

# On Characterization of Cyclic Structures

Milan Randić<sup>†</sup>

National Institute of Chemistry, P.O. Box 3430 Ljubljana, Slovenia

Received July 22, 1997<sup>®</sup>

We consider a characterization of cyclic structures based on the information from graph distance matrix, **D**, and detour matrix, **DD**, of a graph. Using the elements of **D** and **DD** matrices we have constructed a new matrix **D/DD**, the elements of which are given as a quotient of the corresponding elements of **D** and **DD** matrices. The sum of all entries of a matrix above the diagonal of the **D/DD** matrix shows a regular variation for graphs of the same size with the increase of its cyclic structure. Using the cycle graph  $C_n$  and the complete graph  $K_n$  as the extreme cases of cyclic structures we propose an index  $\gamma$  as the measure of molecular cyclic structure. The approach is extended to characterization of molecular local features, such as individual rings embedded in different local environment of a molecule.

## INTRODUCTION

Molecular branching and molecular cyclic structure or cyclicity are the two most visible structural elements that widely vary among molecules. There have been attempts to numerically characterize these molecular skeletal features, molecular branching in particular. The motivation for these efforts is the belief that molecular branching parallels a number of molecular properties. For example, the more branched alkanes have a smaller boiling point, while the less branched isomers have larger boiling points. Wiener in 1947 introduced a molecular descriptor, today known as the Wiener number, as one of the descriptors to quantitatively describe a structure–property relationship.<sup>1</sup> The Wiener number is one of the earliest graph theoretical descriptor used in structure–property studies, even if not recognized as such at its introduction. It parallels molecular branching. Its structural interpretation is somewhat elusive, although Platt<sup>2</sup> suggested already in the late 1940s that the Wiener number relates to molecular volume.

Lovasz and Pelikan suggested the leading eigenvalue of the adjacency matrix as an index of molecular branching.<sup>3</sup> The connectivity index,<sup>4</sup> a bond additive quantity as well as the Hosoya Z number,<sup>5</sup> that counts the way how nonadjacent bonds are distributed throughout molecular skeleton, also reflects in a different way variations of molecular branching. The path matrix, in which the elements are the path subgraphs, recently introduced<sup>6</sup> also leads to novel invariants of interest for the characterization of molecules. If the elements in the path matrix are replaced by the leading eigenvalues of the corresponding paths one obtains (a numerical) matrix. The leading eigenvalue of the so constructed path matrix is even a better index of the molecular branching than the leading eigenvalue of adjacency matrix.<sup>7</sup>

Once we find a satisfactory, mathematically well defined, descriptor of molecular branching then it can be used to rigorously define molecular branching. This is our philosophy in approaching the task of defining molecular cyclicity. First we will search for a suitable structural descriptor which will be sensitive to variations in the mode of fusion of rings,

their relative size and orientation, or details of their interconnection. Then we will use the index to define molecular cyclicity.

In contrast to molecular branching there have been few attempts to characterize numerically the cyclic structure of a molecule. The problem was considered by Bonchev, Mekenyan, and Trinajstić<sup>8</sup> and later by Bonchev, Balaban, Liu, and Klein.<sup>9</sup> They tried to formulate the rules that give the degree of cyclicity among structurally related molecules. They considered the role of the number of atoms, the number of cycles, the number of atoms in a cycle, the number of cycles having a common edge, the mutual position of cycles, etc., on the overall cyclicity. Initially these authors were guided by the variation of the Wiener number with molecular skeletal features. Also the HMO  $\pi$ -electron energies for cyclic structures corresponding to polycyclic conjugated hydrocarbons were considered.

The first difficulty that such an approach entails is the fact that the Wiener number as initially proposed by Wiener is defined only for trees (acyclic structures). Extension of the concept of the Wiener number to cyclic structures is nontrivial in the sense that more than one generalization are not only possible but plausible. One way to generalize the Wiener number for cyclic structures is first to construct the distance matrix of the graph and then sum the elements of the distance matrix above the main diagonal. Such a procedure corresponds to adding all topological distances within a structure. When applied to trees it gives the Wiener number for a tree.<sup>5</sup>

The sum of the topological distances, i.e., the Wiener number, decreases for molecules having more complex cyclic structure or cyclicity in comparison to molecules of the same size and fewer rings. This is to be expected, since in the former case there are many shorter path connecting pairs of vertices, while in the latter there are fewer such shorter paths. The rules of Bonchev et al. have been more recently revised and generalized. The rules based on the Wiener number do not reflect adequately the presence of multiple shorter paths between a pair of vertices. This information can be absorbed by considering instead of the distance matrix other graph matrices. For example, the resistance–distance matrix,<sup>10</sup> in which the metric is based on Kirchhoff's laws for electric networks (assuming uniform edge resistance of 1 ohm),

<sup>†</sup> Visiting scientist from Dept. of Mathematics & Computer Science, Drake University, Des Moines, IA 50311.

<sup>®</sup> Abstract published in *Advance ACS Abstracts*, November 1, 1997.

represents such an alternative route. The work of Bonchev and collaborators implied that the Wiener number,  $W$ , and indirectly the distance matrix from which  $W$  is extracted, is the fundamental quantity that faithfully describes molecular cyclic structure or cyclicity.

While the approach of Bonchev et al. was successful in the classification and comparison of numerous cyclic structures, it is accompanied with some limitations. First, there are structures that have identical Wiener numbers yet show distinct cyclic structure. Such cases require human intervention, which, even when acceptable, does reflect a degree of arbitrariness. An alternative generalization of the Wiener index from acyclic structures to cycle-containing structures may give fewer degeneracies. One may expect that different generalizations of the Wiener number for cyclic structures may produce different numerical results, thus still leaving the precise measure and the meaning of cyclicity unsettled.

Because the graph theoretical distance matrix does not reflect fully the cyclic structure of a molecule we should re-examine alternative routes to characterization of molecular cyclic features. Besides the information on the length of the shortest path between two vertices we need to use the information on the multiplicity of the shortest connections. Multiplicity increases in general with an increased cyclic complexity so it is indirectly taken into account by the distance matrix which is the reason that  $\mathbf{D}$  matrix has lead to satisfactory results. The reason for searching for an alternative is therefore not because the distance matrix is inherently inferior, but because it is associated with a relatively high degeneracy for a number of its invariants. That is, there are many pairs of graphs (often also triples and higher number of multiples) that have the same value for the structural invariant of interest. In that respect the resistance–distance matrix<sup>9</sup> has an advantage. If we compare the Wiener numbers (i.e., the Wiener sum of the  $\mathbf{D}$  matrix) and the Wiener sums of the resistance–distance matrices (i.e., the numbers obtained by adding all entries in the matrix above the diagonal) for many molecules we find fewer coincidences for this invariant in different structures.

#### NOVEL MATRIX FOR CYCLIC STRUCTURES

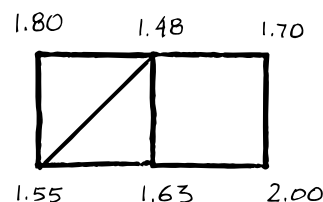
In this paper we report on construction of a novel matrix associated with cyclic graphs which appears to be capable of faithfully characterizing polycyclic systems. The basic structural invariant of the novel matrices that we will consider is the sum of the entries above the main diagonal. We will refer to this number as the Wiener sum of the matrix rather than the Wiener number of the matrix to avoid a confusion with the well-known Wiener number, proposed by Wiener, which relates to the distance matrix of a graph. In this terminology the Wiener number becomes the Wiener sum of the graph distance matrix  $\mathbf{D}$ . The distance matrix  $\mathbf{D}$  was introduced for the first time in graph theory by Harary.<sup>11</sup> Before, however, we outline the construction of the novel matrix we briefly need to describe the detour matrix  $\mathbf{DD}$  also introduced in graph theory by Harary.<sup>11</sup> The metric on graphs, however, dates back to Cayley.

#### DETOUR MATRIX $\mathbf{DD}$

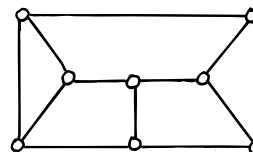
The distance matrix found considerable application in chemistry. In particular it was recognized that it represents a convenient route to the Wiener number of graphs. The

**Table 1.** The Distance Matrix and the Detour Matrix for the Tricyclic Graph of Figure 1 and the Corresponding Row Sums

						row sum
$\mathbf{D}$						
0	1	2	3	2	1	9
1	0	1	2	1	1	6
2	1	0	1	2	2	8
3	2	1	0	1	2	9
2	1	2	1	0	1	7
1	1	2	2	1	0	7
$\mathbf{DD}$						
0	5	5	5	5	5	25
5	0	5	4	3	4	21
5	5	0	5	4	5	24
5	4	5	0	5	4	23
5	3	4	5	0	5	22
5	4	5	4	5	0	23



**Figure 1.** Tricyclic system with all symmetry nonequivalent vertices and the corresponding values for the row sums of  $\mathbf{D}/\mathbf{DD}$  matrix.



**Figure 2.** The polycyclic graph in which the longest path (the detour path) is maximal for any pair of vertices.

other matrix of Harary,<sup>11</sup> the so called detour matrix, however, until very recently was practically overlooked in applied graph theory, including applied graph theory in chemistry. We will use the label  $\mathbf{DD}$  for the detour matrix that is suggestive of close connection with the distance matrix. The element  $(i, j)$  of the  $\mathbf{DD}$  distance is defined by the length of the longest path between the two vertices. In Table 1 we have listed the  $\mathbf{D}$  matrix and the  $\mathbf{DD}$  matrix for the small tricyclic graph of Figure 1.

The detour matrix  $\mathbf{DD}$  was resurrected by Trinajstić and co-workers who introduced a novel graph invariant, the “detour index”.<sup>12</sup> The detour index is the Wiener sum of the detour matrix (i.e., the sum of the elements above the main diagonal of the detour matrix). Properties of the detour index were recently investigated.<sup>12–18</sup> It is not a surprise that graphs were found that have the same detour index<sup>12,13</sup> or were constructed.<sup>18</sup> Use of the detour index in QSAR and structure–property relationship studies is still under exploration. Regardless of the use of the detour index in structure–property applications the detour matrix does deserve attention. Very recently it was shown not only that different graphs may have the same detour index but also that different graphs can have the same detour matrix if the vertices in two graphs are suitably labeled.<sup>18</sup> In Figure 2 we illustrate a graph with the “maximal” detour matrix among the graphs of the same size. The longest path between any pair of vertices of the graph of Figure 2 is the same, equal to the maximal longest path. Hence, if we construct a new graph by adding another edge the detour

**Table 2.** The **D/DD** Matrix for the Tricyclic Graph of Figure 1 and the Corresponding Row Sums

D/DD						row sum
0	1/5	2/5	3/5	2/5	1/5	1.800 00
1/5	0	1/5	2/4	1/3	1/4	1.483 33
2/5	1/5	0	1/5	2/4	2/5	1.700 00
3/5	2/4	1/5	0	1/5	2/4	2.000 00
2/5	1/3	2/4	1/5	0	1/5	1.633 33
1/5	1/4	2/5	2/4	1/5	0	1.550 00

matrix of the new graph will not change the detour matrix at all, since we already have the longest paths between any pair of points. Just as **D** matrix does not record the presence of multiple shortest paths so the **DD** matrix does not record the presence of multiple longest paths. However, that different graphs are associated with the same matrix is a novelty. None of a dozen recently introduced novel graph matrices<sup>19–27</sup> show this peculiar property. The finding is not necessarily surprising, since very dense graphs often will have enough multiple connections to ensure that the longest path between any pair of vertices is of the maximal length  $N - 1$ .<sup>28</sup> If an edge is added to such a graph, one does not create longer paths, and the **DD** matrix is the same, though the two graphs are different.

It is clear that the **DD** matrix also reflects the cyclic structure of a graph just as the **D** matrix does but in a different way. If there are many cycles in a structure the longest path (that represent a detour) can “navigate” around and depending on the positioning of fused rings will lead to distinct longest paths for different molecules even when they have similar overall appearance. It seems therefore that the nondimensional **D/DD** matrix, in which the entries are given as the ratio of  $d_{ij}$  and  $dd_{ij}$ , these being the  $i, j$  elements of the **D** and **DD** matrix, respectively, may offer novel characterizations of molecular cyclic complexity.

#### D/DD QUOTIENT MATRIX OF A GRAPH

In Table 2 we show the **D/DD** matrix of the tricycle graph of Figure 1. The numerators of the fractions shown are the elements of the **D** matrix, and the denominators are the corresponding elements of the **DD** matrix of Table 1. For smaller graphs the construction of the distance and the detour matrices is not involved. However, the construction of the **DD** matrix for larger molecules is inherently difficult and presents an NP complete problem, as pointed out by Lukovits and Razinger.<sup>14</sup> The molecules considered in this paper are all relatively small, or sufficiently simple, so that finding the elements of **D** and **DD** matrices did not present difficulties.

In order to arrive at a characterization of molecular graphs we first have to select a set of matrix or graph invariants. If the selected invariants are ordered, they can be viewed as components of a structure-vector. In this contribution we will consider the Wiener sum of **D/DD** matrices. The Wiener sums are of particular interest as the basic structure invariants primarily because there may be some parallelism with the earlier work of Bonchev and co-workers who based their considerations of molecular cyclicity on the Wiener index (i.e., the Wiener sum derived from the **D** matrix of a graph). Other choices of invariants of the **D/DD** matrix are possible. In particular the properties of the leading eigenvalue of **D/DD** matrix deserves attention in view of the use of the leading

eigenvalue of the path matrices for characterization of molecular branching. The leading eigenvalue of a (symmetric) matrix is bounded from above and from below by the largest and the smallest row sums. Hence, the row sums of the **D/DD** matrix are of interest. In the case of the graph of Figure 1 we have as the bounds on the leading eigenvalue,  $1.48333 < \gamma < 1.8$ , where  $\lambda$  is the largest eigenvalue of the **D/DD** matrix.

#### LOCAL DESCRIPTORS—RING INDICES

The row sums represent local graph invariant and can be used as local descriptors for individual vertices. For graph of Figure 1 all vertices are symmetry nonequivalent resulting in distinct values for the row sums belonging to different vertices. We see that the branching vertices have smaller row sums than the bridging vertices (of degree two). The smallest row sum (1.48) goes with the vertex of degree 4; the next in the magnitude are the rows corresponding to vertices of degree 3 (1.63 and 1.55). The bridging vertices have the largest row sums (1.80, 1.70, and 2.00). This could have been expected because the longest path in the case of branching sites is restricted to cycles on one or the other side of the branching vertex, while the vertices at the bridge have greater latitudes in “navigating” over the molecular skeleton. By summing the contributions (row sums) from the vertices that constitute a ring or a cycle one can arrive at individual ring indices and cycle indices. The ring indices and cycle indices for the graph of Figure 1 are as follows:

cycle defines by vertices:	cycle index:	per vertex:
1, 2, 6	4.833 33	1.611 11
2, 5, 6	4.666 67	1.555 56
2, 3, 4, 5	6.816 67	1.704 15
2, 3, 4, 5, 6	8.366 67	1.673 33
all	10.166 67	1.694 44

The last entry that incorporates all vertices is twice the Wiener sum (which is 5.083 33).

#### CHARACTERIZATION OF $C_N$ AND $K_N$

In this section we will examine the **D/DD** matrices of simple cyclic graphs  $C_n$  and the complete graphs  $K_n$ . We use the word cyclicity to signify cyclic structure of graphs considered. In the following sections we will consider polycyclic graphs having five and six vertices. We will consider the derived Wiener sums of graphs of different cyclic structure and will define cyclicity as a measure of the departure of molecular cyclic structure from that of  $C_n$ . We end with a description of cyclicity in larger polycyclic systems illustrated by a smaller benzenoid where we will look also in the characterization of local structural features.

Monocyclic  $C_n$  graphs are the simplest cyclic structures. The **D/DD** matrix of  $C_n$  is shown in Table 3. Every row of the **D/DD** matrix is identical (thus giving the leading eigenvalue of the matrix exactly) and of a simple structure. It is convenient here to consider separately odd and even member cycles because in the former each entry in each row appears twice, while in the latter the entry for the diametrically opposite vertex appears only once. In Table 4 we listed the Wiener sum and the row sums (which is identical to the leading eigenvalue because all rows of the matrix have the same row sum) for all the cycles  $C_3$ – $C_{12}$ . If

**Table 3.** The **D/DD** Matrix for Cyclic Graph  $C_n$  and  $K_n$ 

<b>D/DD</b> for $C_n$								
0	$1/(n-1)$	$2/(n-2)$	$3/(n-3)$	$4/(n-4)$	...	$3/(n-3)$	$2/(n-2)$	$1/(n-1)$
$1/(n-1)$	0	$1/(n-1)$	$2/(n-2)$	$3/(n-3)$	...	$4/(n-4)$	$3/(n-3)$	$2/(n-2)$
$2/(n-2)$	$1/(n-1)$	0	$1/(n-1)$	$2/(n-2)$	...	$5/(n-5)$	$4/(n-4)$	$3/(n-3)$
$3/(n-3)$	$2/(n-2)$	$1/(n-1)$	0	$1/(n-1)$	...	$6/(n-6)$	$5/(n-5)$	$4/(n-4)$
...	...	...	...	...	...	...	...	...
<b>D/DD</b> for $K_n$								
0	$1/(n-1)$	$1/(n-1)$	$1/(n-1)$	$1/(n-1)$	...	$1/(n-1)$	$1/(n-1)$	$1/(n-1)$
$1/(n-1)$	0	$1/(n-1)$	$1/(n-1)$	$1/(n-1)$	...	$1/(n-1)$	$1/(n-1)$	$1/(n-1)$
$1/(n-1)$	$1/(n-1)$	0	$1/(n-1)$	$1/(n-1)$	...	$1/(n-1)$	$1/(n-1)$	$1/(n-1)$
$1/(n-1)$	$1/(n-1)$	$1/(n-1)$	0	$1/(n-1)$	...	$1/(n-1)$	$1/(n-1)$	$1/(n-1)$
...	...	...	...	...	...	...	...	...

**Table 4.** The Wiener Sum and the Row Sums for **D/DD** Matrix of Cyclic Graphs

$n$	Wiener sum	row sum	$n$	Wiener sum	row sum
3	1.500 000	1.000 000	8	12.609 52	3.152 381
4	3.333 333	1.666 667	9	15.396 43	3.421 429
5	4.583 333	1.833 333	10	19.563 492	3.912 698
6	7.200 000	2.400 000	11	23.121 825	4.203 968
7	9.216 667	2.633 333	12	28.062 338	4.677 056

we consider separately cycles of odd and even ring size we can observe an increase of  $\lambda$  as  $n$  increases with an oscillation in the rate of increase of reflecting the odd–even character of the cyclic structures.

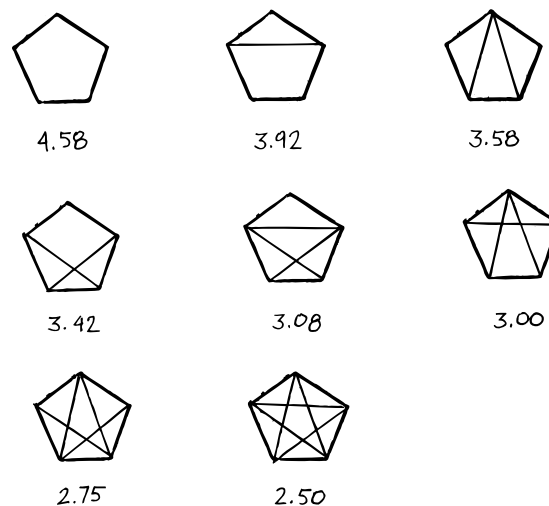
The complete graphs  $K_n$  can be viewed as the most cyclic structure by mere count of the number cycles such graphs contain. However, since every vertex in a graph on  $n$  vertices has degree  $(n-1)$  and the longest path is of length  $(n-1)$ , the **D/DD** matrix of a graphs  $K_n$  possesses simple structure. The **D/DD** matrix of  $K_n$  is shown in Table 3 (bottom). Again every row of the **D/DD** matrix is identical (thus giving the leading eigenvalue of the matrix exactly). The matrix has a particularly simple structure since all the entries of the matrix are the same and the matrix differs from the adjacency matrix of  $K_n$  by a factor of  $1/(n-1)$ . The Wiener sum for **D/DD** matrix of complete graph  $K_n$  equals  $n$  since row sums are independent of  $n$  and equal 1.

The row sums of  $C_n$  and  $K_n$  represent also the limits for the row sums for all cyclic graphs that have no cut points, i.e., graphs which are derived from  $C_n$  by adding additional edges. For a given  $n$  no cyclic graph that has no cut points can have a smaller row sum than the monocyclic graph  $C_n$  nor greater row sum than the complete graph  $K_n$ . Cyclic graphs having spiro-vertex or bridging edges between cyclic components (i.e., cut points) are excluded from our considerations. The row sum for **D/DD** matrix of graphs having a spiro vertex, like the graph of spiropentane, will exceed the row sum of cyclic graphs without such cut points.

#### CYCLOCITY OF SMALLER GRAPHS

The carbon skeleton of bicyclo[1.1.0]pentane represents the smallest polycyclic graph. Its Wiener sum is 2.5000, while the row sums are 1.16667 and 1.3333 for the trivalent and the divalent vertex, respectively. The next in size are the cyclic graphs on five vertices shown in Figure 3. We have excluded the graph of spiropentane which we will consider in a separate report concerned with compounds having spiro-vertices (e.g., the so called cacti). In Table 5 we list the Wiener sum, the average row sum, and the individual row sums, for the graphs of Figure 5.

The cyclic character, defined as the Wiener sum of **D/DD** matrix, increases as we go from  $C_5$  to  $K_5$ . From Table 5 we

**Figure 3.** All cyclic graphs without cut points on five vertices. The corresponding values for the row sums of **D/DD** matrix are shown under each graph.**Table 5.** The Wiener Numbers (WN), the Wiener Sums (WS), the Average Row Sum, the Cyclicity  $\gamma$ , and the Number of Spanning Trees for Graphs of Figure 3

graph/vertex	WN	WS	numeric	av row sum	cyclicity	spanning trees
1	15	55/12	4.583 33	1.833 33	0	5
2	14	47/12	3.916 67	1.566 67	0.32	11
3	13	43/12	3.583 33	1.433 33	0.48	21
4	13	41/12	3.416 67	1.366 67	0.56	24
5	12	37/12	3.083 33	1.233 33	0.72	40
6	12	36/12	3.000 00	1.200 00	0.76	45
7	11	33/12	2.750 00	1.100 00	0.88	75
8	10	30/12	2.500 00	1.000 00	1.00	125

see that the Wiener sums gradually decrease from the value 4.5833 for  $C_5$  to 2.5000 for  $K_5$ . The Wiener number  $W$  for the same graphs parallels the trend by decreasing from the value  $W = 15$  for  $C_5$  to  $W = 10$  for  $K_5$ . Observe that among the eight graphs of Figure 5 several graphs have the same Wiener number (13 and 12, respectively). In contrast the Wiener sum for **D/DD** matrices for all the eight graphs of Figure 5 are distinct. Clearly, the Wiener sum of **D/DD** matrices is more suitable as the basis for a quantification of the molecular cyclicity than the Wiener number, because it requires less (if any) human intervention in order to resolve the cases of the degeneracy for the structures having the same Wiener sum. It is of interest also to notice that the relative magnitudes of Wiener sums parallel the relative magnitudes for the number of spanning trees, which are shown in the last column of Table 5 (taken from ref 29).

The molecular cyclicity is generally an intuitively felt concept that has not yet been rigorously defined. Molecular

cyclicity and complexity of cyclic graphs though conceptually somewhat different are often viewed interchangeably. One measure of the degree of graph complexity suggested in the literature was based on the number of spanning trees of a graph.<sup>29</sup> One can find the number of spanning trees of a graph by considering the Laplacian, i.e., the matrix in which the diagonal has the negative value of the row sum of the corresponding row of the adjacency matrix.<sup>29</sup> The determinant of the Laplacian with one row and the corresponding column deleted then give the number of spanning trees. Hence, the computation of spanning trees is straightforward, but as we see their number can vary considerably, doubling or almost doubling the value as additional edges are inserted in  $C_5$ .

We think that the Wiener sum of **D/DD** matrix has advantages over the Wiener number  $W$  and over the count of spanning trees as a measure of cyclicity/complexity of cyclic graphs. It shows limited degeneracy, as mentioned above, and it does not grow exponentially with the number of cycles. Instead of using directly the Wiener sum (or its reciprocal in view of the decrease of the sum with increase of cyclicity) we can consider the following equation as a definition of cyclicity:

$$\gamma(G_n) = \{\text{WS}(C_n) - \text{WS}(G_n)\} / \{\text{WS}(C_n) - \text{WS}(K_n)\}$$

Here  $\text{WS}$  is Wiener sum of **D/DD** matrices for the listed graphs ( $G$ ,  $C$ , and  $K$ ) and  $\gamma(G_n)$  represents the cyclic complexity (or cyclicity) of the polycyclic graph  $G_n$  having  $n$  vertices. Hence we have the following:

**Definition:** Cyclicity of a graph  $G$  is given by the departure of the cyclic character of the graph from that of the cycle relative to the departure of the complete graph  $K_n$  from the cycle graph  $C_n$ . Cyclic character is given by the sum of all the elements above the diagonal in the **D/DD** matrix of a graph.

Using this definition the cyclic complexity for graphs without cut points or bridging edges is between zero and one. The value zero belongs to  $C_n$  and the value one belongs to  $K_n$ . For the eight graphs of Table 5 we obtain the following values:

0, 8/25, 12/25, 14/25, 18/25, 19/25, 22/25, 1

or in decimal form

0, 0.32, 0.48, 0.56, 0.72, 0.76, 0.88, 1.00

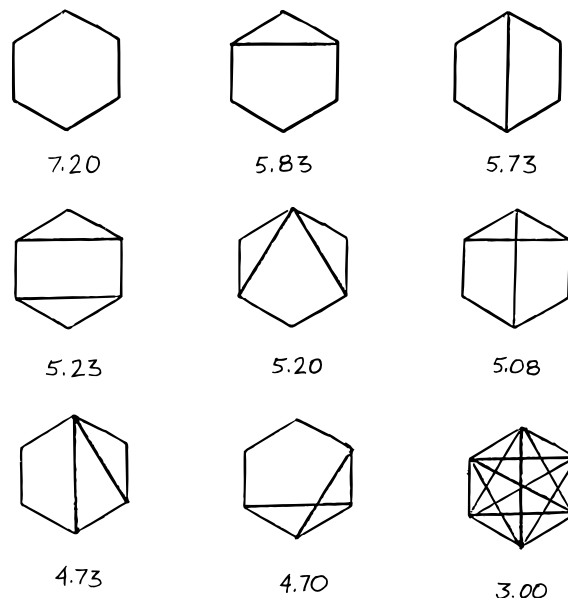
Observe that the index of cyclicity  $\gamma$  is size-independent. This makes it possible to directly compare cyclic structures of different size. If two cyclic graphs show the same cyclicity that means that they have the same relative "distance" from the corresponding extreme standards  $C_n$  and  $K_n$ , although  $n$  will be different for graphs of different size. We believe that the attribute "cyclicity", and this also extends to "branching", should be size independent, since there is already a very simple index for molecular size,  $n$ , the number of vertices.

#### ILLUSTRATIONS

In Table 6 we list the Wiener sum, the average row sum, and the cyclicity index  $\gamma$  for monocyclic, bicyclic, and tricyclic graphs on six vertices. Again we have excluded endocyclic spiro compounds and structures with bridges. The

**Table 6.** The Wiener Sums (WS), the Average Row Sum (RS), and the Cyclicity  $\gamma$ , for Graphs of Figure 4

graph/vertex	WS	RS	$\gamma$
1	7.200 00	2.4	0
2	5.833 33	1.944 444	0.325 40
3	5.733 33	1.911 11	0.349 21
4	5.233 33	1.744 44	0.468 25
5	5.2	1.733 33	0.476 19
6	5.083 33	1.744 44	0.503 97
7	4.733 33	1.577 78	0.587 30
8	4.700 00	1.566 67	0.595 24
$K_6$	3.000 00	1.000 00	1.000 00



**Figure 4.** Cyclic graphs without cut points on six vertices investigated. The corresponding values for the row sums of **D/DD** matrix are shown under each graph.

eight graphs are illustrated in Figure 4. The last row in Table 6 gives the corresponding data for the complete graph  $K_6$ . The graphs have been ordered by increasing cyclicity.

Graphs in the first row of Figure 4 illustrate the rule 8 of ref 8, which still holds but has to be rephrased in terms of the Wiener sum, rather than the Wiener number:

**Modified Rule 8:** The Wiener sum in the molecular polycyclic graph increases, i.e., the molecular cyclicity decreases, with the increase of the difference in the number of atoms within two fused cycles

$$\text{WS}_{n,n} < \text{WS}_{n+1,n-1} < \text{WS}_{n+2,n-2} < \dots < \text{WS}_{n+k,n-k}$$

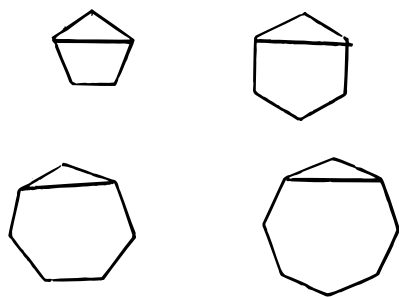
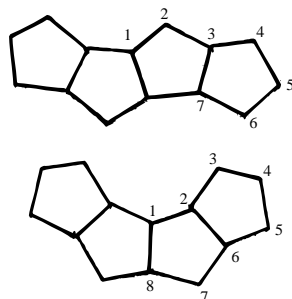
where  $k$  is an integer.

#### COMPARISON OF GRAPHS OF DIFFERENT SIZE

In Table 7 we have listed the cyclicity index  $\gamma$  for the bicyclic graphs of Figure 5 which are obtained from  $C_n$  by introducing an edge between a pair of the vertices separated by two edges. The graphs correspond to those of bicyclo[1.1.0]pentane, bicyclo[2.1.0]hexane, bicyclo[3.1.0]heptane, etc. We have listed separately the  $\gamma$  for graphs having an odd and an even number of vertices because of the already noticed oscillatory behavior of  $\gamma$  for cyclic graphs  $C_n$ . First we see that  $\gamma$  decreases with  $n$ , the size of the graphs. This is in agreement with expectations, because as  $n$  increases the single bridging edge represents smaller and smaller "perturbation". For an infinitely large graph we expect

**Table 7.** The Cyclicity and the Cyclicity per Vertex for Graph of Figure 5

$n$	odd	even	per vertex (%)
4		0.625 01	15.63
5	0.320 00		6.40
6		0.325 40	5.42
7	0.198 25		2.83
8		0.210 45	2.63
9	0.142 90		1.59
10		0.153 16	1.53
11	0.111 51		1.01
12		0.119 55	0.996

**Figure 5.** Bicyclo[ $n.1.0$ ]alkane systems illustrating gradual approach of these molecules to the limiting  $C_n$  case.**Figure 6.** Cata-condensed polycyclic systems built from fused pentagons in an “antiparallel” and the “parallel” mode.

therefore  $\gamma = 0$ , which is true for cyclic graphs  $C_n$  (of any size). While  $\gamma$  decreases with  $n$  we see that even member graphs of Figure 5 have somewhat larger cyclicity than odd member graphs. If however we compute the cyclicity per vertex,  $\gamma/n$ , we obtain for the graphs of Figure 5 a monotonous decrease of the per vertex cyclicity with the increase in  $n$ . Again, however, odd and even values appear to be grouped showing a small change from 5 to 6, or 7 to 8, and 9 to 10 but a larger change between  $n = 6$  and  $n = 7$ ,  $n = 8$  and  $n = 9$ , and  $n = 10$  and  $n = 11$ . Per vertex cyclicity  $\gamma/n$  has been expressed as a percentage change of the cyclicity relative to that of  $C_n$  (normalized to the difference  $WS(C_n) - WS(K_n)$ ).

### CATA-CONDENSED SYSTEMS

The cyclic structure of cata-condensed systems, including also cata-condensed benzenoid hydrocarbons, is relatively simple. Nevertheless, when the corresponding molecular graphs are examined, such systems offer diverse topologies. We will first examine the **D/DD** matrices for cata-condensed systems built from fused five-membered rings, shown in Figure 6. They will illustrate the role of the relative orientation of the rings in polycyclic structures built from the rings of the same type and having the same number of rings.

**Table 8.** Row Sums of the Detour Matrix **DD** and the Distance Matrix **D** for the “Antiparallel” and the “Parallel” Fused Rings of Figure 6, Respectively

vertex	<b>DD</b> row sum	<b>D</b> row sum	<b>DD</b> row sum	<b>D</b> row sum
1	2.719 39	29	2.902 76	31
2	2.790 99	32	3.001 36	35
3	3.173 04	35	3.148 80	35
4	3.646 27	44	3.692 31	44
5	4.312 94	49	4.102 33	46
6	3.610 72	42	3.331 59	39
7	2.988 31	34	2.668 96	30
8			2.536 02	27
	Wiener sum	Wiener number	Wiener sum	Wiener number
	23.241 65	265	22.664 73	258

**Table 9.** The Wiener Sums for Benzenoids of Figure 7

molecule	WS	WS/ $n$	$g$	$\gamma/n$ (%)
benzene	7.200 00	2.400 00	0	0
naphthalene	15.192 06	3.038 41	0.300 16	3.002
phenanthrene	25.172 61	3.596 09	0.415 79	2.970
anthracene	25.917 57	3.702 51	0.391 84	2.799
triphenylene	35.410 33	3.934 48	0.509 37	2.830
chrysene	37.098 97	4.122 11	0.478 00	2.656
benzphenanthrene	37.195 36	4.132 82	0.476 16	2.646
tetracene	37.521 15	4.169 01	0.470 16	2.612

In Table 8 we have listed the Wiener sums, the average row sums, and the cyclicities for the two structures of Figure 6 having four fused pentagonal rings. The results agree with the rule 4 of Bonchev, Mekenyan, and Trinajstić which states that “within ribbons composed of equal numbers of fused cycles each one having the same odd numbers of atoms, the sum of the topological distances is greater, and hence the molecular cyclicity is smaller, when the neighboring cycles have the antiparallel orientation”.<sup>8</sup> The parallelism between the conclusions derived from the **D** matrix and those derived from **D/DD** matrix is encouraging, since already Bonchev and collaborators have shown that the results derived from the **D** matrix are plausible. The advantage of the **D/DD** matrix is that it shows fewer degeneracy and hence requires fewer human interventions in formulating the rules for molecular cyclicity. As one can observe the graph depicting four pentagons fused in a “parallel” mode has two non-equivalent vertices having the same row sums derived from **D** matrix but not from the **D/DD** matrix.

### CYCLOCITY IN SMALLER BENZENOIDS

In Figure 7 we have depicted smaller benzenoids having up to four fused benzene rings. Table 9 shows the Wiener row sums derived from **D/DD** matrices along with the selected benzenoids of Figure 7. The Wiener sums (Table 9) for isomers shows a regularity. The linear acenes, e.g., anthracene and tetracene, have larger Wiener sums than the angularity fused counterparts (phenanthrene and benzphenanthrene, respectively). The larger Wiener sum means a smaller “departure” of a graph from the cyclicity of the simple  $C_n$  ring system, hence, the smaller cyclicity, because for each  $n$  the simple cycle  $C_n$  has the largest Wiener sum. Similarly a comparison of chrysene and benzphenanthrene shows that “more linear” chrysene has a greater Wiener sum ( $WS = 37.403\ 70$ ), hence the lesser cyclicity than “more bent” benzphenanthrene ( $WS = 35.959\ 65$ ). Apparently the WS is sensitive to the mode of fusion of rings. Among benzenoids having four fused rings the smallest Wiener sum

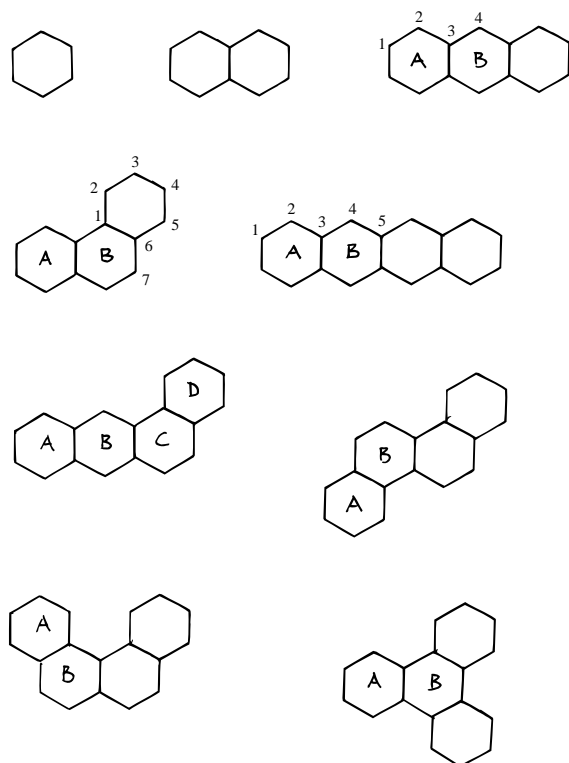


Figure 7. Small cata-condensed benzenoid hydrocarbons.

Table 10. The Wiener Sum of **D/DD** for  $C_n$

$4n+2$	Wiener sum	first difference	second difference
6	7.200 000		
10	19.563492	12.36	
14	38.106 216	18.55	6.19
18	62.829 303	24.71	6.16
22	93.732 978	30.90	6.19
26	130.817 310	37.09	6.09
30	174.082 324	43.26	6.17

is found for the for triphenylene ( $WS = 35.410\ 33$ ) and the largest for tetracene ( $WS = 37.521\ 15$ ).

The cyclicity  $\gamma$  and the cyclicity per vertex  $\gamma/n$  for the considered benzenoids are listed in the last two columns of Table 9. The cyclicity per vertex allows comparison of cyclic systems of different size. It is expressed as the percentage deviation of the cyclicity of a molecule from that of simple  $C_n$ . Thus, for example, the cyclicity of phenanthrene is 0.415 79 and that of triphenylene 0.509 37 pointing to the fact that phenanthrene differs less from the  $C_{14}$  cyclic structure than triphenylene from the  $C_{18}$  ring system. However, when we look at the per vertex cyclicities of triphenylene (2.83%) and phenanthrene (2.97%) we see than

on an absolute scale triphenylene is more similar to the cyclic system  $C_n$  than phenanthrene.

In order to calculate the cyclicity of larger benzenoids one needs to calculate the Wiener sum of **D/DD** matrices for large cyclic systems  $C_n$ . In Table 10 we listed the Wiener sums (WS) for  $4n + 2$  ring systems for  $n = 1$  to  $n = 7$ . We also show the first and the second difference between the successive entries. The WS apparently increases quadratically as is illustrated by the almost constant second difference. In contrast to  $C_n$  graphs the Wiener sum for complete graphs  $K_n$  increases linearly, being given by  $WS_n = n/2$ . Hence, the span  $WS(C_n) - WS(K_n)$ , the denominator of the cyclicity index  $\gamma$ , increases linearly with  $n$ . Although the number of cyclic graphs increases exponentially with the size, the increase in the span  $WS(C_n) - WS(K_n)$  decreases somewhat the chance that two nonisomorphic graphs have the same cyclicity index.

### RING INDICES

The vertex row sums of **D/DD** matrix allow one to construct local structural invariants of cyclic systems corresponding to individual atoms or individual fragments of a molecule. Table 11 shows the row sums for the individual symmetry nonequivalent carbon atoms of the smaller benzenoids of Figure 7. In Table 12 we have listed the characterization for the individual benzene rings for the considered benzenoids. Ring descriptor, each characterizing one ring of a benzenoid system, is obtained by summing local contributions of carbon atoms forming the benzene ring considered. The derived benzene ring indices are descriptors that reflects local geometrical, rather than local electronic environment of the benzenoids. Sometimes it is difficult to separate local geometrical and local electronic contributions in polycyclic conjugated hydrocarbons since both are highly interdependent. However, in our present considerations no electronic factors have been involved, thus we end with descriptors that reflect solely the geometrical features of polycyclic systems. In contrast the local ring indices derived from molecular orbitals, such as the ring indices outlined by Derflinger and Polansky,<sup>32</sup> depend on the electronic and the geometrical factors. The local ring indices derived from the examination of the conjugated circuits model<sup>33–36</sup> are also geometry-dependent although to a lesser degree, since there are molecules, like chrysene and benzphenanthrene, having a 1:1 correspondence between their Kekulé valence structures that produce the same count of conjugated circuits yet correspond to different geometry.

The local ring indices listed in Table 12 show some regularities. In all case the central rings have a smaller index,

Table 11. Row Sums of **D/DD** Matrices for Symmetry Nonequivalent Carbon Atoms of Smaller Benzenoids

	molecule atom					
	anthracene	phenanthrene	tetracene	benzphenanthrene	chrysene	triphenylene
1	3.3971	3.2056	3.5623	2.7605	2.9161	3.2979
2	3.2101	3.2573	3.3668	3.1223	3.3415	3.8812
3	3.5972	4.2306	3.7391	3.8383	4.0608	4.6243
4	4.4528	4.5124	4.1159	4.9206	5.1455	
5		3.7359	5.7576	4.8300	5.5014	
6		3.2220		4.3620	4.4321	
7		3.0086		4.0900	4.1617	
8				4.1574	3.9348	
9				4.6991	3.6049	
10				3.5908		

**Table 12.** Ring Indices for Benzenoids of Figure 7

benzenoid	ring	ring index
benzene		14.400
naphthalene		17.917
anthracene	A	22.520
	B	19.635
phenanthrene	A	22.164
	B	18.873
tetracene	A	27.225
	B	21.336
benzanthracene	A	28.056
	B	21.496
	C	22.536
	D	27.035
chrysene	A	26.643
	B	21.647
benzphenanthrene	A	24.697
	B	21.668
triphenylene	A	23.607
	B	19.787

while the terminal rings show the largest ring index values. The magnitude to some extent depends on the size of the benzenoid or, more exactly, on the number of fused rings in the longest "ring path", i.e., the longest path in the corresponding inner dual of the benzenoid. Thus with four rings in a path the ring index varies approximately between 25 and 28 for the terminal ring, while the central rings have remarkably constant magnitude approximately between 21.5 and 22.5. For systems having at most three linearly or angularly fused rings the terminal ring varies between 22 and 23.5 and the central ring between 18.5 and 19.5. When we compare the derived geometrical ring indices with the recently reported geometrical ring indices obtained by Krygowski and collaborators,<sup>38</sup> the two sets of geometric ring indices do not correlate well. One of the reasons for the limited parallelism is due to considerable variation in the derived parameters in the approach of Krygowski and co-workers. Their ring indices apparently do not show simple regularities. Thus in the case of anthracene their parameter for the central ring has a smaller index than the terminal rings, but in phenanthrene the opposite is the case.

The ring indices of Table 12 and other local descriptors built by considering the corresponding vertices may be of interest in the substructure searching and particularly in the searching for cycle-containing pharmacophores that are imbedded in a larger polycyclic molecule. Structural ring indices have been considered in the literature for characterization of rings located in distinct local environment. For example, the ring ID numbers (the identification numbers) were based on enumeration of weighted paths involving vertices forming the ring of interest.<sup>39</sup> The **D/DD** ring indices can be used similarly in order to identify substructures of interest, particularly in combination with other ring indices, such as the ring imbedding index of Lipkus.<sup>40</sup>

What advantages does the **D/DD** matrix offer over its "components **D** and **DD**"? The distance matrix **D** and the Detour matrix **DD** have limited discriminating power, particularly when one is interested in local molecular features, such as atomic, bond, or ring descriptors. This is seen from Table 1 which gives the row sums for the six nonequivalent vertices of the tricyclic graph of Figure 1. Two pairs of vertices have the same row sums for the **D** matrix and one pair of vertices for the row sum of the **DD** matrix, while all the row sums are distinct for the **D/DD** matrix of the same graph. Hence, the **D/DD** matrix is more sensitive to the

immediate and global environment of vertices or larger graph fragments. In addition the new matrix **D/DD** is a source of novel invariants, such as the outlined  $\gamma$  index, which may find use in structure-property-activity studies as an alternative molecular descriptor.

#### ACKNOWLEDGMENT

The author thanks Professor S. Pejovnik and Professor J. Zupan from the National Institute of Chemistry, Ljubljana (Slovenia) for the hospitality enjoyed at their Institute. This paper is dedicated to Professor J. N. Murrell of University of Sussex Brighton (England), on the occasion of 65th anniversary celebration.

#### REFERENCES AND NOTES

- Wiener, H. Structural determination of paraffin boiling points. *J. Am. Chem. Soc.* **1947**, *69*, 2636.
- Platt, J. R. Prediction of isomeric differences in paraffin properties. *J. Phys. Chem.* **1952**, *56*, 328–336.
- Lovasz, L.; Pelikan, I. On the eigenvalues of trees. *Period. Math. Hung.* **1973**, *3*, 175–182.
- Randić, M. On characterization of molecular branching. *J. Am. Chem. Soc.* **1975**, *97*, 6609–6615.
- Hosoya, H. Topological index. A newly proposed quantity characterizing the topological nature of structural isomers of saturated hydrocarbons. *Bull. Chem. Soc. Jpn.* **1971**, *44*, 2332.
- Randić, M.; Razinger, M.; Plavšić, D. Double invariants. *MATCH* **1997**, *35*, 243–259.
- Randić, M. On molecular branching. *Acta Chim. Slov.* **1997**, *44*, 57–77.
- Bonchev, D.; Mekenyan, O.; Trinajstić, N. Topological Characterization of Cyclic Structures. *Int. J. Quant. Chem.* **1980**, *17*, 845–893.
- Bonchev, D.; Balaban, A. T.; Liu, X.; Klein, D. J. Molecular cyclicity and centrality of polycyclic graphs. I. Cyclicity based on resistance distances or reciprocal distances. *Int. J. Quant. Chem.* **1994**, *50*, 1–20.
- Klein, D.; Randić, M. Resistance distance. *J. Math. Chem.* **1993**, *12*, 81–95.
- Harary, F. *Graph theory*; Addison-Wesley: Reading, MA, 1969.
- Amić, D.; Trinajstić, N. On the detour matrix. *Croat. Chem. Acta* **1995**, *68*, 53–62.
- Lukovits, I.; The detour index. *Croat. Chem. Acta* **1996**, *69*, 873–882.
- Nikolić, S.; Trinajstić, N.; Jurić, A.; Mihalić, Z. The detour matrix and the detour index of weighted graphs. *Croat. Chem. Acta* **1996**, *69*, 1577–1591.
- Lukovits, I.; Razinger, M. On calculation of the detour index. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 283–286.
- Trinajstić, N.; Nikolić, S.; Mihalić, Z. On computing the molecular detour matrix. *Int. J. Quant. Chem.* In press.
- Trinajstić, N.; Nikolić, S.; Lučić, B.; Amić, D.; Mihalić, Z. The detour matrix in chemistry. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 631–638.
- Randić, M.; DeAlba, L. M.; Harris, F. E. Graphs with identical detour matrix. *Croat. Chem. Acta* In press.
- Tratch, S. S.; Stankevitch, M. I.; Zefirov, N. S. Combinatorial models and algorithms in chemistry. The expanded Wiener number - A novel topological index. *J. Comput. Chem.* **1990**, *11*, 899–908.
- Randić, M. Novel molecular descriptor for structure-property studies. *Chem. Phys. Lett.* **1993**, *211*, 478–483.
- Randić, M.; Guo, X.; Oxley, T.; Krishnapriyan, H. Wiener matrix: source of novel graph invariants. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 709–716.
- Randić, M. Restricted random walk on a graph as a source of novel molecular descriptors. *Theor. Chim. Acta.* **1995**, *92*, 97–106.
- Trinajstić, N.; Babić, D.; Nikolić, S.; Plavšić, D.; Amić, D.; Mihalić, Z. The Laplacian matrix in chemistry. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 368–376.
- Randić, M. Hosoya matrix - a source of new molecular descriptors. *Croat. Chem. Acta* **1994**, *67*, 415–429.
- Diudea, M. V. Cluj matrix invariants. *J. Chem. Inf. Comput. Sci.* **1997**, *35*, 300–305.
- Diudea, M. V.; Minailiuc, O. M.; Katona, G.; Gutman, I. Szeged matrices and related numbers. *MATCH* **1997**, *35*, 129–143.
- Diudea, M. V.; Minailiuc, O. M.; Nikolić, S.; Trinajstić, N. Matrices of reciprocal distance. *MATCH* **1997**, *35*, in press.
- Randić, M.; DeAlba, L. M. Dense graphs and sparse matrices. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 1078–1081.



- (29) Nikolić, S.; Trinajstić, N.; Jurić, A.; Mihalić, Z.; Krilov, G. Complexity of some interesting (chemical) graphs. *Croat. Chem. Acta* **1996**, *69*, 883–897.
- (30) Mohar, B.; Babić, D.; Trinajstić, N. A novel definition of the Wiener index for trees. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 153–154.
- (31) John, P.; Sachs, H. Calculating the number of perfect matching and of spanning trees, Pauling's orders, the characteristic polynomial and the eigenvectors of a benzenoid system. *Topic Current Chem.* **1990**, *153*, 145–179.
- (32) Derflinger; Polansky, O. E. *Int. J. Quant. Chem.*
- (33) Randić, M. Conjugated circuits and resonance energies of benzenoid hydrocarbons. *Chem. Phys. Lett.* **1975**, *38*, 68–70.
- (34) Randić, M. A graph theoretical approach to conjugation and resonance energies of hydrocarbons. *Tetrahedron* **1977**, *33*, 1905–1920.
- (35) Randić, M. Aromaticity and conjugation. *J. Am. Chem. Soc.* **1977**, *99*, 444–450.
- (36) Gutman, I.; Randić, M. A correlation between Kekule valence structures and conjugated circuits. *Chem. Phys.* **1979**, *41*, 265–270.
- (37) Randić, M. Resonance energy of very large benzenoid hydrocarbons. *Int. J. Quant. Chem.* **1980**, *17*, 549–586.
- (38) Krygowski, T. M.; Cyranski, M.; Ciesleski, A.; Swirska, B.; Leszczynski, P. Separation of the energetic and geometric contributions to aromaticity. 2. Analysis of the aromatic character of benzene rings in their various topological environments in benzenoid hydrocarbons. Crystal and molecular structure of coronene. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 1135–1141.
- (39) Randić, M. Ring ID numbers. *J. Chem. Inf. Comput. Sci.* **1986**, *26*, 142–147.
- (40) Lipkus, A. H.; A ring-imbedding index and the use in substructure searching. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 92–97.

CI9702407