

CHROM. 11,905

APPLICATION OF TOPOLOGICAL INDICES TO GAS CHROMATOGRAPHIC DATA: CALCULATION OF THE RETENTION INDICES OF ISOMERIC ALKYL BENZENES

D. BONCHEV and Ov. MEKENJAN

Higher School of Chemical Technology, 8010 Burgas (Bulgaria)

G. PROTIĆ

INA-Petroleum Industry, Research and Development Department, P.O.B. 02-849, 41001 Zagreb, Croatia (Yugoslavia)

and

N. TRINAJSTIĆ

Rugjer Bošković Institute, P.O.B. 1016, 41001 Zagreb, Croatia (Yugoslavia)

(First received November 13th, 1978; revised manuscript received March 5th, 1979)

SUMMARY

The retention indices (*RI*) of monoalkyl- and *o*-dialkylbenzenes are calculated using a two-parameter equation of the type $RI = a W(G)^n$. The variable $W(G)$ appearing in this equation is a topological index called the Wiener number, which is dependent on the size and the shape of a given structure. Thus, rather than adopting empirical parameters related to some structural detail of a molecule, the assumption is made that the retention indices depend substantially on the molecular size and shape.

The significance of the proposed equation for the calculation of retention indices lies in the considerable reduction of the structural parameters usually required, in addition to its simplicity in use.

INTRODUCTION

The problem of predicting retention indices in gas chromatography is of special interest owing to the wide application of this technique¹ and the difficulties that occur in many instances of the determination of compounds with complicated molecular structures^{2,3}. Additive schemes, using different structural parameters for their derivation, have been developed for predicting the retention values of several homologous series of organic compounds (*e.g.*, refs. 4–14). In these schemes, certain fragments, or distinct bond types, in a specific part of a molecule are identified and to them are assigned numerical values. Thus, the equations proposed for the calculation of retention indices contain numerous parameters. Therefore, for large hydrocarbon systems, because of the enormous number of branched isomers^{15,16} with the variety of bond forms^{17–19}, the generalizations that are used in this type of method are difficult if not

impossible to handle²⁰. Hence, there is a need to use more general methods for the study of correlations between the structural and chromatographic properties of molecules. A concept based on molecular topology, known as the molecular connectivity method^{21,22} appears to be promising for elucidating the structural origin of chromatographic retention data^{20,21,23-25}. It also opens new perspectives by offering a means for the potential use of alternative topological indices in this and related problems. Kováts²⁶ was the first to discover the close connection that exists between the structural characteristics (*e.g.*, branching) and gas chromatographic data of organic compounds.

Recently, we proposed a topological approach for the study of additive molecular properties which showed great efficiency in discriminating between seemingly similarly branched systems²⁷⁻²⁹. The approach is based on the use of the Wiener number of a given set of structures, $W(G)$. It appears that the Wiener number depends on the size and the shape of a skeletal network of a given molecule²⁹⁻³¹, and the smaller this number is the more compact is the corresponding structure²⁷. It can be shown that the Wiener number of a given structure is equal to half the sum of the off-diagonal elements of the distance matrix of the corresponding graph G , $D(G)$,

$$W(G) = \frac{1}{2} \sum_{i,j} D_{ij}(G) \quad (1)$$

The summation in eqn. 1 is over all the unique distances between the vertices i and j in the graph G . The distance matrix of G is a real $N \times N$ matrix [N = number of atoms (vertices) in a given structure (graph)] which contains elements $D_{ij}(G)$ representing the length (1,2,3,...) of the shortest path between the i th and j th vertex of G ³². All elements of the type $D_{ii}(G)$ are, by definition, zero.

The Wiener number is identical with the path number defined as the sum of the distances between any two carbon atoms in the hydrocarbon molecule in terms of carbon-carbon bonds, which was introduced by Wiener³³⁻³⁶ over 30 years ago in his studies on the additive physical properties of alkanes. This topological index was successfully applied to the characterization of acyclic, cyclic and more complicated structures and for the modelling of their thermodynamic functions^{27-29,37-39}, and it appears to be a useful device for ordering the sequence of isomers²⁷⁻²⁹. In this work, this approach is used for predicting the retention indices of a set of test molecules, *viz.*,

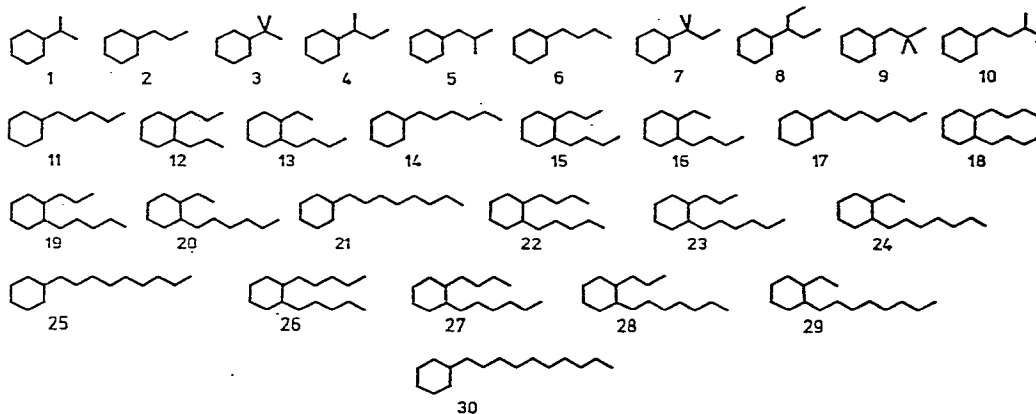


Fig. 1. Structures of molecules studied.

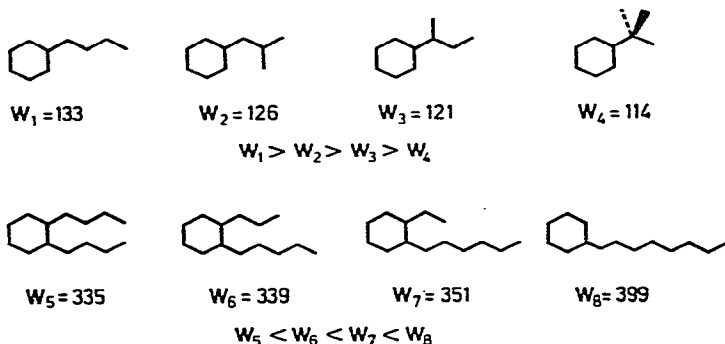
monoalkyl- and *o*-dialkylbenzenes. The hydrogen-suppressed graphs^{30,31} of these molecules are shown in Fig. 1.

These compounds, although limited in number, have been selected for a detailed description of the present approach because of the abundant experimental data available^{40,41}. In this choice we also followed the ideas of our approach²⁷⁻²⁹ according to which the best quantitative correlations can be obtained within a series of isomeric compounds that undergo only a certain identified structural change and, accordingly, are described by one of the rules of molecular branching and cyclicity²⁹ (see later for a discussion on the nature of these structural changes). It therefore appears that monoalkyl- and *o*-dialkylbenzenes represent a convenient set of isomeric molecules for structural and chromatographic studies^{*4-14}.

RESULTS AND DISCUSSION

The Wiener numbers of isomeric monoalkyl- and *o*-dialkylbenzenes, which were simply obtained from their distance matrices (see eqn. 1), are reported in Table I.

In all instances the isomeric structures were correctly differentiated. For example:



In the above samples two types of structural changes occur, both of which are well followed by changes in the Wiener numbers. The first is concerned with the increasing number of branches, which shortens the length of the alkyl side-chain, and with the displacement of a certain branch to a more internal position on the chain. In this instance the topological index $W(G)$ decreases on going from the least to the most branched structure (see $C_{10}H_{14}$ monoalkylbenzenes). The second type of structural change occurs with the isomeric $C_{14}H_{12}$ *o*-dialkylbenzenes, namely, an increase in the length of one of alkyl chains at the cost of a decrease in the length of the other alkyl chain produces regular increase in the $W(G)$ index. On this basis, one can predict the relative order of values of some molecular properties of alkylbenzenes, including the retention indices. Quantitative correlations could be also deduced, as these indices are very sensitive to structural changes in the molecules.

* The *o*-methylalkylbenzenes are not considered as they deviate in their topological behaviour from the other isomers of the *o*-dialkylbenzene series because they do not obey the same rule of molecular branching²⁷.

TABLE I

WIENER NUMBERS OF MONOALKYLBENZENES AND *o*-DIALKYL BENZENES

Graph*	$N(G)$ **	$W(G)$	Graph*	$N(G)$ **	$W(G)$
1	9	88	16	13	274
2	9	94	17	13	314
3	10	114	18	14	335
4	10	121	19	14	339
5	10	126	20	14	351
6	10	133	21	14	399
7	11	150	22	15	418
8	11	158	23	15	426
9	11	160	24	15	450
10	11	174	25	15	498
11	11	182	26	16	512
12	12	206	27	16	516
13	12	210	28	16	528
14	12	242	29	16	548
15	13	266	30	16	612

* Numbers correspond to graphs listed in Fig. 1.

** N = number of carbon atoms in the molecule.

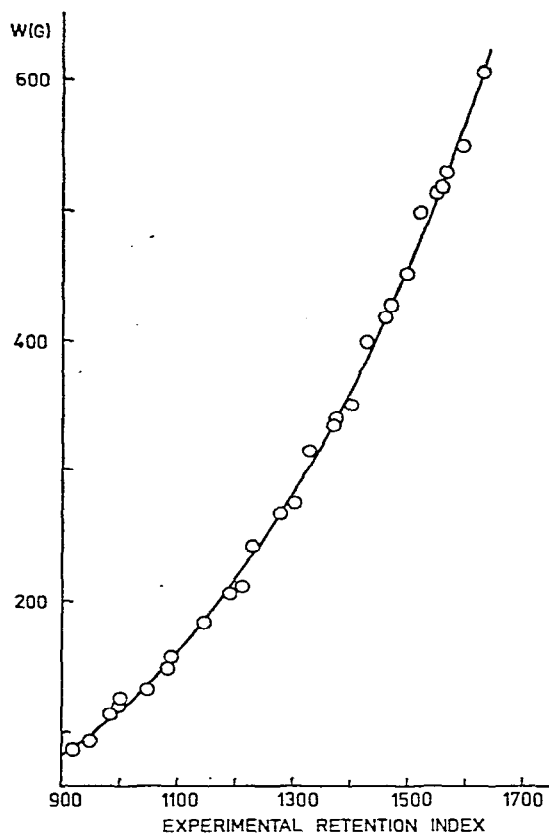


Fig. 2. Plot of the Wiener numbers of 28 monoalkylbenzenes and *o*-dialkylbenzenes (see Fig. 1) versus their experimental retention indices.

There are several possibilities open for establishing correlation between retention indices (*RI*), Wiener numbers and other topological quantities. We assumed the simplest, namely retention indices *versus* Wiener numbers of the molecules studied. This plot is shown in Fig. 2 for 28 compounds for which reliable retention indices on squalane at 368.4 and 413 °K were available.

The least-squares fit of this curve produced the following two-parameter equation:

$$RI = (244 \pm 4) W(G)^{(0.297 \pm 0.003)} \quad (2)$$

This equation was used for the calculation of the retention indices of monoalkyl- and *o*-dialkylbenzenes and the values obtained are given in Table II.

TABLE II
THEORETICAL AND EXPERIMENTAL RETENTION INDICES

Molecule*	Retention indices		Molecule*	Retention indices	
	Experimental ^{40,41}	Theoretical**		Experimental ^{40,41}	Theoretical**
1	918.6	919.5	16	1302.1	1288.0
2	947.4	937.7	17	1327.7	1341.2
3	984.9	992.9	18	1370.0	1367.2
4	1002.1	1010.7	19	1371.9	1372.0
5	1001.8	1022.9	20	1397.7	1386.2
6	1047.2	1039.4	21	1427.3	1440.0
7	1085.1	1077.2	22	1460.6	1460.0
8	1089.5	1093.9	23	1466.8	1468.2
9	—	1098.0	24	1496.4	1492.3
10	—	1124.6	25	1527.2	1537.8
11	1145.6	1140.8	26	1551.8	1550.5
12	1187.5	1183.5	27	1555.1	1554.1
13	1207.9	1190.3	28	1564.3	1564.8
14	1228.0	1241.4	29	1595.0	1582.1
15	1279.3	1276.7	30	1626.8	1634.8

* Numbers correspond to structures in Fig. 1.

** Calculated using eqn. 2.

Fig. 3 shows the correlation between the theoretical and experimental retention indices. The two sets of values are remarkably close (correlation coefficient = 0.9991) considering the simple equation used for the calculation. This good agreement suggests that the gas chromatographic behaviour of alkylbenzenes may depend critically on the size and the shape of the molecular network.

In our earlier studies on the topological parameters of molecules⁴², it was found that the number of atoms (*i.e.*, the size) and the number of bonds (*i.e.*, the shape represented by the connectivity) are the most important topological characteristics of a molecule. The connectivity in a molecule is represented by the (vertex) adjacency matrix of a structure^{29-32,42}, $A(G)$. The adjacency matrix is a real symmetrical matrix which has non-zero elements (equal to unity) only between bonded atoms

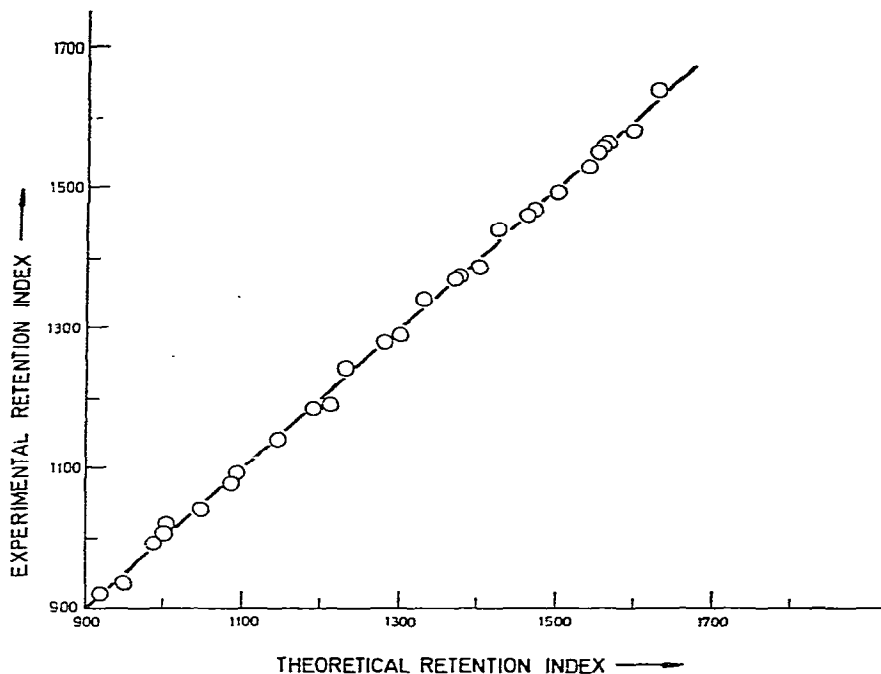


Fig. 3. Plot of experimental *versus* theoretical retention indices.

(adjacent vertices). The distance matrix is related to adjacency matrix by the relation*

$$D(G) = A(G) + \Sigma B(G) \quad (3)$$

where $B(G)$ are topological matrices containing only those non-zero entries which represent unique paths between the second, third, etc., neighbours. Thus, the distance matrix also represents the connectivity in a given structure, but in addition it has a greater information content than the corresponding adjacency matrix. The molecular connectivity method^{21,22} is related to the adjacency matrix in an indirect way.

This method is based on the calculation of the Randić's²¹ connectivity index $\chi_R(G)$:

$$\chi_R(G) = \sum_{\text{edges}} (d_i d_j)^{-\frac{1}{2}} \quad (4)$$

where d_i and d_j are the valencies (degrees)^{29,30,41} of the adjacent vertices (atoms) in a graph (structure). The sum of the non-zero entries in the row (column) of the adjacency matrix specifies the degree of a particular vertex in the graph G . However, Randić's index was obtained after a search for a specific index to characterize the branching of acyclic saturated hydrocarbons instead of analysing the information concealed in the structure of the adjacency matrix^{28,29,43,44}. Thus, the $\chi_R(G)$ index cannot be considered

* We are grateful to Damir Kasum (Zagreb) for inspiring discussions about the relationship between the distance and adjacency matrices.

as a truly graph-theoretical invariant as is the case with the Wiener number. It is a conveniently, but arbitrarily, adjusted graph-theoretical index which appears to be a very powerful tool for studying molecular additive properties²¹⁻²⁵. Randić, in his study of the origin of the chromatographic retention data of alkanes, used, in addition to the molecular connectivity index, the number of terminal paths of length three, T_3 , *i.e.*, the number of paths between two methyl groups three bonds apart. The Randić expression for the calculation of the retention indices of alkanes is

$$RI_{(\text{Randić})} = 200(Z_R - 1.4142) + (T_3)^2 + 300 \quad (5)$$

where the factor 200 converts the Kováts retention scale to that of the connectivity values, and 1.4142 and 300 are the connectivity index and the Kováts retention value for propane, respectively. Theoretical results obtained by means of eqn. 5 are in good agreement with the experimental values for alkanes²⁰.

As the molecular connectivity indices and Wiener numbers for a given set of isomers are roughly linearly correlated²⁷, both indices can be used for tracing the structural origin of chromatographic retention data. However, the applicability of both methods should be thoroughly investigated (such a study is in progress) before giving preference to either of them. However, from the pure graph-theoretical point of view, the advantage is in favour of the Wiener number because it is a well defined graphical invariant. From the practical point of view the accuracy of the simple equation proposed here is certainly attractive and competitive with the molecular connectivity method and with the empirical correlations. Hence the significance of the present study is three-fold. Firstly, it is very simple to use. Secondly, it demonstrates that even with the considerable reduction in the structural parameters ordinarily required⁴⁻¹⁴, gas chromatographic indices can be accurately reproduced. Thirdly, it clearly exhibits the virtues of the alternative topological index in studying the chromatographic properties of molecules.

ACKNOWLEDGEMENTS

We thank Drs. N. Dimov (Sofia) and D. Papazova (Burgas) for helpful discussions, and Professors A. T. Balaban (Bucuresti) and M. Randić (Ames, Iowa) for sending us copies of their unpublished work. We also thank Drs. Dj. Deur-Šiftar and V. Švob for useful lectures on gas chromatography. One of us (N.T.) was supported in part by N.S.F. Grant No. F6F006Y.

REFERENCES

- 1 H. J. Noebels, R. F. Wall and N. Brenner (Editors), *Gas Chromatography*, Academic Press, New York, 1961.
- 2 R. Porter (Editor), *Gas Chromatography in Biology and Medicine*, Churchill, London, 1969.
- 3 G. J. Dickes and P. V. Nicholas, *Gas Chromatography in Food Analysis*, Butterworths, London, 1976.
- 4 G. Schomburg and G. Dielman, *J. Chromatogr. Sci.*, 11 (1973) 151.
- 5 J. M. Takács, *J. Chromatogr. Sci.*, 11 (1973) 210.
- 6 G. Castello, M. Lunardelli and M. Berg, *J. Chromatogr.*, 76 (1973) 31.
- 7 M. Chastratte, G. Leufant, A. Remy and M. Cohen-Makabeh, *J. Chromatogr.*, 84 (1973) 275.
- 8 P. Souter, *J. Chromatogr. Sci.*, 12 (1974) 418.

- 9 R. Vanheertum, *J. Chromatogr. Sci.*, 13 (1975) 150.
- 10 G. Castello and C. D'Amato, *J. Chromatogr.*, 107 (1975) 1.
- 11 N. Dimov, *J. Chromatogr.*, 119 (1976) 109.
- 12 J. R. Chréten and J. E. Dubois, *J. Chromatogr.*, 126 (1976) 171.
- 13 D. Papazova and N. Dimov, *J. Chromatogr.*, 137 (1977) 259.
- 14 N. Dimov and D. Papazova, *J. Chromatogr.*, 137 (1977) 265.
- 15 A. Cayley, *Phil. Mag.*, 47 (1874) 444.
- 16 R. C. Read, in A. T. Balaban (Editor), *Chemical Applications of Graph Theory*, Academic Press, London, 1976, p. 25.
- 17 G. I. Spivakovskii, A. I. Tishchenko, I. I. Zaslavskii and N. S. Wulfson. *J. Chromatogr.*, 144 (1977) 1.
- 18 J. E. Dubois, D. Laurent and H. Veillard, *C.R. Acad. Sci.*, 263 (1966) 764 and 1245.
- 19 J. E. Dubois and D. Laurent, *Bull. Soc. Chim. Fr.*, (1969) 2449.
- 20 M. Randić, *J. Chromatogr.*, 161 (1978) 1.
- 21 M. Randić, *J. Amer. Chem. Soc.*, 97 (1975) 6609.
- 22 L. B. Kier and L. H. Hall, *Molecular Connectivity in Chemistry and Drug Research*. Academic Press, New York, 1976.
- 23 R. Kaliszan, *Chromatographia*, 10 (1977) 529.
- 24 R. Kaliszan and H. Foks, *Chromatographia*, 10 (1977) 346.
- 25 R. Kaliszan and H. Lamparczyk, *J. Chromatogr. Sci.*, 16 (1978) 246.
- 26 E. Kováts, *Z. Anal. Chem.*, 181 (1961) 351.
- 27 D. Bonchev and N. Trinajstić, *J. Chem. Phys.*, 67 (1977) 4517.
- 28 D. Bonchev and N. Trinajstić, *Int. J. Quantum Chem., Symp.*, S 12 (1978) 293.
- 29 D. Bonchev, J. V. Knop, and N. Trinajstić, in P. J. Plath (Editor), *Proceedings of the Bremen Conference on Mathematical Structures in Chemistry (Bremen, June 1978)*, Institut für Strahlenchemie im Max-Planck-Institut für Kohlenforschung, Mülheim/Ruhr, in press.
- 30 A. Graovac, I. Gutman and N. Trinajstić, in *Lecture Notes in Chemistry*, Springer-Verlag, Berlin, 1977, No. 4.
- 31 I. Gutman and N. Trinajstić, *Topics Curr. Chem.*, 42 (1973) 49.
- 32 D. H. Rouvray, in A. T. Balaban (Editor), *Chemical Applications of Graph Theory*, Academic Press, London, 1976, p. 175.
- 33 H. Wiener, *J. Amer. Chem. Soc.*, 69 (1947) 17.
- 34 H. Wiener, *J. Amer. Chem. Soc.*, 69 (1947) 2636.
- 35 H. Wiener, *J. Phys. Colloid Chem.*, 52 (1948) 425.
- 36 H. Wiener, *J. Phys. Chem.*, 52 (1948) 1082.
- 37 J. R. Platt, *J. Chem. Phys.*, 15 (1947) 419.
- 38 J. R. Platt, *J. Phys. Chem.*, 56 (1952) 328.
- 39 H. Hosoya, *Bull. Chem. Soc. Jap.*, 44 (1971) 2332.
- 40 L. Soják, J. Janák and J. A. Rijks, *J. Chromatogr.*, 135 (1977) 71.
- 41 G. Protić, *M.Sc. Thesis*, University of Zagreb, 1979.
- 42 N. Trinajstić, in G. A. Segal (Editor), *Semiempirical Methods of Electronic Structure Calculations. Part A: Techniques*, Plenum Press, New York, 1977, p. 1.
- 43 I. Gutman, B. Ruščić, N. Trinajstić and C. F. Wilcox, Jr., *J. Chem. Phys.*, 62 (1975) 3399.
- 44 D. Cvetković and I. Gutman, *Croat. Chem. Acta*, 49 (1977) 115.