

Short Communication

Application of topographic indices to chromatographic data: calculation of the retention indices of alkanes

NADA BOŠNJAK

Pliva Research Institute, I. L. Ribara 89, 41 000 Zagreb, Croatia (Yugoslavia)

ZLATKO MIHALIĆ

Faculty of Science and Mathematics, University of Zagreb, Strossmayerov trg 14, 41 000 Zagreb, Croatia (Yugoslavia)

and

NENAD TRINAJSTIĆ*

Rugjer Bošković Institute, P.O. Box 1016, 41 001 Zagreb (Yugoslavia)

(First received June 13th, 1990; revised manuscript received October 30th, 1990)

ABSTRACT

Gas chromatographic (GC) data on alkanes have been re-examined from a structural point of view with emphasis on the distances in optimum (minimum-energy) conformations. The information on the distances in an alkane *G* is embodied in the topographic (geometric) distance matrix **D**(*G*) and the related 3-D Wiener number ${}^3W(G)$ of *G*. The optimum 3-D structures of alkanes were obtained from the molecular mechanics computations. The GC retention indices (*I*) of 157 alkanes were calculated using a three-parameter equation of the form $I = a[{}^3W(G)]^b + c$. The calculated *I* values are in excellent agreement with the experimental values. A comparison between the reported results and those obtained with the 2-D Wiener number and the connectivity index is also discussed.

INTRODUCTION

Quantitative structure–chromatographic retention relationships (QSCRR) with topological (graph-theoretical) indices have evolved in the past 15 years into a convenient model for studying the correlations between chromatographic and structural properties of molecules [1]. Topological indices [2] appear in this model as molecular descriptors which can be used for predicting the gas chromatographic (GC) retention indices (*I*) via an appropriate statistical equation. Ordinarily, indices employed in QSCRR are obtained for the two-dimensional (2-D) diagrams (graphs) serving as models for molecules [1,2]. Molecules, however, are three-dimensional (3-D) objects. Hence, in addition to their topological and combinatorial contents, their 3-D character is of profound importance [3]. In view of the considerable success of

topological indices in describing structure–chromatographic data correlations, it seemed desirable to investigate how to extend these models to include 3-D structural features of molecules. One such attempt is described in this paper.

In this paper we introduce the use of topographic (geometric) indices [4] in QSCRR. These indices contain information about the spatial 3-D architecture of molecules. Here we investigate how the introduction of topographic indices into the framework of QSCRR compares with the QSCRR models with topological indices. The specific topographic index considered is the 3-D Wiener number. As test molecules we selected alkanes because they are not particularly complicated molecules, their GC data are well known and several QSCRR with topological indices for them are available in the literature [5,6].

COMPUTATION OF THE 3-D WIENER NUMBER

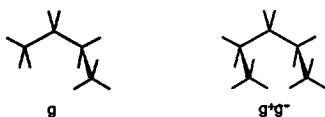
The original 2-D Wiener number (2W) was introduced in 1947 by Wiener [7]. This number (called the path number by Wiener) was defined as the number of bonds between all pairs of atoms in an acyclic molecule. Twenty-five years later, Hosoya [8] gave a graph-theoretical definition of the Wiener number. The 2-D Wiener number ${}^2W(G) = {}^2W$ of a structure G is equal to the half-sum of the elements of the distance matrix [9] ${}^2D(G) = {}^2D$ of G :

$${}^2W = \frac{1}{2} \sum_{k>l} ({}^2D)_{kl} \quad (1)$$

where the distance matrix entry $({}^2D)_{kl}$ denotes the topological (graph-theoretical) distance in G between sites k and l , *i.e.*, the length of the shortest path between k and l .

The distance matrix of a molecule G may also be based on the topographic (geometric) distances, *i.e.*, the matrix elements $({}^3D)_{kl}$ may represent “true” spatial distances between atoms k and l in G expressed in some units of length [10]. This topographic (geometric) distance matrix ${}^3D(G) = {}^3D$ contains information on the 3-D structure of a molecule. The corresponding Wiener number is named the 3-D Wiener number [11], and is representative of topographic invariants [4]. It is denoted by ${}^3W(G) = {}^3W$.

The actual computations of the 3-D Wiener numbers for alkanes were done in the following way. Owing to free rotation about single bonds, alkanes possess more than one conformation, each in conformational equilibrium with other conformations. To simplify the treatment, we considered only the most stable conformation, present in the equilibrium mixture in the greatest amount. For each alkane the intuitively most stable (minimum-energy) conformation was constructed using standard bond lengths and bond angles. Usually it was the most extended conformation, with the minimum number of *gauche*-butane (g) interactions, without the “forbidden” pentane (g^+g^-) interactions.



These starting geometries were then refined using standard molecular mechanics calculations [12], giving the chosen conformation optimum bond lengths, bond angles and interatomic distances. The alkane geometry obtained in such a way was then used to set up the corresponding topographic distance matrix and to calculate the 3-D Wiener number of the alkane by summing the elements in the upper triangle of the matrix. The elements of ${}^3\mathbf{D}$, *i.e.*, the distances between atoms in the alkane (${}^3\mathbf{D}$)_{kl}, were calculated from the Cartesian coordinates, which are the part of the standard output of the molecular mechanics program:

$$({}^3\mathbf{D})_{kl} = [(x_k - x_l)^2 + (y_k - y_l)^2 + (z_k - z_l)^2]^{1/2} \quad (2)$$

Schematically, the computation procedure may be represented by

$$\begin{array}{ccccccc} \text{intuitive} & & \text{optimum} & & \text{topographic} & & \\ \text{conformation} & \rightarrow & \text{conformation} & \rightarrow & \text{distance matrix} & \rightarrow & {}^3W(G) \\ \text{of an alkane } G & & \text{of } G & & \text{of } G & & \end{array}$$

The 3-D Wiener number was computed for two models representing alkanes. In the first model only carbon atoms in the most stable conformation of an alkane were considered. The 3W number obtained for this model of alkanes is denoted by 3W_C . In the second model all atoms in the most stable alkane conformation were taken into account. The corresponding 3-D Wiener number is denoted by ${}^3W_{CH}$.

The 2-D Wiener number 2W_C was calculated for the bare carbon skeletons of alkanes by setting up the distance matrix using the appropriate computer code [13] and then employing eqn. 1. This procedure may be schematized as follows:

$$\begin{array}{ccc} \text{carbon skeleton} & & \text{topological} \\ \text{of an alkane } G & \rightarrow & \text{distance matrix} \rightarrow {}^2W_C(G) \\ & & \text{of } G \end{array}$$

RESULTS AND DISCUSSION

Table I gives the 2-D Wiener numbers, 2W_C , the 3-D Wiener numbers, 3W_C and ${}^3W_{CH}$, connectivity indices, ${}^2\chi_C$, and experimental, I_{exp} , and calculated, I_{calc} , retention indices for 157 alkanes. Experimental GC retention indices measured on squalane at 333 K were taken from a paper by Chrétien and Dubois [14] and on squalane at 373 K from a paper by Schomburg and Dielman [15].

We tested several relationships between the experimental GC retention indices and the Wiener number. The best relationship was found to be an adjusted Walker-type relationship [16,17] of the form

$$I = a W^b + c \quad (3)$$

where a , b and c are statistical parameters to be determined by non-linear least-squares regression, and W represents the Wiener numbers 2W_C , 3W_C and ${}^3W_{CH}$. The values of the statistical parameters are given in Table II.

The statistical characteristics of all three correlations are of fairly good quality,

TABLE I

2-D WIENER NUMBER, 2W_C , 3-D WIENER NUMBERS, 3W_C AND ${}^3W_{CH}$, CONNECTIVITY INDICES, ${}^2\chi_C$, EXPERIMENTAL RETENTION INDICES, I_{exp} , AND CALCULATED RETENTION INDICES, I_{calc} , USING THE ${}^3W_{CH}$ NUMBER

Experimental retention indices measured on squalane at 333 K taken from ref. 14 and on squalane at 373 K from ref. 15.

Compound	2W_C	3W_C	${}^3W_{CH}$	${}^2\chi_C$	I_{exp}	I_{calc}
Methane	0	0.000	15.36	0.000	100.0	64.3
Ethane	1	1.532	56.88	1.000	200.0	190.1
Propane	4	5.607	131.32	1.141	300.0	300.1
Butane	10	13.60	251.72	1.914	400.0	406.4
2,2-Dimethylpropane	16	21.26	377.74	2.000	412.6	483.8
2-Methylbutane	18	23.37	395.20	2.270	475.4	493.0
Pentane	20	26.69	426.97	2.414	500.0	509.1
2,2-Dimethylbutane	28	35.64	571.17	2.561	537.6	572.8
2,3-Dimethylbutane	29	36.95	583.43	2.643	568.1	577.7
2-Methylpentane	32	41.04	620.14	2.770	569.8	591.9
3-Methylpentane	31	39.73	608.00	2.808	584.6	587.2
Hexane	35	46.22	669.74	2.914	600.0	610.1
2,2-Dimethylpentane	46	57.83	844.90	3.061	626.3	667.7
2,4-Dimethylpentane	48	59.07	853.31	3.126	630.1	670.3
2,2,3-Trimethylbutane	42	52.49	797.28	2.943	641.1	653.0
3,3-Dimethylpentane	44	55.30	822.11	3.121	660.2	660.7
2-Methylhexane	52	66.12	919.34	3.270	666.8	689.7
2,3-Dimethylpentane	46	57.81	845.15	3.181	672.5	667.8
3-Methylhexane	50	63.83	900.30	3.308	676.5	684.2
3-Ethylpentane	48	57.76	838.30	3.346	686.6	665.7
2,2,4-Trimethylpentane	66	79.92	1121.20	3.417	690.9	743.4
Heptane	56	73.38	989.89	3.414	700.0	709.3
2,2-Dimethylhexane	71	88.45	1200.13	3.561	719.9	762.6
2,2,3-Trimethylpentane	63	77.86	1107.57	3.481	738.6	740.0
2,3-Dimethylhexane	70	87.30	1191.73	3.681	760.8	760.6
2,3,3-Trimethylpentane	62	76.69	1097.57	3.504	761.4	737.5
3-Ethyl-2-methylpentane	67	80.32	1124.20	3.719	762.4	744.2
2-Methylheptane	79	100.21	1308.98	3.770	765.0	787.6
4-Methylheptane	75	95.50	1268.76	3.808	767.4	778.6
3,4-Dimethylhexane	68	85.22	1175.40	3.719	771.6	756.7
3-Methylheptane	76	96.61	1277.61	3.808	772.6	780.6
2,2,4,4-Tetramethylpentane	88	105.04	1435.09	3.707	774.6	814.8
3,3-Dimethylhexane	67	80.68	1128.11	3.621	775.7	745.1
2,2,4-Trimethylhexane	98	118.77	1556.10	3.955	777.3	839.4
2,2,5-Trimethylhexane	94	114.22	1518.57	3.917	790.7	831.9
Octane	84	109.51	1399.79	3.914	800.0	807.4
2,4,4-Trimethylhexane	92	111.97	1500.13	3.977	809.7	828.2
2,3,5-Trimethylhexane	96	115.85	1530.93	4.037	813.2	834.4
2,2-Dimethylheptane	104	129.44	1658.81	4.061	816.2	859.2
2,2,5,5-Tetramethylhexane	127	151.41	1941.58	4.207	820.1	909.7
2,4-Dimethylheptane	102	125.98	1627.06	4.164	821.2	853.2
2,2,3,4-Tetramethylpentane	86	103.72	1429.17	3.854	821.9	813.6
2,2,3-Trimethylheptane	92	112.97	1511.19	3.981	823.3	830.5
2,2-Dimethyl-3-ethylpentane	88	106.36	1450.39	4.019	824.4	818.0
4-Ethyl-2-methylhexane	98	115.12	1522.37	4.202	824.9	832.7
2,6-Dimethylheptane	108	133.06	1687.31	4.126	827.5	864.6

(Continued on p. 434)

TABLE I (continued)

Compound	2W_C	3W_C	${}^3W_{CH}$	${}^2\chi_C$	I_{exp}	I_{calc}
4,4-Dimethylheptane	96	120.18	1580.25	4.121	828.6	844.2
2,5-Dimethylheptane	104	128.84	1652.47	4.164	833.7	858.0
3,5-Dimethylheptane	100	123.59	1605.47	4.202	834.4	849.1
3,3-Dimethylheptane	98	122.41	1598.06	4.121	837.5	847.6
2,4-Dimethyl-3-ethylpentane	90	107.35	1456.11	4.091	838.4	819.2
2,3,3-Trimethylhexane	90	110.77	1493.25	4.004	841.7	826.8
3-Ethyl-2-methylhexane	96	114.77	1522.62	4.219	844.4	832.8
2,3,4-Trimethylhexane	92	112.17	1503.55	4.091	849.7	828.9
3,3,4-Trimethylhexane	88	108.46	1474.27	4.042	855.1	823.0
2,3-Dimethylheptane	102	126.89	1637.14	4.181	855.5	855.1
3-Ethyl-4-methylhexane	94	112.56	1504.60	4.257	855.6	829.1
2,2,3,3-Tetramethylpentane	82	103.73	1429.19	3.811	855.8	813.6
3-Ethyl-3-methylhexane	92	110.56	1487.49	4.182	856.0	825.7
3,4-Dimethylheptane	98	122.46	1599.99	4.129	858.0	848.0
4-Ethylheptane	102	123.55	1603.71	4.346	858.2	848.7
2,3,3,4-Tetramethylpentane	84	101.73	1413.01	3.887	861.1	810.2
4-Methyloctane	108	137.21	1735.40	4.308	863.3	873.5
2-Methyloctane	114	144.39	1797.29	4.270	864.8	884.6
3-Ethylheptane	104	125.58	1619.73	4.346	867.4	851.8
2,4,6-Trimethylheptane	135	162.21	2039.83	4.520	870.1	926.0
3-Methyloctane	110	139.69	1757.22	4.308	870.8	877.4
2,2,4,5-Tetramethylhexane	124	148.13	1916.57	4.327	872.1	905.4
2,2,6-Trimethylheptane	139	168.09	2092.28	4.417	873.0	934.5
2,2,3,5-Tetramethylhexane	123	146.95	1906.31	4.337	873.3	903.7
2,3-Dimethyl-3-ethylpentane	86	102.93	1420.38	4.065	875.0	811.8
2,2,4-Trimethylheptane	131	159.92	2027.35	4.455	875.7	923.9
2,2,5-Trimethylheptane	134	163.04	2052.19	4.455	878.1	928.0
3,3-Diethylpentane	88	104.35	1430.61	4.243	880.2	813.9
2,2-Dimethyl-4-ethylhexane	126	146.35	1896.25	4.493	881.3	902.0
2,2,4,4-Tetramethylhexane	119	143.07	1875.23	4.268	888.6	898.3
2,4,4-Trimethylheptane	127	155.35	1988.87	4.477	889.4	917.6
2,5-Dimethyl-3-ethylhexane	127	147.98	1912.26	4.575	891.4	904.7
2,5,5-Trimethylheptane	131	159.64	2023.18	4.477	891.7	923.3
Nonane	120	155.82	1909.68	4.414	900.0	904.3
2,2-Dimethyl-3-ethylhexane	122	146.34	1905.67	4.519	902.1	903.6
2,3,3,5-Tetramethylhexane	120	144.15	1884.96	4.360	903.3	900.0
3-Ethyl-2,2,4-trimethylpentane	115	134.81	1801.01	4.392	903.9	885.3
2,4,5-Trimethylheptane	130	158.01	2009.28	4.575	906.7	921.0
4-Ethyl-2-methylheptane	134	158.72	2012.39	4.702	907.4	921.5
3,3,5-Trimethylheptane	126	154.02	1976.33	4.515	907.7	915.5
2,2,3,4-Tetramethylhexane	118	142.73	1877.27	4.392	908.8	898.7
2,3,5-Trimethylheptane	131	158.83	2015.02	4.575	912.9	921.9
2,2,3-Trimethylheptane	130	159.49	2026.14	4.481	914.4	923.7
2,2-Dimethyloctane	146	181.61	2224.40	4.561	914.9	955.2
2,4-Dimethyl-3-isopropylpentane	117	138.26	1833.36	4.464	915.1	891.0
3-Isopropyl-2-methylhexane	124	146.59	1905.28	4.591	915.5	903.5
2,4-Dimethyloctane	142	175.75	2171.92	4.664	915.8	947.0
4,4-Dimethyloctane	134	167.58	2104.49	4.621	918.0	936.4
2,3,6-Trimethylheptane	136	164.67	2064.61	4.537	919.0	930.0
2,4-Dimethyl-4-ethylhexane	122	146.13	1904.27	4.538	920.7	903.3
2,2,3,4,4-Pentamethylpentane	111	132.22	1782.18	4.155	921.7	881.9
3,5-Dimethyloctane	143	171.13	2132.56	4.664	921.8	940.9

TABLE I (continued)

Compound	2W_C	3W_C	${}^3W_{CH}$	${}^2\chi_C$	I_{exp}	I_{calc}
2,5-Dimethyloctane	138	177.29	2186.04	4.702	921.8	949.2
2,3,4,5-Tetramethylhexane	121	146.37	1907.95	4.464	923.1	904.0
5-Ethyl-2-methylheptane	138	163.67	2053.96	4.702	924.8	928.3
4-Isopropylheptane	131	156.58	1997.73	4.719	925.0	919.1
2,7-Dimethyloctane	151	186.84	2268.53	4.626	928.5	961.9
2,2,3,3-Tetramethylhexane	115	139.85	1852.33	4.311	928.8	894.3
3,6-Dimethyloctane	141	175.26	2169.96	4.702	929.0	946.7
2,4-Dimethyl-3-ethylhexane	122	144.73	1890.57	4.629	929.8	901.0
2,6-Dimethyloctane	146	186.83	2268.43	4.664	931.5	961.9
2,3,3-Trimethylheptane	127	156.01	1996.23	4.504	931.7	918.8
3,3-Dimethyloctane	138	172.43	2146.83	4.621	932.0	943.1
3,4,4-Trimethylheptane	122	150.07	1945.35	4.542	932.2	910.3
2,3,4-Trimethylheptane	128	156.37	1997.87	4.591	933.4	919.1
2,3,4,4-Tetramethylhexane	116	140.54	1858.26	4.415	935.0	895.4
4-Ethyl-3-methylheptane	129	154.75	1981.17	4.757	935.7	916.3
3,4-Dimethyloctane	137	171.13	2136.17	4.719	936.0	941.4
3,3,4-Trimethylheptane	123	151.18	1954.57	4.542	936.6	911.9
4-Ethyl-4-methylheptane	126	151.62	1954.76	4.682	937.6	911.9
3,3-Dimethyl-4-ethylhexane	118	142.72	1877.96	4.580	937.8	898.8
3-Ethyl-4-methylheptane	130	155.37	1984.91	4.757	940.5	916.9
3-Ethyl-2-methylheptane	134	160.20	2026.81	4.719	941.0	923.8
4,5-Dimethyloctane	135	168.68	2114.29	4.719	943.1	938.0
3,4,5-Trimethylheptane	125	152.63	1965.00	4.629	945.0	913.6
3,4-Diethylhexane	125	145.65	1894.07	4.795	945.8	901.6
2,3,3,4-Tetramethylhexane	115	139.09	1846.20	4.425	949.1	893.3
2,3-Dimethyl-4-ethylhexane	123	144.13	1882.67	4.629	949.4	899.6
4-Ethylheptane	141	171.24	2130.45	4.846	951.5	940.5
2,3-Dimethyloctane	143	177.74	2191.63	4.681	952.1	950.1
2-Ethyl-2-methylheptane	129	154.78	1980.23	4.621	953.0	916.2
2,2,3,3,4-Pentamethylpentane	108	129.68	1764.19	4.193	953.4	878.7
3,3-Diethylhexane	121	142.70	1872.36	4.743	954.1	897.8
5-Methylnonane	149	189.15	2302.73	4.808	957.4	967.0
4-Methylnonane	150	190.38	2313.66	4.808	960.0	968.7
2-Methylnonane	158	200.10	2398.37	4.770	963.9	981.2
3-Ethylheptane	145	175.85	2170.31	4.846	964.0	946.8
3,4-Dimethyl-3-ethylhexane	117	139.51	1847.09	4.603	964.6	893.4
3-Ethyl-2,2,3-trimethylpentane	110	130.85	1770.99	4.371	965.7	879.9
3-Ethyl-2,3,4-trimethylpentane	112	132.68	1786.25	4.447	969.4	882.7
3-Methylnonane	153	194.11	2346.10	4.808	969.6	973.5
3,3,4,4-Tetramethylhexane	111	135.63	1818.76	4.371	983.7	888.4
Decane	165	213.63	2531.95	4.914	1000.0	1000.2
Undecane	220	284.23	3277.49	5.414	1100.0	1095.3
6-Methylundecane	261	331.10	3805.18	5.808	1151.8	1153.7
4-Methylundecane	265	336.12	3849.17	5.808	1158.6	1158.3
2-Methylundecane	277	350.81	3978.30	5.770	1164.0	1171.6
3-Methylundecane	270	342.50	3905.70	5.808	1169.6	1164.2
Dodecane	286	368.76	4157.00	5.914	1200.0	1189.5
5,7-Dimethylundecane	308	384.71	4399.31	6.202	1190.4	1213.0
4,6-Dimethylundecane	310	386.91	4418.04	6.202	1193.0	1214.8
3,5-Dimethylundecane	316	394.67	4486.86	6.202	1207.2	1221.2
2,4-Dimethylundecane	326	406.46	4589.45	6.164	1208.2	1230.7

(Continued on p. 436)

TABLE I (continued)

Compound	2W_C	3W_C	${}^3W_{CH}$	${}^2\chi_C$	I_{exp}	I_{calc}
2,5-Dimethylundecane	324	403.83	4565.88	6.164	1210.4	1228.6
2,6-Dimethylundecane	324	403.96	4567.77	6.164	1210.4	1228.7
2,7-Dimethylundecane	326	406.49	4589.98	6.164	1215.8	1230.8
5,6-Dimethylundecane	306	383.75	4393.91	6.219	1223.4	1212.5
4,5-Dimethylundecane	310	388.57	4435.84	6.219	1230.4	1216.4
2,9-Dimethylundecane	336	419.21	4702.51	6.164	1232.6	1241.1
3,4-Dimethylundecane	318	398.05	4519.20	6.219	1247.0	1224.2
2,3-Dimethylundecane	330	411.92	4639.25	6.181	1251.4	1235.3

but the best correlation is obtained between I_{exp} and ${}^3W_{CH}$. Hence, the calculated retention indices, I_{calc} , reported in Table I were obtained using the equation

$$I_{calc} = 119.5({}^3W_{CH})^{0.2947} - 202.9 \quad (4)$$

A plot of I_{exp} vs. ${}^3W_{CH}$ is shown in Fig. 1 and Fig. 2 is a plot of residuals $I_{calc} - I_{exp}$ vs. ${}^3W_{CH}$.

It is interesting that eqn. 3 with 2W_C unexpectedly possesses slightly better statistical characteristics than eqn. 3 with 3W_C .

We also compared our calculations with the results that one would obtain by using the connectivity index [5] $\chi(G) = \chi$ instead of the Wiener number. We selected the connectivity index because it is a well established topological descriptor in QSCRR [1,18-23].

The connectivity index χ is defined as a bond additivity quantity. First one finds contributions from the individual bonds and these are defined by the weights $(d_k d_l)^{-1/2}$, where d_k and d_l are valencies of the atoms k and l making up the bond $k-l$. Then these bond weights are added:

$${}^2\chi_C = \sum_{k,l} (d_k d_l)^{-1/2} \quad (5)$$

TABLE II

STATISTICAL CHARACTERISTICS OF THE ADJUSTED WALKER-TYPE RELATIONSHIP BETWEEN THE RETENTION INDICES OF 157 ALKANES AND THE WIENER NUMBERS

Parameter	2W_C	3W_C	${}^3W_{CH}$
<i>a</i>	171.2 ± 15.7	170.6 ± 17.0	119.5 ± 21.8
<i>b</i>	0.3347 ± 0.0128	0.3245 ± 0.0133	0.2947 ± 0.0174
<i>c</i>	48.6 ± 27.3	31.8 ± 30.2	-202.9 ± 57.0
<i>r</i>	0.9844	0.9817	0.9862
<i>F</i>	2403	2048	2728
<i>s</i>	32.97	35.62	31.00

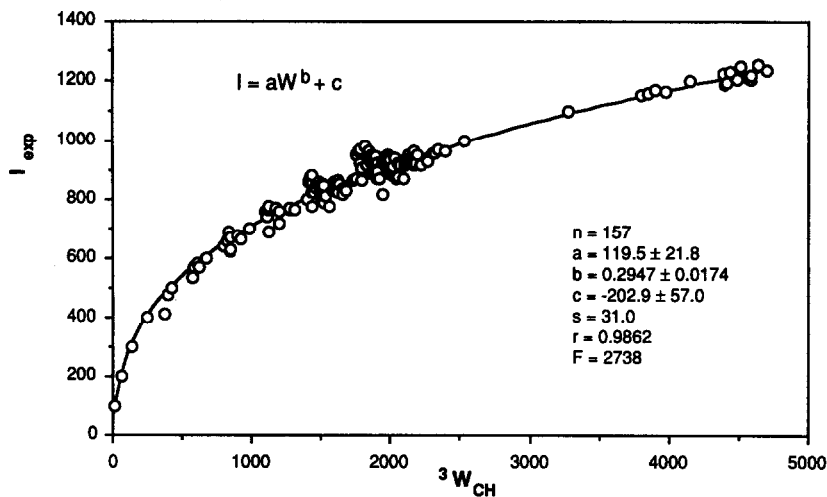


Fig. 1. Plot of I_{exp} vs. ${}^3W_{CH}$ for 157 alkanes with statistical characteristics.

where ${}^2\chi_C$ is the 2-D connectivity index corresponding to a carbon skeleton of an alkane. For alkanes, eqn. 5 can be given in a closed form. Thus, in the alkane carbon skeletons four types of sites appear, *i.e.*, sites with valency 1, 2, 3 and 4. These give rise to ten types of bonds with the following weights: $(1,1) = 1$, $(1,2) = 0.7071$, $(1,3) = 0.5773$, $(1,4) = 0.5$, $(2,2) = 0.5$, $(2,3) = 0.4082$, $(2,4) = 0.3536$, $(3,3) = 0.3333$, $(3,4) = 0.2887$ and $(4,4) = 0.25$. If we denote the number of bonds of each type by b_{11} , b_{12} , b_{13} , b_{14} , b_{22} , b_{23} , b_{24} , b_{33} , b_{34} and b_{44} , respectively, and if we use the related weights from the above, then eqn. 5 reduces to

$${}^2\chi_C = b_{11} + 0.7071b_{12} + 0.5773b_{13} + 0.5b_{14} + 0.5b_{22} + 0.4082b_{23} + 0.3536b_{24} + 0.3333b_{33} + 0.2887b_{34} + 0.25b_{44} \quad (6)$$

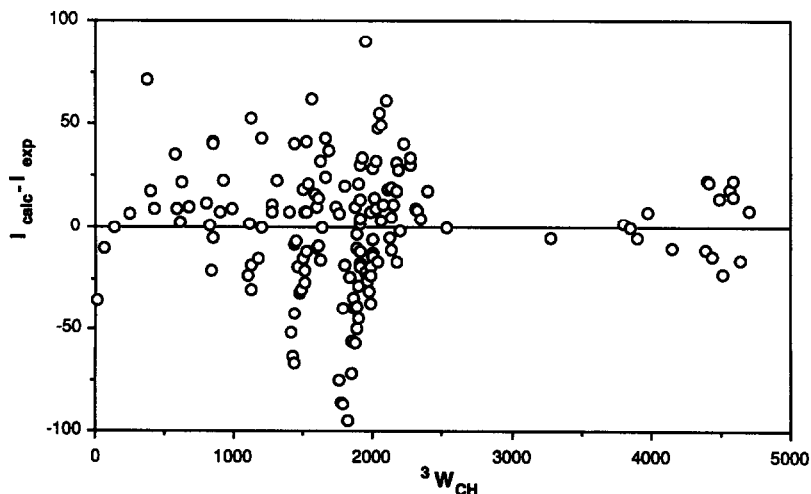


Fig. 2. Plot of $I_{calc} - I_{exp}$ vs. ${}^3W_{CH}$.

This expression reveals that the connectivity indices are fully determined by the numbers of the bond types that may appear in alkanes. The computed connectivity indices for 157 alkanes are also given in Table I.

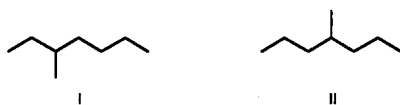
In the case of the connectivity index, the best correlation between GC retention indices of alkanes and their connectivity indices is linear:

$$I = k \ ^2\chi_C + l \quad (7)$$

The following values of the statistical parameters are obtained: $n = 157$, $k = 186.93 \pm 2.11$, $l = 69.81 \pm 9.31$, $r = 0.9902$, $F = 7827$ and $s = 26.0$.

A plot of I_{exp} vs. $\ ^2\chi_C$ is given in Fig. 3.

The comparison between $\ ^3W_{\text{CH}}$ and $\ ^2\chi_C$ points to some advantages of $\ ^3W_{\text{CH}}$ over $\ ^2\chi_C$. For example, $\ ^3W_{\text{CH}}$ is a more discriminatory descriptor than $\ ^2\chi_C$. All 157 alkanes considered possess different values of $\ ^3W_{\text{CH}}$. $\ ^2\chi_C$ is less effective than $\ ^3W_{\text{CH}}$ in discriminating between structurally very close isomeric alkanes. For example, 3-methylheptane and 4-methylheptane both have the same value of the connectivity index (3.808), but their shapes (I and II) are different and their I values are also slightly different (I, 772.6; II, 767.4).



Both $\ ^3W_{\text{CH}}$ and $\ ^2\chi_C$ model the shape and the size of alkanes [11,24-27] and they are highly intercorrelated indices. Hence both of these indices express similar constitutional information and the difference is related to the degree of topographic information. This result is perhaps indicative of why the $\ ^2\chi_C$ index does so well in the

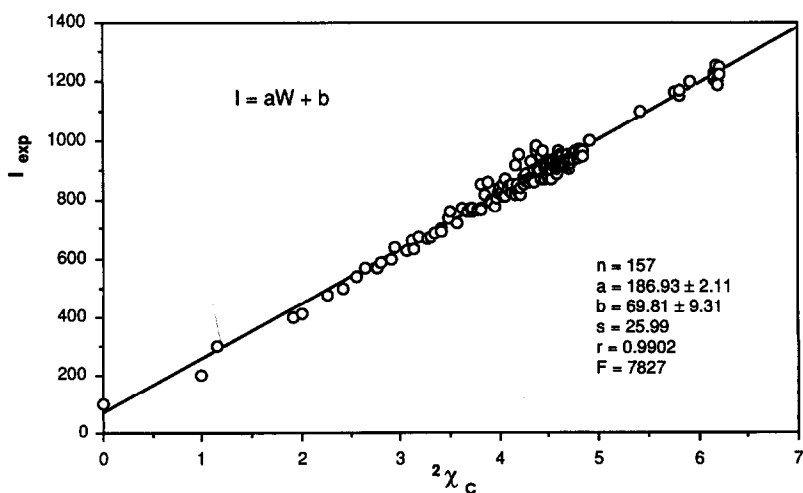


Fig. 3. Plot of I_{exp} vs. $\ ^2\chi_C$ for 157 alkanes with statistical characteristics.

QSCRR models, and even in the more general models such as in quantitative structure-activity relationships (QSAR) [28]. The connectivity index for the carbon skeleton of an alkane mimics very well the 3-D Wiener number for the molecule with all atoms present in it.

To conclude, we suggest that pragmatically speaking, the connectivity index has many advantages for use in the QSCRR models. However, the 3-D Wiener number is a physically more acceptable molecular descriptor as it reflects well the shape changes of alkanes and produces results of a quality comparable to that given by the connectivity index. In addition, its extension to more complex hydrocarbons and heteroatom systems is straightforward.

CONCLUSIONS

We have developed the 3-D Wiener number as a unique descriptor for alkanes. It was used to correlate with the GC retention indices of alkanes via a QSCRR model of the Walker type. The statistical characteristics of this model are of good quality.

The QSCRR- ${}^3W_{CH}$ model was compared with a model based on the 2-D connectivity index. The results indicate that this model is comparable to the well established QSCRR- ${}^2\chi_C$ model. Although the QSCRR- ${}^2\chi_C$ model is simpler, we suggest, however, that our approach be used in building up the QSCRR models because the ${}^3W_{CH}$ index is constructed from the optimum (minimum-energy) conformation of the molecule and it therefore reflects well the structure of the molecules, which appears to be one of the important factors responsible for molecular migration in the chromatographic process. Also, an advantage of using the ${}^3W_{CH}$ index is its straightforward extension to more complex types of molecules.

REFERENCES

- 1 R. Kaliszan, *Quantitative Structure Chromatographic Retention Relationships*, Wiley, New York, 1988.
- 2 N. Trinajstić, *Chemical Graph Theory*, Vol. II, CRC, Boca Raton, FL, 1983, ch. 4.
- 3 N. Turro, *Angew. Chem., Int. Ed. Engl.*, 25 (1986) 882.
- 4 M. Randić, in R. C. Lacher (Editor), *MATH/CHEM/COMP 1987*, Elsevier, Amsterdam, 1988, p. 101.
- 5 M. Randić, *J. Am. Chem. Soc.*, 97 (1975) 6609.
- 6 N. Bošnjak, N. Adler, M. Perić and N. Trinajstić, in Z. B. Maksić (Editor), *Modelling of Structures and Properties of Molecules*, Ellis Horwood, Chichester, 1987, p. 103.
- 7 H. Wiener, *J. Am. Chem. Soc.*, 69 (1947) 17.
- 8 H. Hosoya, *Bull. Chem. Soc. Jpn.*, 44 (1971) 2332.
- 9 F. Harary, *Graph Theory*, Addison-Wesley, Reading, MA, 1971, p. 203.
- 10 E. R. Canfield, R. W. Robinson and D. H. Rouvray, *J. Comput. Chem.*, 6 (1985) 598.
- 11 B. Bogdanov, S. Nikolić and N. Trinajstić, *J. Math. Chem.*, 3 (1989) 299.
- 12 U. Burkert and N. L. Allinger, *Molecular Mechanics*, American Chemical Society, Washington, DC, 1982.
- 13 W. R. Müller, K. Szymanski, J. V. Knop and N. Trinajstić, *J. Comput. Chem.*, 8 (1987) 170.
- 14 J. R. Chrétien and J.-E. Dubois, *J. Chromatogr.*, 126 (1976) 171.
- 15 G. Schomburg and G. Dielman, *J. Chromatogr. Sci.*, 11 (1973) 151.
- 16 J. Walker, *J. Chem. Soc.*, 65 (1894) 193.
- 17 J. Walker, *J. Chem. Soc.*, 65 (1894) 725.
- 18 R. Kaliszan and H. Foks, *Chromatographia*, 10 (1977) 346.
- 19 R. Kaliszan, *Chromatographia*, 10 (1977) 529.
- 20 M. Randić, *J. Chromatogr.*, 161 (1978) 1.
- 21 R. Kaliszan and H. Lamparczyk, *J. Chromatogr. Sci.*, 16 (1978) 246.

- 22 A. Sabljic, *J. Chromatogr.*, 314 (1984) 1.
- 23 A. Sabljic, *J. Chromatogr.*, 319 (1985) 1.
- 24 M. Randić, B. Jerman-Blažić and N. Trinajstić, *Comput. Chem.*, 14 (1990) 237.
- 25 D. H. Rouvray, *Sci. Am.*, 254 (1986) 40.
- 26 P. G. Seybold, M. May and U. A. Bagal, *J. Chem. Educ.*, 64 (1987) 575.
- 27 M. I. Stankevich, I. V. Stankevich and N. G. Zefirov, *Russ. Chem. Rev.*, 57 (1988) 191.
- 28 L. B. Kier and L. H. Hall, *Molecular Connectivity in Structure-Activity Analysis*, Research Studies Press, Chichester, 1986.