coding reaction transforms and can be made more or less discriminating by using an appropriate number of eigenvalues. However, the method is not reversible in that the index cannot be used to derive the original structure. On the other hand, it can be used to derive a unique numbering for substructures by ordering the eigenvalues of the defined connectivity matrix.

Another useful benefit is associated with the eigenvectors, which can be used to determine the attribution of each atom to substructures upon disconnection of the main structure into distinct fragments. This could be useful when the possible disconnections of a main structure into reacting substructures are considered.

REFERENCES AND NOTES

(1) Wiswesser, W. J. *J. Chem. Inf. Comput. Sci.* **1982**, *22*, 88.
(2) Wiswesser, W. J. *J. Chem. Inf. Comput. Sci.* **1985**, *25*, 258.
(3) Randić, M. *J. Chem. Inf. Comput. Sci.* **1984**, *24*, 164.
(4) Ash, J. E.; Chubb, P. A.; Ward, S. E.; Welford, S. M.; Willett, P. *Communication, Storage and Retrieval of Chemical Information*; Ellis Horwood Ltd.: Chichester, U.K., 1985.
(5) Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. *Numerical Recipes: The Art of Scientific Computing*; Cambridge University Press: Cambridge, U.K., 1986.
(6) Henze, H. R.; Blair, C. M. *J. Am. Chem. Soc.* **1931**, *53*, 3077.
(7) Lederberg, J.; Sutherland, G. L.; Buchanan, B. G.; Feigenbaum, E. A.; Robertson, A. V.; Duffield, A. M.; Djerassi, C. *J. Am. Chem. Soc.* **1969**, *91*, 2973.

Topological Organic Chemistry. 1. Graph Theory and Topological Indices of Alkanes

HARRY P. SCHULTZ

Department of Chemistry, University of Miami, Coral Gables, Florida 33124

Received January 24, 1989

By use of the adjacency, degree, and distance matrices that describe the structure of an alkane, a method is outlined for deriving with simple matrix algebra a molecular topological index number of the alkane. Additionally, the calculations offer a picture of the relative intricacy (branched) value for each carbon atom of the alkane.

INTRODUCTION

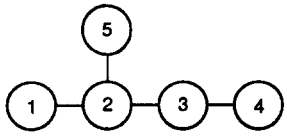
Graph theory offers the means to numerically characterize structures of chemicals. Single-number representations of alkanes have been reported by Hosoya[1] and Randić.[2] Balaban,[3] Randić,[4] and Hansen and Jurs[5] have summarized methods of calculating molecular topological indices, observing that problems, mainly of degeneracy, exist in the methods they reviewed.

This paper outlines a technique to determine molecular topological indices (MTI) by subjecting the adjacency, valence, and distance matrices describing the structures of alkanes to matrix algebraic operations. The procedure results in a single-number solution of high discrimination for each molecular graph. The method is simple to use. It is entirely objective, and the results are essentially monotonic. The representative list in Table I includes many of those pairs of compounds (**17** and **18**, **19** and **20**, **24** and **25**, **27** and **28**, **29** and **31**, **31** and **32**, **21** and **33**, **41** and **42**) for which earlier work presented ambiguous indices.

COMPUTATIONS

The method described is illustrated with a specific example, 2-methylbutane (**7**), hydrogen-suppressed, and with the interatomic carbon–carbon bond distances set at unity. The procedure uses the simplest of matrix algebraic operations.[6]

(**1**) The structural (molecular) graph for 2-methylbutane is drawn.



(**2**) Its distance (**D**) matrix is constructed.

$$\mathbf{D} = \begin{bmatrix} 0 & 1 & 2 & 3 & 2 \\ 1 & 0 & 1 & 2 & 1 \\ 2 & 1 & 0 & 1 & 2 \\ 3 & 2 & 1 & 0 & 3 \\ 2 & 1 & 2 & 3 & 0 \end{bmatrix}$$

(**3**) Its adjacency (**A**) matrix is constructed.

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

(**4**) Its valency (**v**) (degree) matrix is constructed.

$$\mathbf{v} = [1 \quad 3 \quad 2 \quad 1 \quad 1]$$

Müller et al.[7] demonstrated that the **D** matrix can be derived from the **A** matrix. Summing the elements in either the rows or columns of the **A** matrix yields the elements of the **v** vector; inspection of the molecular graph is the fastest way to construct the **v** vector.

(**5**) The **D** and **A** matrices of 2-methylbutane are summed to give the (**D** + **A**) matrix.

$$\underset{\mathbf{D}}{\begin{bmatrix} 0 & 1 & 2 & 3 & 2 \\ 1 & 0 & 1 & 2 & 1 \\ 2 & 1 & 0 & 1 & 2 \\ 3 & 2 & 1 & 0 & 3 \\ 2 & 1 & 2 & 3 & 0 \end{bmatrix}} + \underset{\mathbf{A}}{\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}} =$$

$$\underset{(\mathbf{D}+\mathbf{A})}{\begin{bmatrix} 0 & 2 & 2 & 3 & 2 \\ 2 & 0 & 2 & 2 & 2 \\ 2 & 2 & 0 & 2 & 2 \\ 3 & 2 & 2 & 0 & 3 \\ 2 & 2 & 2 & 3 & 0 \end{bmatrix}}$$

(**6**) Its (**D** + **A**) matrix is multiplied by the row vector **v** used as the premultiplier, thus affording conformability to the expression

0095-2338/89/1629-0227$01.50/0

Table I. Molecular Topological Indices and Boiling Points for Representative C_1–C_{15} Alkanes

no.	compd	MTI	bp, °C
1	methane	0	−164
2	ethane	4	−88.6
3	propane	16	−42
4	butane	38	−0.5
5	2-methylpropane	36	−11.7
6	pentane	74	36.1
7	2-methylbutane	68	27.9
8	2,2-dimethylpropane	64	9.4
9	hexane	128	68.7
10	2-methylpentane	118	60.3
11	3-methylpentane	114	63.2
12	2,3-dimethylbutane	108	58
13	2,2-dimethylbutane	106	49.7
14	heptane	204	98.4
15	2-methylhexane	190	90
16	3-methylhexane	182	92
17	2,4-dimethylpentane	176	80.5
18	3-ethylpentane	174	93.5
19	2,2-dimethylpentane	170	79.2
20	2,3-dimethylpentane	168	89.8
21	3,3-dimethylpentane	162	86
22	2,2,3-trimethylbutane	156	80.9
23	octane	306	125.7
24	3-methylheptane	276	119
25	4-methylheptane	272	117.7
26	2,5-dimethylhexane	270	109.1
27	2,2-dimethylhexane	260	106.9
28	2,4-dimethylhexane	258	109.5
29	3,4-dimethylhexane	246	117.7
30	3,3-dimethylhexane	244	112
31	3-ethyl-2-methylpentane	242	115.7
32	3-ethyl-3-methylpentane	232	118.3
33	2,2,3,3-tetramethylbutane	214	106.3
34	nonane	438	150.8
35	decane	604	174.1
36	undecane	808	196.8
37	dodecane	1054	216.3
38	tridecane	1346	235.4
39	tetradecane	1688	253.7
40	pentadecane	2084	270.6
41	2,6-dimethyl-5-ethyl-3-isopropyloctane[a]	1366	
42	2,3-dimethyl-6-ethyl-5-isopropyloctane[a]	1358	

[a] Suggested by Hansen and Jurs.[5]

$$\begin{matrix} & & & & & & \mathbf{(D+A)} & & & & \\ & & & & & \begin{bmatrix} 0 & 2 & 2 & 3 & 2 \\ 2 & 0 & 2 & 2 & 2 \\ 2 & 2 & 0 & 2 & 2 \\ 3 & 2 & 2 & 0 & 3 \\ 2 & 2 & 2 & 3 & 0 \end{bmatrix} & & & & \\ [1 & 3 & 2 & 1 & 1] & & & & & = & & & & \\ & & & & & & \mathbf{v(D+A)} & & & & \\ & & & & & & [15 & + & 10 & + & 12 & + & 16 & + & 15] \end{matrix}$$

The elements of the product, $\mathbf{v(D + A)}$, picture a relative intricacy (branched) value for each carbon atom of 2-methylbutane. The lower the number, the more intricate is the atom. Note that the intricacy value (16) of the number four primary carbon atom indicates that the atom is slightly less intricate than are the numbers one and five primary carbon atoms (15 each).

(7) The elements of the product, $\mathbf{v(D + A)}$, are summed to yield the molecular topological index (MTI) number (68) of 2-methylbutane.

$$\text{MTI} = [15 + 10 + 12 + 16 + 15] = 68$$

SUMMARY

Table I displays the decreasing magnitude of the topological index number for each alkane isomer of four complete series, the butanes, pentanes, hexanes, and heptanes, as the molecular intricacy increases. The data also demonstrate the discrimination power of this mode of computing the topological index number of an alkane. The boiling points of most of the substances are listed for comparison purposes. Correlations of boiling points with the Randić χ connectivity index values[2,4] and with the matrix MTI values of this paper present contrasts that favor the connectivity values. Twenty-nine compounds of Table I, **6–34**, with χ values of 2.414–4.414 have the following boiling point correlation values: coefficient of determination R^2 value, 0.976; coefficient of multiple correlation, 0.988; standard error of estimate, 5.087; equation coefficients, constant −98.785 and variable (1) 57.810. The same compound correlations of the MTI values (ranging from 74 to 438) were, in like order, 0.887, 0.942, 11.02, 17.922, and 0.365. A similar contrast of the nine heptane isomers of Table I, **14–22**, which cancels size as a factor, shows χ and MTI values, respectively, in the same order as above: 0.802 and 0.487; 0.896 and 0.698; 3.142 and 5.061; −25.058, 35.169 and 32.459, 0.315.[8]

Within the purview of matrix operations other routes to the topological index exist. For example, $(\mathbf{v} \times \mathbf{D})$ and $(\mathbf{v} \times \mathbf{A})$ can be separately calculated, the elements of each row or column of the product summed, and the total sums of each added together, resulting in the same topological index number $\mathbf{v(D + A)}$. The value of $(\mathbf{v} \times \mathbf{A})$ can be computed rapidly by summing the squares of each element of the $\mathbf{v}$ vector. Other shortcuts to the index can be taken, but all lack the peripheral advantage illustrated above in step 6, which indicates the relative intricacy of each carbon atom of a molecule in a similar descending pattern as is pictured for increasingly intricate molecular isomers.

REFERENCES AND NOTES

(1) Hosoya, H. *Bull. Chem. Soc. Jpn.* **1971**, *44*, 2332–2339.
(2) Randić, M. *J. Am. Chem. Soc.* **1975**, *97*, 6609–6615.
(3) Balaban, A. T. *J. Chem. Inf. Comput. Sci.* **1985**, *25*, 334–343, and references cited therein.
(4) Randić, M. *J. Chem. Inf. Comput. Sci.* **1988**, *28*, 142–147, and references cited therein.
(5) Hansen, P. J.; Jurs, P. C. *J. Chem. Educ.* **1988**, *65*, 574–580, 661–664, and references cited therein.
(6) Campbell, H. G. *An Introduction to Matrices, Vectors, and Linear Programming*, 2nd ed.; Prentice-Hall; Englewood Cliffs, NJ, 1977.
(7) Müller, W. R.; Szymanski, K.; Knop, J. V.; Trinajstić, N. *J. Comput. Chem.* **1987**, *8*, 170–173.
(8) I thank a reviewer for providing the statistics that supported the above comparisons.