

## Molecular Branching Topology & Quantum Mechanical Quantities

S C RAKSHIT\*, M BANERJEE & BIREN HAZRA

Chemistry Department, University of Burdwan, Burdwan 713 104

Received 12 March 1987; revised and accepted 17 July 1987

Topological branching indices of *n*-alkanes and their isomers have been shown to lead to standard curves when some quantum mechanical properties, such as binding energies, two-electron exchange integrals (CNDO), charge densities, etc are plotted against them. These curves can be used to obtain the values of molecular level quantities for an unexplored alkane without resorting to quantum treatment. Either the graphs or the regression equations may be used for this purpose. This method has been extended to calculate the quantum mechanical properties of some primary alcohols, carboxylic acids and silanes also. Tentative reasons have been suggested for the link between the specific quantum quantities and the branching indices. The validity of the treatment, when the alkanes are substituted by alcohols or carboxylic acids, has also been justified.

The degree of molecular branching in the members of a homologous series of a general formula or in the isomers of a given molecular composition is a topic of considerable importance in graph theory. Different methods have been employed to set up measures, called branching indices<sup>1-4</sup>, using the topological parameters of the molecular graphs. Randic<sup>4</sup> has shown that a function of the degrees of vertices of the molecular graph of the straight-chain and branched alkane isomers produces a numerical order that runs parallel to an ordering of the alkanes based on their topological matrices when the vertex labelling is done in a unique way. This unique labelling can be done either by a few trials for not too big molecules or by use of a computer programme, developed by Randic<sup>5</sup>, for bigger molecules.

Bonchev and Trinajstic<sup>3</sup> have used information theory in conjunction with distance matrices of the molecules to set up information theoretic graphical indices of molecular branching. Besides these two, a few more branching indices have been proposed.

It has been shown<sup>4</sup> that the experimental values of many physical and thermodynamic properties, such as melting points, boiling points, vapour pressure, enthalpies of formation, etc., when plotted against Randic branching indices (RBI) of the alkane isomers, either give a smooth curve or, in some cases, even yield linear graphs. We thought it worthwhile to explore the relation between the branching indices proposed by Randic (RBI) and Bonchev and Trinajstic (BTBI) and some molecular level properties.

### Theoretical

Interestingly enough, the response was found to be positive in cases of some important quantum-mechanically derived quantities like binding energies

and diatomic electron-electron exchange energy of CNDO molecular orbital method.

The binding energies of molecules and ions in the singlet sigma state may be calculated with a satisfactory accuracy by a semiclassical method<sup>6</sup>. This semiclassical expression is virtually the CNDO expression for energy in which some parameters of the latter are substituted by their effective classical analogues. A further improvement in accuracy has been made possible by a recent modification<sup>7</sup> of  $k_n$  of the original formulation<sup>6</sup>. The binding energy (in eV) is given by Eq. 1,

$$E_{\text{bin}} = \frac{1}{2} \left[ \left\{ \sum_i \frac{q_i^2}{2r_i} + \sum_{i < j} \frac{q_i q_j}{r_{ij}} \right\} - \sum_{A < B} \left\{ \frac{1}{2} (1 + k_n n) \sum_{\mu}^A \sum_{\nu}^B P_{\mu\nu}^2 \gamma_{AB} \right\} \pm \Delta U_D \right] \times 27.2 \quad \dots (1)$$

where the symbols have their usual significances as given by Pople<sup>8</sup> in his original valence shell basis treatment.

Eq. (1) may be written as<sup>6</sup>,

$$E_{\text{bin}}(\text{eV}) = \frac{1}{2} (Q + E_{\text{ex}} + k_n n \cdot E_{\text{ex}} \pm 0.05) \times 27.2 \quad \dots (2)$$

where  $Q$  is the electrostatic work of charging and assembling the atoms as these exist in the molecule and is calculated from CNDO charge distribution data.  $E_{\text{ex}}$  is the total diatomic electron-electron exchange integral, 'n' represents the number of valence shell ba-

sis orbitals and  $k_n$ , a calibrated quantity connected with resonance energy, is obtained from Eq. 3,

$$k_n = A + (B/n) - \lambda \ln(n) \quad \dots (3)$$

The topological branching indices of Randić<sup>4</sup> and of Bonchev and Trinajstić<sup>3</sup> are given by (for alkanes) Eqs 4 and 5, respectively,

$$\text{RBI} = \sum_{i < j} (v_i v_j)^{-1/2}, \text{ for all } ij \text{th bonds} \quad \dots (4)$$

where  $v_i$  and  $v_j$  are the degrees of vertices in the carbon-carbon chain without the H-attaching bonds in the molecular graphs.

$$\text{BTBI} = \bar{I}_D = - \sum \frac{2k_i}{N(N-1)} \log_2 \frac{2k_i}{N(N-1)} \quad \dots (5)$$

where  $\bar{I}_D$  = mean information content for a distance in the distance matrix for alkane chain,  $k_i$  = the number or times a particular distance 'i' occurs in the probability scheme based on the distance matrix and  $N^2$  equals the total number of elements in the matrix.

## Results and Discussion

When the molecular level properties of Eq. (2) are plotted against RBI values of a homologous series of alkanes, interesting graphs are obtained. Extremely good straight lines result for plots of (i) binding energies (ii) total diatomic electron exchange energies and (iii) quantity Q of Eq. (2) (Fig. 1). The plots of binding energies against RBI for a set of alkane isomers are also linear (Fig. 2).

The linear graphs (drawn for the first few members) offer the unique advantage of serving as standard curves for predicting the binding energies of unexplored alkanes and their two electron diatomic exchange energies obviating the need of onerous molecular orbital computation. For the sake of graphic demonstration, quantum mechanical value of binding energy of an arbitrarily chosen alkane (*n*-octane) is indicated by the point P (Fig. 1). The binding energy and the diatomic exchange energy of *n*-hexane of the  $C_6$ -alkane isomer series are denoted by the points P and R (Figs 2a, 2b) respectively. However, to demonstrate the general predictability for the alkanes, a comprehensive Table 1 is provided, the data of which are based on the regression equation of the method of least squares, with correlation coefficient  $r = 0.9995$ . The equation is,

$$y = 24.413x + 6.425 \quad \dots (a)$$

where  $x$  represents the Randić branching index and  $y$ , the binding energy.

The binding energy value for the alkane *n*-eicosane involving a basis set of 122 atomic orbitals (much beyond the limit of the present day programmes avail-

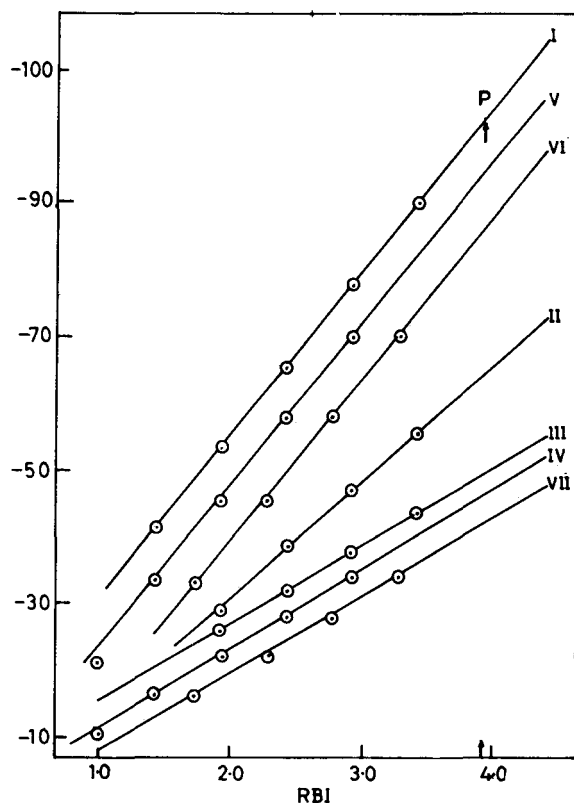


Fig. 1—Plots of calculated molecular level properties against RBI values. [Curves I, V and VI: B.E. of *n*-alkanes, alcohols and carboxylic acids respectively; curves III, IV and VII:  $10^3 E_{ex}$  of *n*-alkanes, alcohols and carboxylic acids respectively; curve II:  $10^{-5}$  Q of *n*-alkanes; prediction for point P: *n*-octane (RBI = 3.914), lit. B.E. (eV) = -102.43, graphical B.E. (eV) = -101.98].

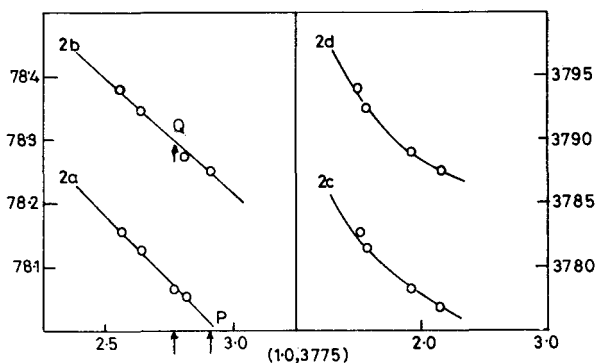


Fig. 2—Plots of calculated molecular level properties of *n*-hexane and its isomers against RBI and BTBI [Curve 2a: -B.E. vs RBI; curve 2b:  $-10^3 E_{ex}$  vs RBI; curve 2c: -B.E. vs BTBI; curve 2d:  $-10^3 E_{ex}$  vs BTBI; prediction for point P (hexane): lit B.E. = -78.09 eV, graphical B.E. = -78.02 eV; prediction for point R (2-methylpentane): calc.  $E_{ex} = -3788 \times 10^{-3}$  a.u., graphical  $E_{ex} = -3789 \times 10^{-3}$  a.u.]

able for 80 orbital basis set of the conventional CNDO technique) has been elegantly obtained by the use of the regression equation (Table 1). The importance of the graphical method, apart from its simplicity, is further emphasised by the fact that for calculat-

Table 1—Values of Binding Energies for *n*-Alkanes

Molecule ( <i>n</i> -Alkane)	Randic branching indices	Calc. B.E. (eV) using Eq. 2	Value of B.E. (eV) predicted* from regression Eq. (a)	Thermo- dynamic B.E. (eV)	Standard error
Ethane	1.00	-29.68	-30.839	-29.33	
Propane	1.414	-41.28	-40.95	-41.50	
<i>n</i> -Butane	1.914	-53.62	-53.158	-53.69	
<i>n</i> -Pentane	2.414	-66.09	-65.365	-65.887	
<i>n</i> -Hexane	2.914	-78.033	-77.570	-78.082	
<i>n</i> -Heptane	3.414	-89.95	-89.778	-90.255	
<i>n</i> -Octane	3.914	-100.998	-101.98	-102.43	
<i>n</i> -Nonane	4.414	—	-114.19*	-114.62	
<i>n</i> -Decane	4.914	—	-126.39*	-126.80	
<i>n</i> -Undecane	5.414	—	-138.60*	-139.03	0.697
<i>n</i> -Dodacane	5.914	—	-150.81*	-151.17	
<i>n</i> -Tridecane	6.414	—	-163.02*	-163.35	
<i>n</i> -Tetradecane	6.914	—	-175.22*	-175.53	
<i>n</i> -Pentadecane	7.414	—	-187.43*	-187.72	
<i>n</i> -Hexadecane	7.914	—	-199.64*	-199.90	
<i>n</i> -Heptadecane	8.414	—	-211.84*	-212.08	
<i>n</i> -Octadecane	8.914	—	-224.05*	-224.27	
<i>n</i> -Nonadecane	9.414	—	-236.26*	-236.45	
<i>n</i> -Eicosane	9.914	—	-248.46*	-248.63	

ing binding energy of even *n*-heptane, with just 44 atomic orbitals as the basis, around 15 min of a fairly fast computer are required. It is easy to comprehend the time and cost involved in CNDO calculations when 122 basis orbitals are involved, as in the case of *n*-eicosane.

In Fig. (3) are displayed the plots of the B.E.,  $10 E_{ex}$  and Q against BTBI (relation-5). In all cases (except for Q) the curves are smooth, though not always linear. Although extrapolation is permissible for the unexplored alkanes, RBI linear curves decidedly have an advantage over the BTBI curves in this respect.

This convenient way of obtaining quantum mechanical quantities of unexplored alkanes and their isomers from standard curves can be extended to domains of some other homologous series deemed to be derived from the alkanes. The homologous series of the normal aliphatic alcohols ( $RCH_2OH$ ) and carboxylic acids ( $RCOOH$ ) have been chosen for this study and both quantum mechanical calculations have been done and graphs have been drawn for them. The rationale behind this extension is explained below.

The molecular graph in a normal alkane is of the type  $\text{---}\text{---}\text{---}$ , where the vertices (all carbon atoms) are of degrees either 2 or 1. The terminal vertices have the degree value 1. If one such terminal atom be replaced by O (as in the formation of alcohol), this new vertex may be considered as that of a pseudo C-atom having the unaltered degree 1. Hence, the branching index (RBI) of the new chain remains the same as th

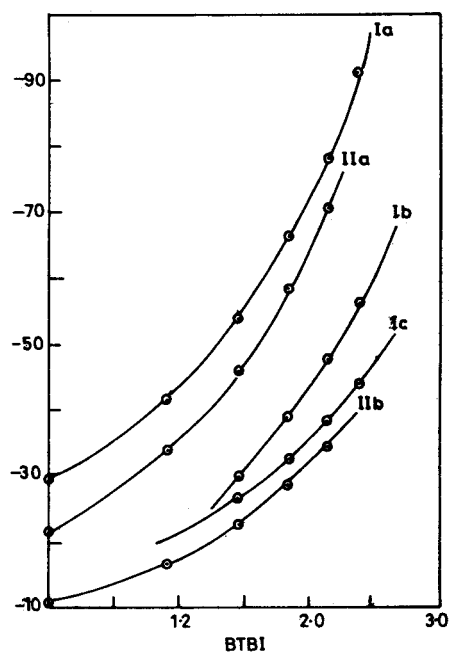


Fig. 3—Plots of calculated molecular level properties against BTBI [Curve Ia: B.E. of *n*-alkanes; curve IIa: B.E. of alcohols; curve Ic:  $10 E_{ex}$  of *n*-alkanes; curve IIb:  $10 E_{ex}$  of alcohols; curve Ib:  $10^{-5} Q$  of *n*-alkanes].

of the original alkane. The BTBI branching index, based on information theoretic graphical scheme, should also remain unchanged if the O-atom plays the role of a pseudo C-atom. It is in this perspective that the previous molecular level properties (viz. binding

energies, etc.) are plotted as functions of the RBI or BTBI.

The results obtained for primary alcohols and the carboxylic acids are represented graphically in Figs 1 and 3. All the curves faithfully mimic the pattern of the corresponding *n*-alkane series. The linear curves, as in the earlier cases, serve as standards for evaluating the corresponding molecular level properties. Alternatively, one may use the regression equations (b) and (c) to predict the values of the binding energies of the primary alcohols and of the carboxylic acids, respectively. Tables 2 and 3 exhibit such data with the correlation factors, the regression equations appearing at the head of each table.

While comparing the data predicted for the alcohols and the acids with the thermodynamic data, one ought to remember that the latter intrinsically include the effect of the possible polymeric forms of the species due to hydrogen bonding in the vapour state<sup>11</sup>. However, as the percentage of the polymeric species is expected to be rather low, the effect on the overall thermodynamic value is negligible. The binding ener-

gy calculated with the help of Eq. (2) refers to that of the single species in the vapour state.

While the potential uses of such graphs are evident, a suitable explanation is needed to account for the linear variation of the quantum molecular properties with the topology based property, viz., the branching index. A rigorous explanation cannot be offered at this stage, but plausible reasons for linearity are not difficult to divine.

Actual computation shows that the main contribution to  $E_{bin}$  in Eq. (2) comes from  $E_{ex}$  which essentially involves a series of integrals over diatomic overlap functions and distance parameter  $r_{\mu\nu}$ . The binding energy of an alkane may be thought to be proportional to the total overlap integral comprising a series of definite integrals representing  $\Phi_{Csp^3} - \Phi_{Csp^3}$  overlaps and  $\Phi_{Csp^3} - \Phi_{H1s}$  overlaps. That is,

$$E_{bin} = k[S_{(n-alkane)}] \\ = k[(n'-1)S_{C-C} + (2n'+2)S_{C-H}] \quad \dots (6)$$

(where  $n'$  = number of carbon atoms)

Table 2—Binding Energies of Primary Alcohols (correlation coefficient,  $r = 0.9996$ ) Obtained by the Regression Eq.  $y = 24.998x - 2.712 \dots (b)$

Molecule	Branching index (Randic)	Calc. value of binding energy (eV) Eq. 2	Value of B.E. (eV) predicted from regression Eq. (b)	Thermodynamic value of the heat of atomisation (eV)	Standard error
Methanol	1.00	-21.32	-22.286	21.14	
Ethanol	1.4142	-33.51	-32.64	33.46	
<i>n</i> -Propanol	1.9142	-45.361	-45.139	45.67	
<i>n</i> -Butanol	2.4142	-57.927	-57.638	57.92	0.570
<i>n</i> -Pentanol	2.9142	-70.08	-70.137	70.229	
<i>n</i> -Hexanol	3.4142	-82.28	-82.636	—	
<i>n</i> -Heptanol	3.9142	—	-95.135	95.495	
<i>n</i> -Octanol	4.4142	—	-107.634	107.77	

Table 3—Binding Energies of Carboxylic Acids (correlation coeff. = 0.9997); Obtained by the Regression Eq.  $y = 24.249x - 8.967 \dots (c)$

Molecule	Branching index (Randic)	Calc. value of B.E. (eV) Eq. 2	Value of B.E. (eV) predicted from regression Eq. (c)	Thermodynamic value of the heat of atomisation (eV)	Standard error
Acetic acid	1.7321	-33.363	-33.034	33.61	
Propionic acid	2.2701	-45.48	-46.08	45.76	
Butyric acid	2.7702	-58.149	-58.207	58.44	
Valeric acid	3.270	-70.905	-70.327	70.71	
Caproic acid	3.7701	-82.206	-82.45	—	0.416
Heptylic acid	4.270	—	-94.576	—	
Caprylic acid	4.770	—	-106.70	—	
Pelargonic acid (Nonanoic)	5.27	—	-118.82	—	
Capric acid	5.77	—	-130.94	131.64	

The overlap integral involving two, *i*th and *j*th carbon atoms of degrees 1 and 2, respectively, will have an  $S_{C-C}$ , deemed to be a geometric mean (positive) of two real number contributions,

$$(S_{C-\frac{1}{2}S_C})^{1/2} = (1.2)^{-1/2} S_C = (v_i v_j)^{-1/2} S_C$$

related to the graphical degree values as in RBI. Similarly if *k* and *l* atoms be of degrees 2 and 3 respectively (as is possible in a branched hydrocarbon containing tertiary C-atoms) each,

$$S_{C-C} = (\frac{1}{2} S_C \frac{1}{3} S_C)^{1/2} = (2.3)^{-1/2} S_C = (v_k v_l)^{-1/2} S_C$$

This equality to geometric mean is based on an assumption that carbon atoms of topological degrees 1, 2 and 3 respectively, contribute an effective multiple of  $S_C$ , each consistent with the corresponding connectivity number. It implicitly neglects the overlap contributions coming from the classically unconnected C-atoms, which are really small, notwithstanding the fact that the exchange integral involves all possible pairs of orbitals.

The remaining part of contribution to  $E_{bin}$  due to all  $S_{C-H}$ 's (Eq. 6) increases additively by a constant amount as, *n'*, the number of C-atoms in the chain, increases.

It is thus possible to conceive of a linear dependence of the exchange integrals and binding energies upon the RBI index values and, in practice, this turns out to be the case (Fig. 1).

The reason why such quantum properties fail to show linear dependence on BTBI is that the latter branching index, unlike RBI, is not merely a sum of two vertex properties, but depends on a more elaborate probability scheme.

The reasons for linear variation of the binding energies with the branching indices (RBI) of the alcohol set is easily understood if we accept the truth of the Eq. (6). Taking the diatomic exchange energies to be the determinant and assuming that such exchange energies are proportional to diatomic overlap integrals, the binding energies of alcohols derived from the corresponding alkanes can be written, following Eq. (6), as

$$E_{bin} = k[(n'' - 1)S_{C-C} + S_{C-O} + (2n'' + 1)S_{C-H} + S_{O-H}] \quad \dots (7)$$

where,  $n'' + 1 = n' =$  total number of C-plus O-atoms in the primary alcohol. The difference of Eqs (6) and (7) leads to,

$$E_{bin(alkane)} - E_{bis(alcohol)} = k[(S_{C-C} - S_{C-O}) + (3S_{C-H} - S_{O-H})] \quad \dots (8)$$

for all *n'*-alkanes and the corresponding alcohols. This means a constant difference,  $\Delta E_{bin}$ , between an alkane and the alcohol irrespective of the value of *n'*. We ought to expect not only linearity for the alcohol graphs but also parallel runs for the sets of alkanes and alcohols (cf. Fig. 1, curves I and V, and curves III and IV).

Referring to Mulliken *et al.*<sup>9</sup>, data on diatomic overlap enable us to calculate  $(S_{C-C} - S_{C-O})$  and  $(3S_{C-H} - S_{O-H})$  of Eq. (8) and hence *k* from the known values of binding energies of the LHS.

Keeping in view the increasing attention received by silanes and their derivatives and the fact that molecular orbital studies on silanes are rather limited, we have calculated the binding energies of the first four normal silanes<sup>12</sup> by the present method. Following the procedure adopted for the alkane series and including the *d*-orbitals of the silicon atom in the basis set for the CNDO charge and exchange energy computation, we have obtained binding energies of silanes (Si 1 to 4):

Molecule	Branching index (Radic)	Binding energy (eV)
SiH <sub>4</sub>	0	-10.73
Si <sub>2</sub> H <sub>6</sub>	1.00	-21.067
Si <sub>3</sub> H <sub>8</sub>	1.4142	-30.11
Si <sub>4</sub> H <sub>10</sub>	1.9142	-38.256

As in the case of alkanes, the binding energies of the silanes, quite expectedly, turn out to be a linear function of the branching indices.

## References

- Wiener H, *J Am chem Soc*, **69** (1947) 17; **69** (1947) 2636; *J phys Chem*, **52** (1948) 425.
- Hosoya H, *Bull chem Soc Jpn*, **44** (1971) 2332.
- Bonchev D & Trinajstic N, *J chem Phys*, **67** (1977) 4517.
- Radic M, *J Am chem Soc*, **97** (1975) 6609.
- Radic M, *J chem Phys*, **60** (1974) 3920.
- Rakshit S C & Mukherjee Bani, *Indian J Chem*, **22A** (1983) 273.
- Rakshit S C & Hazra B, *Indian J Chem* (communicated).
- Pople J A & Segal G A, *J chem Phys*, **43** (1965)S 136; **44** (1966) 3289.
- Mulliken R S, Rieke C A, Orloff D & Orloff H, *J chem Phys*, **17** (1949) 1248.
- Pople J A & Beveridge D L, *Approximate molecular orbital theory* (McGraw Hill, NY) 1970.
- Pauling Linus, *The nature of the chemical bond* (Cornell University Press, Ithaca) 1967.
- Cotton F A, *Advanced inorganic chemistry* (Wiley Eastern, New Delhi) 1969.