Wiener Index and Vibrational Energy

Ivan Gutman and Igor G. Zenkevich[a]

Faculty of Science, University of Kragujevac, P. O. Box 60, YU-34000 Kragujevac, Yugoslavia
[a] Chemical Research Institute, St. Petersburg State University,
Universitetsky pr. 26, St. Petersburg 198504, Russia

Reprint requests to Prof. I. G.; Fax: +381 34 335040, E-mail: gutman@knez.uis.kg.ac.yu

Z. Naturforsch. **57 a,** 824–828 (2002); received June 4, 2002

The Wiener index (W) is known to represent a rough measure of the surface area of molecules, which is used to explain why W is correlated with numerous physico-chemical properties of organic compounds. We now point out another aspect of W, unrelated to surface area, namely the relation between W and the internal (mainly vibrational) energy of organic molecules. For isomeric alkanes a theoretical justification of this result is offered. Another example – the dithioderivatives $C_3H_8S_2$ – indicates that this regularity may be applicable also to other groups of isomeric acyclic organic compounds.

Key words: Wiener Index; Intramolecular Energy; Vibrational Energy;
Acyclic Organic Compounds; Alkanes.

Introduction

The Wiener index (W) is one of the oldest molecular-graph-based structure-descriptors [1, 2] and its chemical applications are well documented [3 - 6]. If G is a molecular graph, u and v are two vertices of G, and $d(u, v|G)$ is their distance (= number of edges in a shortest path connecting u and v), then the Wiener index is defined as

$$W = W(G) = \sum_{u<v} d(u, v|G)$$

with the summation going over all pairs of vertices of G.

The fact that there are good correlations between W and a variety of physico-chemical properties of organic compounds (boiling point, heat of evaporation, heat of formation, chromatographic retention times, surface tension, vapor pressure, partition coefficients, etc.) could be rationalized by the assumption that W is roughly proportional to the van der Waals surface area of the respective molecule [7]. As a consequence, in the case of non-polar molecules (such as alkanes), W is proportional to the intermolecular forces [6, 8], implying relations between W and the intermolecular-force-dependent physico-chemical properties. This explanation, however, could not be employed for some other empirically established correlations, in particular those regarding the total electron energy of polymers [9 - 11], ultrasonic sound velocity [12], and stability constants of some complexes [13].

In this work we demonstrate that in the case of alkanes there is a relationship between the Wiener index and the dynamic parameters of the respective molecule. In particular, we establish an approximate linear dependence between W and the molecular vibrational energy, resulting in a linear correlation between W and the intramolecular energy.

Internal Energy of a Molecule and Its Evaluation

In a series of studies [14 - 18] one of the present authors discovered the existence of a linear relation between gas chromatographic (GC) retention indices (at least on standard non-polar stationary phases) and the intramolecular energies of the respective molecules. Such relations were observed for a variety of classes of organic compounds. Moreover, it was found that there exist mutual correlations between GC retention indices, boiling points, intramolecular energies and topological parameters (in particular, Wiener and Hosoya indices) [19]. Already these observations permit us to attribute new physico-chemical meaning to the topological indices – as parameters related to intramolecular energies.

Thus, GC retention indices could be predicted from the known values of the internal molecular energy (E_{int}). The calculation of the total intramolecular energy at a chosen temperature is based on the computer modeling of the intramolecular rotational and vibrational processes. (For this, the *HyperChem* software was found to be suitable; for details see [14 - 19] as well as the discussion below.)

The general methods of computing E_{int} are well known (see, for instance, [20]). On the other hand, obtaining reliable E_{int}-values requires very long computational times. Another problem is the existence of several stable conformers of the same compound with different intramolecular energies. In such cases, obtaining a single E_{int}-value with high precision becomes practically impossible.

In order to overcome these computational difficulties, a drastic, seemingly naive, yet appropriately chosen simplification was put forward [16 - 18]; its quality may be judged from Figure 1. The approximation is based on the fact that intramolecular vibrational energies (U_{vibr}) exceed many times the corresponding rotational energies (U_{rot}). It is important that U_{vibr} values, at least in the case of acyclic molecules, can be estimated by simpler methods (see below) than U_{rot} [21].

From another point of view, the total intramolecular energy E_{int} can be divided into two parts: a "variable" term U_{var} and a "constant" term U_{con}. The term U_{var} contains the energy contributions coming from (stretching) vibrations of carbon-carbon bonds. The term U_{con} contains the contributions of all other vibrational and rotational modes. The former is sensitive to the details of molecular structure and assumes different values for various isomers (see below). The latter one is nearly constant within a group of isomeric compounds.

Adopting the harmonic approximation [16, 19], the frequency of the (stretching) vibration of a carbon-carbon bond is given by [20]

$$\nu = \frac{1}{2\pi}\sqrt{\frac{k}{\mu}},$$

where k is the respective force constant and μ the reduced mass. Assuming further that all carbon-carbon bonds have nearly equal force constants (equal to k), one concludes that U_{var} is equal to

$$U = \sum \frac{1}{\sqrt{M_1\,M_2/M}} \tag{1}$$

times a pertinently chosen constant. The actual value of this constant is immaterial as far as correlations between U_{var} and the retention indices (or some other physico-chemical parameters) are concerned.

In formula (1) $M_1\,M_2/M$ is the reduced mass corresponding to the vibration of a particular carbon-carbon bond; M is the mass of the molecule considered, whereas M_1 and M_2 are the masses of the fragments on the two sides of the vibrating bond; $M_1 + M_2 = M$. Clearly, the masses M, M_1, and M_2 may be expressed in atomic mass units. The summation in (1) goes over all carbon-carbon bonds.

Connection between U and the Wiener Index

In the case of alkanes, (1) can be significantly simplified. Let H and C denote the relative atomic masses of hydrogen and carbon, respectively, $H = 1.0$, $C = 12.0$. (Within the present approach, using more accurate values for these parameters seems not to be necessary.) Let further n be the number of carbon atoms. Then

$$U = \sum \sqrt{\frac{(C+2H)n + 2H}{[(C+2H)n_1 + H][(C+2H)n_2 + H]}},$$

where n_1 and n_2 are the number of carbon atoms on the two sides of the vibrating carbon-carbon bond; $n_1 + n_2 = n$. (A graph-theoretical interpretation of the numbers n_1 and n_2 can be found in [22].) Direct multiplication yields

$$U = \sum \sqrt{\frac{(C+2H)n + 2H}{(C+2H)^2\,n_1\,n_2 + H(C+2H)n + H^2}}, \tag{2}$$

which should be compared with the well known relation [1, 22]

$$W = \sum n_1\,n_2. \tag{3}$$

In order to linearize the summand on the right-hand side of (2), note that the minimal and maximal values of the product $n_1\,n_2$ are $n-1$ and $(n/2)^2$, respectively. The middle between these extremes is $\xi = n^2/8 + (n-1)/2$. Therefore we expand the function $[(C +$

$2H)^2\, x + H(C+2H)n + H^2]^{-1/2}$ into a power series with respect to x, around the point $x = \xi$. This yields

$$[(C+2H)^2\, x + H(C+2H)n + H^2]^{-1/2}$$

$$= \frac{2}{Q} - \frac{4(C+2H)^2}{Q^3}\,(x-\xi) + \text{higher order terms},$$

where

$$Q = \left\{ \frac{1}{2}\,(C+2H)^2\, n^2 + 2(C+2H)(C+4H)n - 2(C^2 + 4\,H\,C + 2\,H^2) \right\}^{1/2}.$$

The above expression can be rewritten as

$$[(C+2H)^2\, x + H(C+2H)n + H^2]^{-1/2}$$

$$= \frac{3}{Q} - 4\,\frac{H(C+2H)n + H^2}{Q^3} - 4\,\frac{(C+2H)^2}{Q^3}\, n_1 n_2 + \text{higher order terms},$$

which substituted back into (2) and by taking into account (3) and the fact that there are $n-1$ carbon-carbon bonds, results in

$$U \approx A - B\,W, \tag{4}$$

where

$$A = \left[\frac{3}{Q} - 4\,\frac{H(C+2H)n + H^2}{Q^3} \right] \cdot (n-1)\,\sqrt{(C+2H)n + 2H},$$

$$B = 4\,\frac{(C+2H)^2}{Q^3}\,\sqrt{(C+2H)n + 2H}$$

or, by setting $H = 1$ and $C = 12$,

$$A = \left[\frac{3}{Q} - \frac{56n+4}{Q^3} \right] (n-1)\,\sqrt{14\,n + 2}, \tag{5}$$

$$B = \frac{784}{Q^3}\,\sqrt{14n + 2}, \tag{6}$$

$$Q = \sqrt{98\,n^2 + 448\,n - 388}.$$

Relation (4) should be understood as the main result of this paper. Its quality may be judged from Figure 3.

Table 1. Internal molecular energy (E_{int}, in kJ/mol) at 400 K, the auxiliary quantity U (defined via (1)), and the Wiener index W of the isomeric alkanes C_7H_{16}; for details see text.

Isomer	E_{int}	U	W
n-heptane	259.0 ± 7.5	1.405	56
2-methylhexane	275.3 ± 7.1	1.465	52
3-methylhexane	266.9 ± 7.5	1.482	50
2,2-dimethylpentane	273.6 ± 7.5	1.546	46
2,3-dimethylpentane	277.8 ± 7.5	1.546	46
2,4-dimethylpentane	275.7 ± 7.1	1.524	48
3,3-dimethylpentane	279.9 ± 7.5	1.561	44
3-ethylpentane	272.8 ± 7.1	1.501	48
2,2,3-trimethylbutane	284.9 ± 7.5	1.602	42

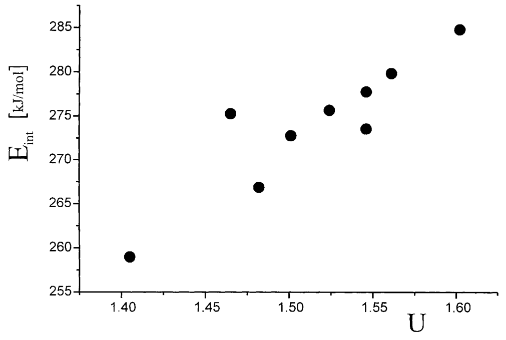

Fig. 1. Correlation between the internal energy (E_{int}) and its approximation U, (1), in the case of isomeric heptanes. Least-squares fitting gives $E_{\mathrm{int}} = (116 \pm 21)\,U + (99 \pm 31)$ with correlation coefficient $R = 0.903$. Note that (for reasons not fully understood) 2-methylhexane is an outlier; by leaving out this data point, the correlation would be significantly better ($R = 0.975$).

Relationships between U and W, of the form (4), were previously noticed for some groups of acyclic compounds [19]. By means of the considerations outlined in this section we provide a theoretical explanation of these empirical findings. In what follows we corroborate our results by additional examples.

Numerical Work and Details of Computer Modeling

In order to test the approximation (4) and the expressions (5) and (6) we have computed the intramolecular energies (E_{int} at 400 K) of all isomeric heptanes ($n = 7$), as well as their U and W values. Calculations of E_{int} have been carried out by means of the *HyperChem 5.1* software. The results obtained are given in Table 1. The quality of the correlation between the internal energy and U, as well as between

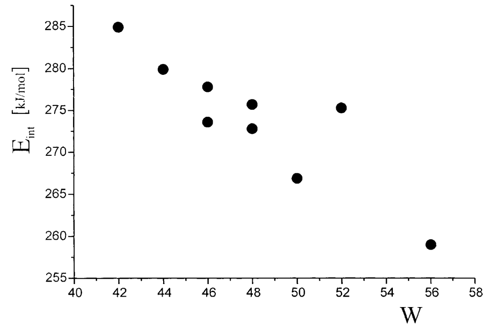

Fig. 2. Correlation between the internal energy (E_{int}) and the Wiener number W for isomeric heptanes. The respective regression line reads: $E_{\text{int}} = (-1.6 \pm 0.3)\,W + (349 \pm 15)$ with $R = -0.884$. Without the data point for 2-methylhexane one would have $R = -0.971$, cf. Figure 1.

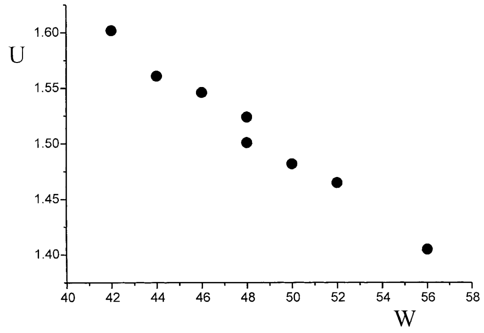

Fig. 3. Correlation between the molecular-graph-based structure-descriptors U and W in the case of isomeric heptanes, illustrating the precision of the approximation (4). The respective regression line reads: $U = (-0.01372 \pm 0.0007)\,W + (2.17 \pm 0.03)$ with $R = -0.992$.

E_{int} and W, and U and W, is seen from Figures 1, 2, and 3.

Note that by least-squares fitting of the data from Table 1 we get $U = 2.17 - 0.01372\,W$ (cf. Fig. 3), which is in surprisingly good agreement with $U = 2.04 - 0.01195\,W$ obtained by direct application of (5) and (6).

The same regularities are observed not only for isomeric alkanes, but also within different groups of isomeric organic compounds containing heteroatoms. For obvious reasons, a very large number of examples of this kind could be produced. In this paper we present only one: the isomeric dithiols and alkylthiothiols of the general formula $C_3H_8S_2$. Only five isomers were included in our sample, because ethylmethyl-disulfide $CH_3SSC_2H_5$ and di(methylthio)-methane $CH_3SCH_2SCH_3$ belong to another class of disulfides and should not be considered together with compounds possessing SH-groups. For results see Table 2; the respective regression lines read: $U = (-0.029 \pm 0.002)\,W + (1.44 \pm 0.03)$; $R = -0.995$ and $E_{\text{int}} = (-1.7 \pm 0.6)\,W + (178 \pm 13); R = -0.842$.

Table 2. Same data as in Table 1 for the isomeric dithioderivatives $C_3H_8S_2$.

Isomer	E_{int}	U	W
1,2-propanedithiol	146.0 ± 4.2	0.890	19
1,3-propanedithiol	140.6 ± 2.9	0.806	22
1-(methylthio)-ethanethiol	149.8 ± 2.5	0.952	17
2-(methylthio)-ethanethiol	143.5 ± 3.3	0.875	20
(ethylthio)-methanethiol	141.4 ± 3.3	0.898	19

Also in this example we find a very good correlation between U and W, corroborating our earlier conclusions. However, the correlation between E_{int} and W is somewhat weaker. A possible cause for this may be sought in the difficulties connected with the computation of dynamic properties of molecules possessing heteroatoms.

It is worth noting that the calculation of Wiener indices of organic compounds other than saturated hydrocarbons (possessing double or triple bonds, heteroatoms, etc.) usually requires some appropriate modification [19]. In particular, in the example considered above, every sulfur atom has been treated as being equivalent to two carbon atoms, resulting in the W-values reported in Table 2.

Concluding Remarks

By deducing the approximate expression (4) we have demonstrated a novel, hitherto unnoticed feature of the Wiener index: the quantity W (which in a very simple manner is computed from the molecular graph) reflects certain dynamic properties of organic molecules, at least in the case of alkanes. The finding that the Wiener index is related to the vibrational energy sheds a new light not only on the previous successful empirical applications of this half-century-old molecular structure descriptor, but will also stimulate further investigations.

The masses M_1 and M_2 (and therefore also the numbers n_1 and n_2) are well defined only in the case of acyclic molecules. Furthermore, the identity (3) plays a crucial role in our derivation of the relation (4), which connects U and W. The validity of the identity (3) is also restricted to acyclic molecules (or more precisely: to acyclic graphs). In view of this, the extension of the theory outlined in this work to cyclic molecules seems to be a significantly more difficult task. At the present moment, it is not known if such an extension is possible at all.

[1] H. Wiener, J. Amer. Chem. Soc. **69**, 17 (1947).

[2] R. Todeschini and V. Consonni, Handbook of Molecular Descriptors, Wiley-VCh, Weinheim 2000.

[3] Z. Mihalić and N. Trinajstić, J. Chem. Educ. **69**, 701 (1992).

[4] I. Gutman, Y. N. Yeh, S. L. Lee, and Y. L. Luo, Indian J. Chem. **32A**, 651 (1993).

[5] S. Nikolić, N. Trinajstić, and Z. Mihalić, Croat. Chem. Acta **68**, 105 (1995).

[6] I. Gutman and J. H. Potgieter, J. Serb. Chem. Soc. **62**, 185 (1997).

[7] I. Gutman and T. Körtvélyesi, Z. Naturforsch. **50a**, 669 (1995).

[8] I. Gutman and J. H. Potgieter, South Afr. J. Sci. **92**, 47 (1996).

[9] D. Bonchev and O. Mekenyan, Z. Naturforsch. **35a**, 739 (1980).

[10] D. Bonchev, O. Mekenyan, and O. E. Polansky, Z. Naturforsch. **36a**, 643 (1981).

[11] D. Bonchev, O. Mekenyan, and O. E. Polansky, Z. Naturforsch. **36a**, 647 (1981).

[12] D. H. Rouvray and W. Tatong, Z. Naturforsch. **41a**, 1238 (1986).

[13] S. Nikolić and N. Raos, Croat. Chem. Acta **74**, 621 (2001).

[14] I. G. Zenkevich, Zh. Org. Khim. **34**, 1463 (1998).

[15] I. G. Zenkevich, Fresenius J. Anal. Chem. **365**, 305 (1999).

[16] I. G. Zenkevich, Rus. J. Phys. Chem. **73**, 797 (1999).

[17] I. G. Zenkevich, Zh. Anal. Khim. **55**, 1091 (2000).

[18] I. G. Zenkevich, Zh. Org. Khim. **37**, 283 (2001).

[19] I. G. Zenkevich and A. N. Marinichev, Zh. Struct. Khim. **42**, 893 (2001).

[20] D. A. McQuarrie and J. D. Simon, Physical Chemistry: A Molecular Approach, University Science, Sausalito (Calif.), 1997.

[21] W. J. Orville-Thomas (Ed.), Internal Rotation in Molecules, Wiley 1974.

[22] I. Gutman and O. E. Polansky, Mathematical Concepts in Organic Chemistry, Springer-Verlag, Berlin 1986, p. 127.