### DISTANCE COMPARISONS FOR THE 76 BENZENOIDS OF UP TO 6 RINGS

The final comparison of the graph and Euclidean distances was performed for the 76 benzenoid hydrocarbons of up to six benzenoid rings presented in Figure 11. This test set is more diverse than the previous ones, and can indicate if our earlier findings are more general or represent only a particular situation encountered only for cycles and for linear and zigzag benzenoids. The geometries of all 76 benzenoids were optimized with the MM+ force field from HyperChem 5. Due to the presence of bay regions, some compounds from this set are not planar, as indicated also by the final geometry obtained from the molecular mechanics computations. On the other hand, all these compounds are rigid, and the problem of selecting the global minimum conformer is avoided. In Table 1 we present the AC dissimilarity index between five graph distance measures and the Euclidean distance, computed for the 76 benzenoids from Figure 11. These numbers exhibit the similarity with the Euclidean distance decreases in the following order:  $d$ ,  $\Omega$ ,  $\rho_{qe}$ ,  $d_{1/2}$ ,  $\Omega_{1/2}$ . The similarity between the Euclidean distance and the shortest-path and resistance distances is very high, as can be seen from the

TABLE 1. The AC index of dissimilarity between five graph distance measures and the geometric distance, computed for the 76 benzenoid hydrocarbons from Figure 11

BH	AC( $d$ ,3D)	AC( $d_{1/2}$ ,3D)	AC( $\Omega$ ,3D)	AC( $\Omega_{1/2}$ ,3D)	AC( $\rho_{qe}$ ,3D)
3	0.1508	0.0750	0.0362	0.1479	0.0611
4	0.1131	0.1616	0.0694	0.2089	0.0915
5	0.0942	0.1823	0.0943	0.2259	0.1226
6	0.0910	0.2011	0.0703	0.2410	0.0992
7	0.0869	0.1731	0.1079	0.2178	0.1216
8	0.1080	0.1793	0.1425	0.2203	0.1807
9	0.0812	0.2052	0.0934	0.2475	0.1215
10	0.0792	0.2040	0.0883	0.2426	0.1223
11	0.0772	0.2223	0.0669	0.2577	0.1009
12	0.0963	0.1685	0.1012	0.2310	0.0968
13	0.0959	0.1836	0.1333	0.2270	0.1603
14	0.1665	0.1775	0.2177	0.2192	0.2604
15	0.0956	0.2026	0.1248	0.2396	0.1619
16	0.0698	0.2147	0.0901	0.2585	0.1180
17	0.0680	0.2117	0.0884	0.2517	0.1241
18	0.0713	0.2206	0.0849	0.2586	0.1189
19	0.0759	0.1905	0.0999	0.2303	0.1286
20	0.1097	0.1885	0.1451	0.2262	0.1837
21	0.0825	0.2008	0.1125	0.2355	0.1452
22	0.0685	0.2095	0.0844	0.2454	0.1250
23	0.0697	0.2204	0.0802	0.2549	0.1152
24	0.0677	0.2345	0.0628	0.2667	0.1011
25	0.0873	0.1806	0.1152	0.2331	0.1290
26	0.0871	0.1748	0.1211	0.2290	0.1356
27	0.0822	0.1964	0.1064	0.2491	0.1272
28	0.1007	0.1819	0.1451	0.2257	0.1689
29	0.0942	0.2137	0.1247	0.2429	0.1608
30	0.0958	0.1890	0.1322	0.2280	0.1591
31	0.1439	0.1921	0.1933	0.2328	0.2352
32	0.1604	0.1945	0.2016	0.2310	0.2431
33	0.1469	0.1886	0.2004	0.2328	0.2380
34	0.0853	0.2010	0.1205	0.2426	0.1570
35	0.0894	0.2182	0.1098	0.2507	0.1550
36	0.0903	0.2061	0.1202	0.2462	0.1491
37	0.0901	0.2009	0.1381	0.2460	0.1456
38	0.0604	0.2225	0.0857	0.2670	0.1150
39	0.0943	0.2029	0.1286	0.2410	0.1643
40	0.0753	0.1974	0.1110	0.2358	0.1494
41	0.0616	0.2227	0.0836	0.2629	0.1190
42	0.0622	0.2227	0.0822	0.2629	0.1146

TABLE 1. (Continued)

<b>43</b>	0.0646	0.2316	0.0781	0.2665	0.1144
<b>44</b>	0.0829	0.2206	0.1086	0.2540	0.1410
<b>45</b>	0.0811	0.2021	0.1155	0.2418	0.1324
<b>46</b>	0.0685	0.1925	0.0949	0.2298	0.1274
<b>47</b>	0.1021	0.1814	0.1396	0.2212	0.1762
<b>48</b>	0.1015	0.1990	0.1275	0.2324	0.1738
<b>49</b>	0.0994	0.2006	0.1370	0.2400	0.1757
<b>50</b>	0.1909	0.1858	0.2365	0.2247	0.2766
<b>51</b>	0.1278	0.1795	0.1600	0.2155	0.2029
<b>52</b>	0.1385	0.1875	0.1827	0.2228	0.2122
<b>53</b>	0.0763	0.2174	0.1019	0.2496	0.1367
<b>54</b>	0.0838	0.1997	0.1159	0.2322	0.1497
<b>55</b>	0.0593	0.2182	0.0817	0.2556	0.1230
<b>56</b>	0.0621	0.2220	0.0811	0.2576	0.1162
<b>57</b>	0.0619	0.2181	0.0780	0.2509	0.1201
<b>58</b>	0.0644	0.2310	0.0764	0.2650	0.1154
<b>59</b>	0.0679	0.2102	0.0899	0.2454	0.1294
<b>60</b>	0.1054	0.2045	0.1376	0.2387	0.1690
<b>61</b>	0.0687	0.2169	0.0930	0.2479	0.1221
<b>62</b>	0.0628	0.2314	0.0731	0.2627	0.1088
<b>63</b>	0.0606	0.2419	0.0588	0.2715	0.1009
<b>64</b>	0.0818	0.1963	0.1229	0.2450	0.1226
<b>65</b>	0.0918	0.1872	0.1433	0.2355	0.1876
<b>66</b>	0.0718	0.2084	0.1070	0.2576	0.1446
<b>67</b>	0.0775	0.1870	0.1173	0.2360	0.1329
<b>68</b>	0.0790	0.1986	0.1137	0.2443	0.1318
<b>69</b>	0.0762	0.2021	0.1050	0.2453	0.1481
<b>70</b>	0.0937	0.1823	0.1409	0.2308	0.1716
<b>71</b>	0.0735	0.2160	0.1024	0.2652	0.1070
<b>72</b>	0.0935	0.1802	0.1510	0.2330	0.1704
<b>73</b>	0.0763	0.1981	0.1127	0.2404	0.1607
<b>74</b>	0.0903	0.1985	0.1369	0.2457	0.1789
<b>75</b>	0.0711	0.2081	0.1030	0.2593	0.1106
<b>76</b>	0.0716	0.2168	0.0981	0.2599	0.1414
<b>77</b>	0.0864	0.1736	0.1231	0.2381	0.1168
<b>78</b>	0.0832	0.1879	0.1098	0.2515	0.1020
Mean	0.0891	0.1996	0.1141	0.2411	0.1444

average AC values:  $AC(d,3D)_{av} = 0.0891$ ,  $AC(d_{1/2},3D)_{av} = 0.1996$ ,  $AC(\Omega,3D)_{av} = 0.1141$ ,  $AC(\Omega_{1/2},3D)_{av} = 0.2411$ ,  $AC(\rho_{qe},3D)_{av} = 0.1444$ . This finding is significant for the computation of distance-based QSAR/QSPR structural descriptors, suggesting that either  $d$  or  $\Omega$  should

make a viable substitution for the Euclidean distance in computing structural indices, at least for rigid organic compounds.

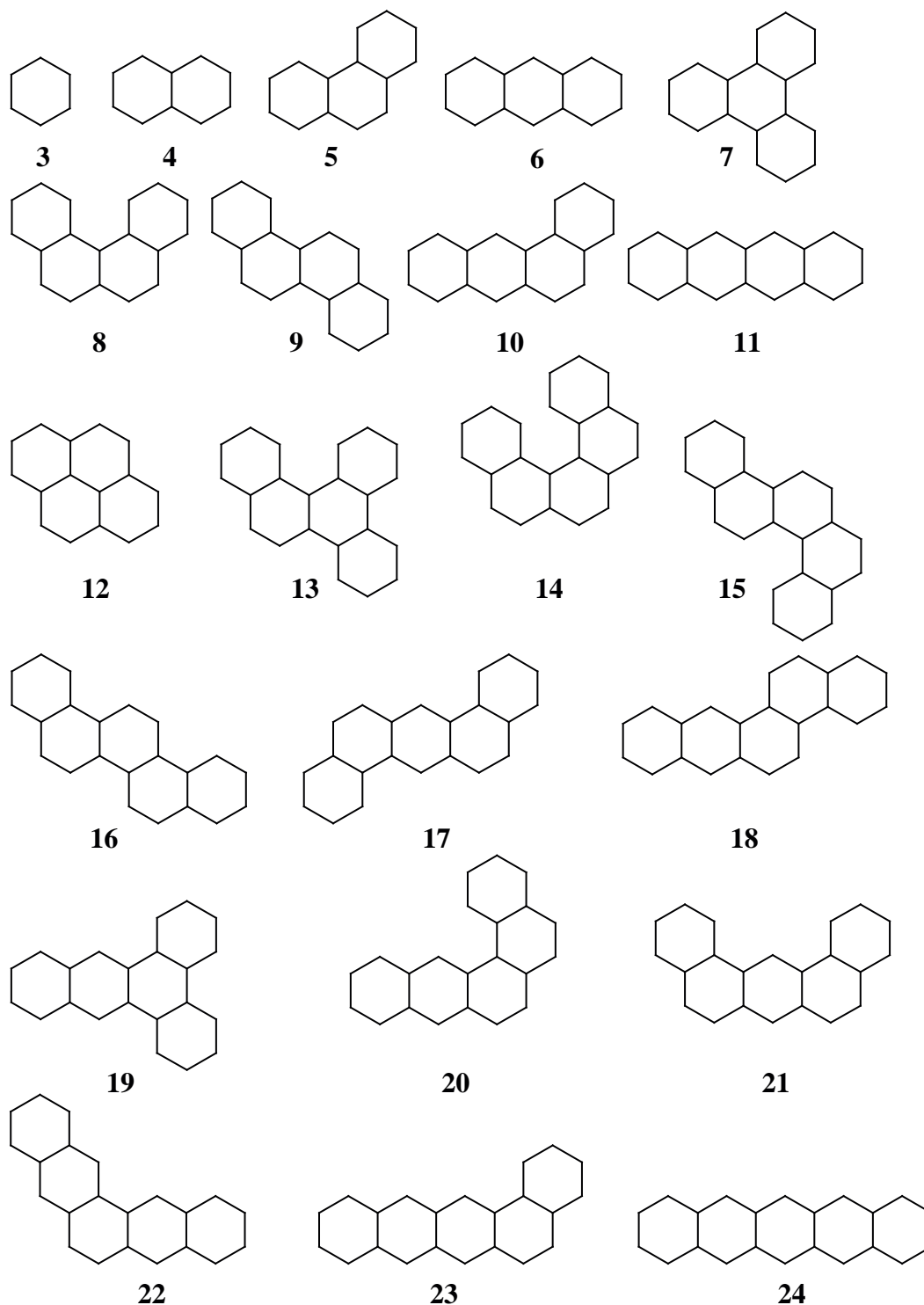


FIGURE 11. Benzenoid hydrocarbons used to measure the similarity of different distance functions.

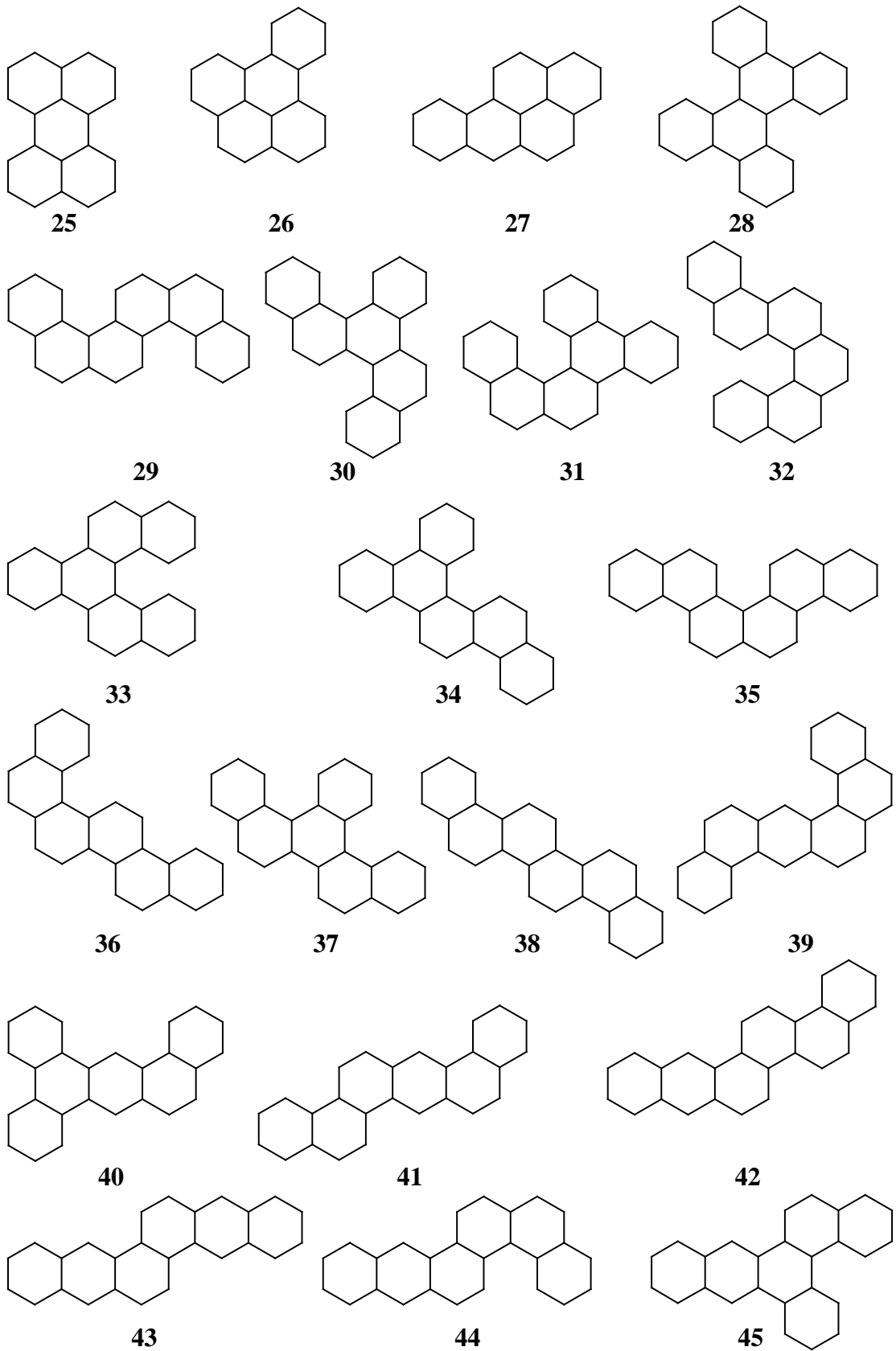


FIGURE 11. (Continued)

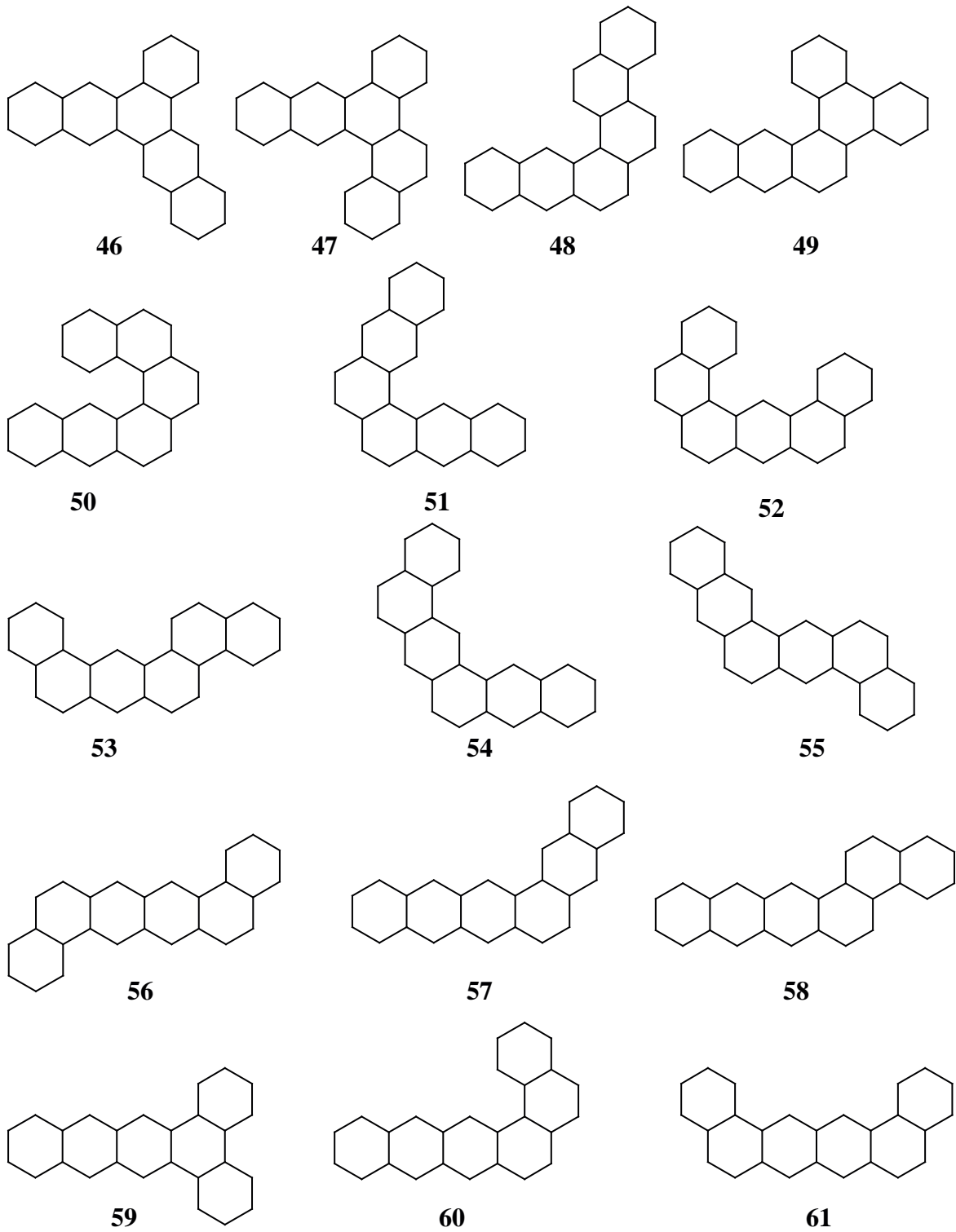


FIGURE 11. (Continued)



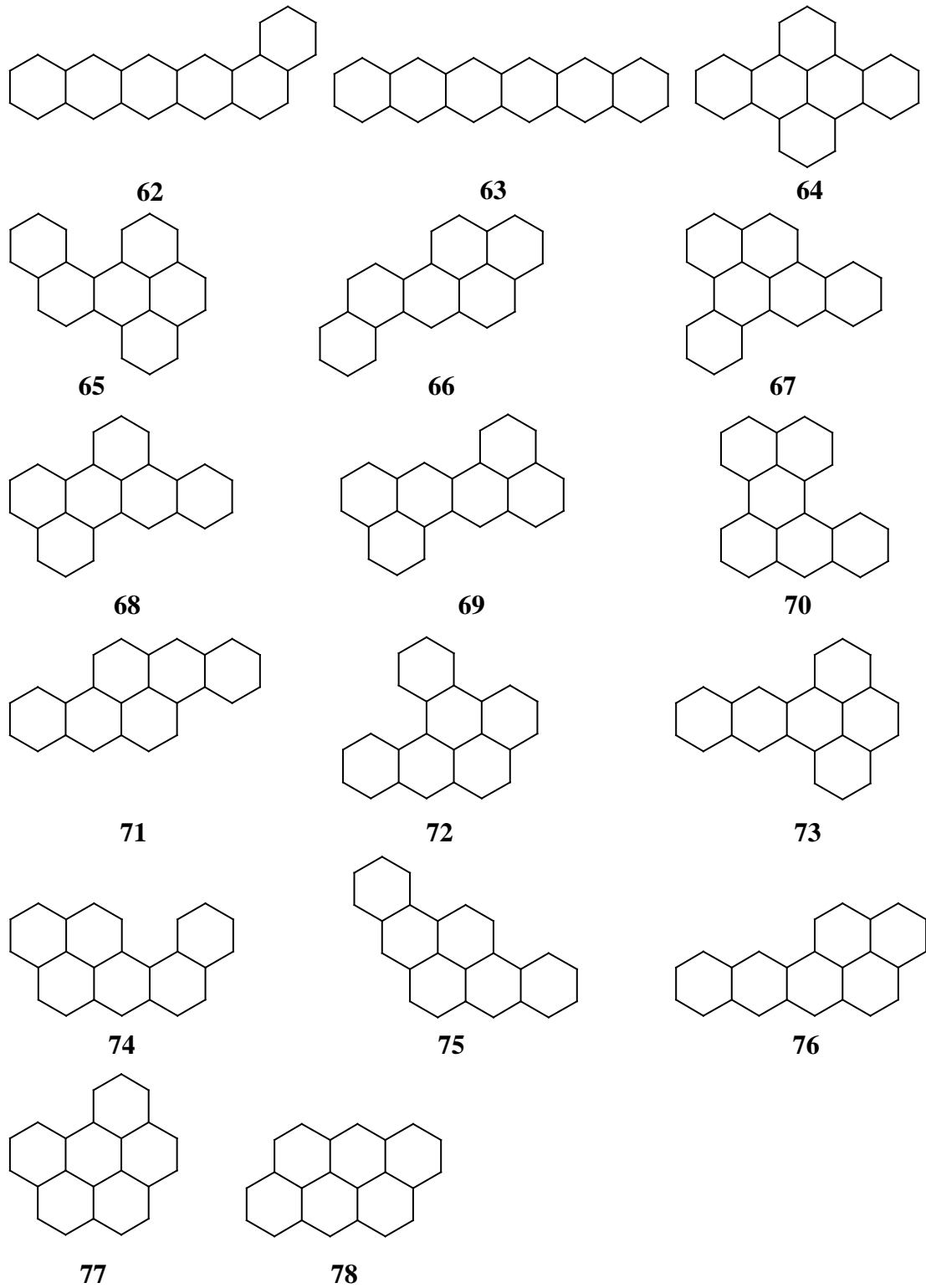


FIGURE 11. (Continued)

## CONCLUSIONS

The scope of this investigation was to compare the extent of similarity between five intrinsic graph distance functions and the Euclidean distance for several sets of molecular graphs. The comparisons to Euclidean distances of course depends upon the graph embedding, and our chemically plausible selection considered only cases where the conformation can be unambiguously determined: the embedding of a regular polygon into a circle, linear and zigzag benzenoid hydrocarbons with up to 20 benzenoid rings, and all 76 benzenoids with up to six rings. Due to the steric interactions between atoms situated in bay regions, some of the benzenoids from the last set are not planar, as indicated also by the molecular mechanics computations; however, they are rigid, and their three dimensional geometry can be uniquely characterized.

Notably the square-root distance functions  $d_{1/2}$  and  $\Omega_{1/2}$  are commonly the least closely in agreement with the Euclidean distances. Of the remaining three,  $d$ ,  $\Omega$ , and  $\rho_{qe}$ , one or the other may be better in different circumstances. Particularly the favorable comparison for  $d$  is in consonance with an earlier suggestion [42]. Quite often it is the shortest-path graph distance  $d$  which is most similar to the Euclidean distance, followed by the resistance distance and the quasi-Euclidean distance. These results have a significant importance for QSAR and QSPR models that use distance-based structural descriptors, indicating that the shortest-path distance  $d$  or equally well the resistance distance  $\Omega$  can substitute the Euclidean distance in computing structural indices.

## ACKNOWLEDGMENT

The authors acknowledge the financial support of this research by the Welch Foundation of Houston, Texas.

## REFERENCES AND NOTES

- [1] 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.
- [2] 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, The Netherlands, 1999, pp. 59-167.
- [3] 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, The Netherlands, 1999, pp. 169-220.
- [4] 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, The Netherlands, 1999, pp. 221-277.
- [5] M. V. Diudea and I. Gutman, *Croat. Chem. Acta*, **1998**, *71*, 21-51.
- [6] O. Ivanciuc, S. L. Taraviras, and D. Cabrol-Bass, *J. Chem. Inf. Comput. Sci.*, **2000**, *40*, 126-134.
- [7] S. Taraviras, O. Ivanciuc, and D. Cabrol-Bass, *J. Chem. Inf. Comput. Sci.*, **2000**, *40*, 1128-1146.
- [8] H. Wiener, *J. Am. Chem. Soc.*, **1947**, *69*, 17-20.
- [9] H. Wiener, *J. Am. Chem. Soc.*, **1947**, *69*, 2636-2638.
- [10] O. Ivanciuc, *J. Chem. Inf. Comput. Sci.*, **2000**, *40*, 1412-1422.
- [11] O. Ivanciuc, T. Ivanciuc, and M. V. Diudea, *SAR QSAR Environ. Res.*, **1997**, *7*, 63-87.
- [12] D. J. Klein, *J. Math. Chem.*, **1995**, *18*, 321-348.
- [13] D. J. Klein, I. Lukovits, and I. Gutman, *J. Chem. Inf. Comput. Sci.*, **1995**, *35*, 50-52.
- [14] H.-Y. Zhu, D. J. Klein, and I. Lukovits, *J. Chem. Inf. Comput. Sci.*, **1996**, *36*, 420-428.
- [15] D. J. Klein, *Commun. Math. Comput. Chem. (MATCH)*, **1997**, *35*, 7-27.
- [16] D. J. Klein and I. Gutman, *J. Chem. Inf. Comput. Sci.*, **1999**, *39*, 534-536.
- [17] O. Ivanciuc, *Rev. Roum. Chim.*, **1989**, *34*, 1361-1368.
- [18] T. S. Balaban, P. A. Filip, and O. Ivanciuc, *J. Math. Chem.*, **1992**, *11*, 79-105.
- [19] O. Ivanciuc, T.-S. Balaban, and A. T. Balaban, *J. Math. Chem.*, **1993**, *12*, 309-318.

- [20] M. V. Diudea, O. Ivanciuc, S. Nikolić, and N. Trinajstić, *Commun. Math. Comput. Chem. (MATCH)*, **1997**, 35, 41-64.
- [21] O. Ivanciuc, T. Ivanciuc, and A. T. Balaban, *J. Chem. Inf. Comput. Sci.*, **1998**, 38, 395-401.
- [22] O. Ivanciuc, M. V. Diudea, and P. V. Khadikar, *Ind. J. Chem.*, **1998**, 37A, 574-585.
- [23] D. J. Klein and M. Randić, *J. Math. Chem.*, **1993**, 12, 81-95.
- [24] H.-Y. Zhu and D. J. Klein, *J. Chem. Inf. Comput. Sci.*, **1996**, 36, 1067-1075.
- [25] D. J. Klein and H.-Y. Zhu, *J. Math. Chem.*, **1998**, 23, 179-195.
- [26] D. Bonchev, A. T. Balaban, X. Liu, and D. J. Klein, *Int. J. Quantum Chem.*, **1994**, 50, 1-20.
- [27] A. T. Balaban, X. Liu, D. J. Klein, D. Babić, T. G. Schmalz, W. A. Seitz, and M. Randić, *J. Chem. Inf. Comput. Sci.*, **1995**, 35, 396-404.
- [28] O. Ivanciuc and A. T. Balaban, *Commun. Math. Comput. Chem. (MATCH)*, **1994**, 30, 141-152.
- [29] O. Ivanciuc, *Rev. Roum. Chim.*, **1999**, 44, 519-528.
- [30] O. Ivanciuc, T. Ivanciuc, and A. T. Balaban, *A C H - Model. Chem.*, **2000**, 137, 57-82.
- [31] A. T. Balaban, D. Mills, O. Ivanciuc, and S. C. Basak, *Croat. Chem. Acta*, **2000**, 73, 923-941.
- [32] M. Randić, *Int. J. Quantum Chem.: Quantum Biol. Symp.*, **1988**, 15, 201-208.
- [33] B. Bogdanov, S. Nikolić, and N. Trinajstić, *J. Math. Chem.*, **1989**, 3, 299-309.
- [34] S. Nikolić, N. Trinajstić, Z. Mihalić, and S. Carter, *Chem. Phys. Lett.*, **1991**, 179, 21-28.
- [35] Z. Mihalić and N. Trinajstić, *J. Mol. Struct. (Theochem)*, **1991**, 232, 65-78.
- [36] M. V. Diudea, D. Horvath, and A. Graovac, *J. Chem. Inf. Comput. Sci.*, **1995**, 35, 129-135.
- [37] M. Randić, *New J. Chem.*, **1995**, 19, 781-791.
- [38] O. Ivanciuc and A. T. Balaban, *Rev. Roum. Chim.*, **1999**, 44, 539-547.
- [39] L. Bytautas, D. J. Klein, M. Randić, and T. Pisanski, *DIMACS Ser. Discr. Math. Theor. Comput. Sci.*, **2000**, 51, 39-61.
- [40] HyperChem 5, Hypercube, Inc., Florida Science and Technology Park, 1115 N.W. 4th Street Gainesville, Florida 32601, U.S.A., [www http://www.hyper.com](http://www.hyper.com).
- [41] Z. Mihalić and N. Trinajstić, *J. Chem. Educ.*, **1992**, 69, 701-712.
- [42] D. J. Klein, *Int. J. Quantum Chem. Quantum Chem. Symp.*, **1986**, 20, 153-171.