Linear Correlation between Properties and Graph Theoretical Indices of Homologous Molecules

S. C. RAKSHIT* and BIREN HAZRA

Chemistry Department, Burdwan University, Burdwan-713 104

The excellent linear relations between sets of specific quantum mechanical quantities, certain molecular properties and selective topological indices have previously been explained on the basis of the additive nature of both properties and the indices. Such explanations, based on parallelism, appear to be extrinsic. A more intrinsic explanation having its root in the tridiagonalities of the matrix forms of the involved entities seems to be a more general and unified interpretation. Besides new interrelations are also derived for these quantities.

Ever since graph theory was brought to the doorway of chemistry, it was shown that various topological indices could be divined or contrived which would correlate with different physical properties of such molecules. Molecules belonging to homologous series are specially adapted for such correlations. Certain molecular properties, quantum mechanical quantities and thermodynamic functions that are additive in character show excellent linear dependence on suitably chosen topological invariants of such molecules[1-4].

No basic interpretation, apart from the conceivable additivity of definite and constant contributions for each addition of $-CH_2-$ group to both the molecular property and the corresponding topological index, has been offered. Considering the diversity and the gamut of such quantities, a deeper level unifying interpretation is overdue. A previous attempt in this direction by us[4] resulted in an extrinsic explanation only. It hinged on an analysis of the theoretical structures of the quantum mechanical quantities and molecular (also thermodynamic) properties and, incidentally, also of the topological invariants (viz. Z = half the modular sum of eigenvalues, N = number of vertices in the graph, RBI = Randic branching index) to show how constant lumps of contributions arise from each unit addition of $-CH_2-$ group. Although the definite contributions were thus established mathematically and piecewise for each property, quantum mechanical quantity and Z, N, RBI, the process still lacks the aura of a unified and intrinsic explanation.

A truly topological graph of a molecule consisting of vertices and connectivities only is such a high degree of abstraction that hardly any physics or chemistry is left in it. Consequently, in any sort of interpretation of correlation involving the pure molecular topology and its matrix, it is futile to search for any physical or chemical link as the basic underlying cause of correlation. Expectedly, therefore, the topological matrix or some matrix property becomes our natural suspect as the possible cause. Proceeding along this line, we may arrive at what may be termed as an intrinsic explanation for the host of observed linear correlations[1,3,4]. The quantities presented in Table 1, amongst many others, show intra- and inter-correlations in the forms of excellent linear graphs[4].

It is known that for conjugated hydrocarbons, the modular sum of energy eigenvalues $\sum |\varepsilon_1|$ is given by the modular sum of their topological eigenvalues, called the energy of the graph E_g, which according to McClelland[7], is

$$E_g = \sum_1 |\lambda_1| = \sum_1 |\varepsilon_1| = 0.91 \sqrt{2MN} \qquad (1)$$

where M an N represent the number of edges and vertices respectively, in the topological graph. For the hydrogen suppressed graphs (HSG) of the normal alkanes, equation (1) will still be valid although λ_1's are merely eigenvalues and not identical with energy eigenvalues of the alkanes,

$$2Z = \sum_1 |\lambda_1| = E_g \simeq 0.91 \sqrt{2MN} \qquad (2)$$

TABLE 1

Topological quantities	Quantum mechanical quantities	Molecular/Thermodynamic properties
1. RBI (branching index)	1. E_{ex} = Total two-electron exchange energy (CNDO/2)	1. Molar diamagnetic susceptibility χ_m
2. Half the modular sum of topological eigenvalues $= (1/2) \sum_i \| \lambda_i \|$	2. E_b = Binding energy[5]	2. Enthalpy of formation
3. N = Number of vertices in a topological graph	3. $\sum S_{ab}$ = Total adjacent diatomic overlap[6]	3. Entropy of formation
		4. Enthalpy of atomisation

To understand the excellent linear relationships, let us first pick the pair E_{ex} and Z, the total diatomic exchange integral (CNDO) and half the modular sum of topological eigenvalues respectively of the alkanes.

As an illustration, we choose n-hexane. The diatomic-exchange integral (d.e.i.) matrix $\mathbf{D}$ is (20×20) dimensional and its topological matrix a (6×6) one. In order that some comparison be feasible between the two, let us agree to write a condensed form of d.e.i. matrix (henceforth, abbreviated as c.d.e.i. matrix) to reduce it to a (6×6) dimension. In achieving this end, we divide the atoms of hexane into six groups CH_3 (gr. 1), CH_2 (gr. 2), CH_2 (gr. 3) ... CH_3 (gr. 6) and write the c.d.e.i. matrix in such a way that each diagonal element of the matrix is a group sum of original d.e.i. matrix elements of the atoms pertaining to a group and that an off-diagonal matrix element of c.d.e.i. matrix represents a sum of d.e.i. matrix elements of the atoms pertaining to two different groups (say gr. 1 and gr. 2 so on). The $\mathbf{D}^{(cond)}$ matrix elements thus retain the basic information of the d.e.i. matrix elements only in the form of their sums over atoms of a group and not in detail relating to each pair of atoms.

We may write down specifically the c.d.e.i. matrix computed by us and the topological matrix of n-hexane to help clarify our arguments,

$\mathbf{D}^{(cond)}$ = c.d.e.i. matrix of n-hexane

$$= \begin{pmatrix} -0.63 & -0.17 & 0.0 & 0.0 & 0.0 & 0.0 \\ -0.17 & -0.42 & -0.17 & 0.0 & 0.0 & 0.0 \\ 0.0 & -0.17 & -0.42 & -0.17 & 0.0 & 0.0 \\ 0.0 & 0.0 & -0.17 & -0.42 & -0.17 & 0.0 \\ 0.0 & 0.0 & 0.0 & -0.17 & -0.42 & -0.17 \\ 0.0 & 0.0 & 0.0 & 0.0 & -0.17 & -0.63 \end{pmatrix} \quad (3)$$

$\mathbf{T}$ = Topological matrix of n-hexane

$$= \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

All exchange integrals between any two classically non-bonded atoms are, as actual CNDO/2 calculations show, very small compared to those of the bonded ones and are thus omitted in writing the $\mathbf{D}^{(cond)}$ matrix.

Let $\mathbf{D}^{(cond)}$ represent the two-centre (CNDO/2) c.d.e.i. matrix and $\mathbf{T}$, the topological matrix of any particular member with N vertices of a homologous series, say, alkanes. Then, in general $\mathbf{D}^{(cond)} = a\,\mathbf{T} + b\,\mathbf{1} + \mathbf{M}$; $\mathbf{M}$ is a matrix where all the elements are zero except the (11)th. and (NN)th. elements, each of the latter being equal to a C–H exchange energy (–0.21 a.u.). $\mathbf{1}$ is unit matrix $[(\mathbf{1})_{ii} = 1]$ and a and b are two numerical constants [for any alkane $a = -0.17$ (a.u.) = C–C exchange energy and $b = -0.21 \times 2$ (a.u.) = two times the C–H exchange energy; vide Table 2].

At this point we diversify our treatment under two heads, viz. (A) cause of linear relationships and (B) discovery of further correlations.

(A) Intrinsic cause of linear graphs :

The total diatomic exchange E_{ex} (CNDO) is the sum of all the matrix elements in $\mathbf{D}^{(cond)}$ which is *tridiagonal* in form exactly similar to its topological matrix $\mathbf{T}$. The relation, viz.

$$E_{ex} = N\,b + 2\,(N-1)a + 2\,(b/2)$$
$$= N\,b + [2(N-1)a + b] \quad (5)$$

TABLE 2 – MATRIX ELEMENTS OF D'_n AND T' (ALKANES)

Molecule	d'_n	t'_n
n-Hexane	–0.113 7	–1.801 9
	–0.208 0	–1.247 0
	–0.344 3	–0.445 0
	–0.495 7	0 445 0
	–0.632 0	1.247 0
	–0.726 3	1.801 9
n-Pentane	–0.125 6	–1.732 1
	–0.250	–1.00
	–0.420	0.0
	–0.590	1.0
	–0.714 4	1.732 1
n-Butane	–0 144 9	–1.618 0
	–0.314 9	–0.618 0
	–0.525 1	0.618 0
	–0.695 6	1.618 0
Propane	–0.179 6	–1.414 2
	–0.420	0.0
	–0.660 4	1.414 2

is easily scooped from the tridiagonal $D^{(cond)}$ matrix of an alkane with a topology of N vertices. The two (h/2) terms are the (11)th. and (NN)th. matrix elements of the M matrix. Since h is a constant (= –0.42 a.u.) for the alkane series representing twice the value of a C–H exchange energy, the linearity of E_{ex} on N is at once evident. Again since $2 Z = 0.91 \sqrt{2 M N}$, (cf. eq. 2) and $M = N - 1$ for a HSG of the alkane series

$$Z \simeq 0.91 \left(\frac{\sqrt{2N^2}}{2} \right) \simeq \frac{0.91\,N}{\sqrt{2}} \qquad (6)$$

whenever $N^2 \gg N$, a condition that is readily attained for $N = 3$ or more. Recasting equation (5) with the help of equation (6), where $c = \sqrt{2}/0.91$, we have

$$E_{ex} = cZb + 2\,(cZ-1)\,a + b \qquad (7)$$

This accounts for the linear dependence of E_{ex} on Z.

The intrinsic explanation thus rests on tridiagonalities of both the T and $D^{(cond)}$ matrices. A tridiagonal T essentially means adjacent atom connectivities. The E_{ex}, being amenable to be ex-

pressed as tridiagonal $D^{(cond)}$ matrix, virtually means that only the exchange integral elements associated with two bonded atoms are significant, the non-bonded atom contributions are effectively zero. It is this common feature, viz. the importance of the adjacency factor, that sneaks into the domain of E_{ex} and helps in the emergences of equations (5) and (7).

We may now consider other members of the intra- and inter-relationships, viz. E_b, $\sum_{a<b} S_{ab}$ (adjacent atoms), RBI and χ_m mentioned earlier. Since $E_b = (1/2) (1 + n\,k_n)\,E_{ex}$ (cf. Ref. 5) and E_{ex} is topology dependent (eqs. 5 and 7), E_b will be linearly dependent on both N and Z for the alkane series. Actual computations of S_{ab} for non-bonded atoms[4,6] in alkane series show these to be negligible compared to the values of the bonded ones. Hence, an $S^{(cond)}$ matrix, tridiagonal in form and dimensionally the same as the tridiagonal topological matrix of HSG, can be written down as

$$S^{(cond)} = c\,T + d\,1 + P \qquad (8)$$

where, $c = 0.65$, $d = 1.38$ and the P matrix has the elements $P_{11} = P_{NN} = (d/2)$ and others, zero. Once again noting its similarity with $D^{(cond)}$ matrix and employing similar arguments it can be established that $\sum S_{ab}$ is a linear function of both N and Z, the topological invariants.

The Randic branching index matrix R with RBI = $\sum_{i<j} (v_i\,v_j)^{(1/2)}$ can directly be built up. For n-hexane (HSG) it will have the tridiagonal structure

$$R = \begin{pmatrix} 0.0 & 0.7 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.7 & 0.0 & 0.5 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.5 & 0.0 & 0.5 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.5 & 0.0 & 0.5 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.5 & 0.0 & 0.7 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.7 & 0.0 \end{pmatrix} \qquad (9)$$

For a general alkane one can write,

$$R = d''\,T + Q \qquad (10)$$

where, $q_{12} = q_{21} = q_{\overline{N-1}\,N} = q_{N\,\overline{N-1}} = 0.2$ and the other matrix elements of Q are zero. The linear dependences of other foregoing quantities on RBI are manifestly evident.

Diamagnetic susceptibility matrix (χ) : The expres-

sion for diamagnetic susceptibility[8] of a molecule $A_lB_kC_j$ in the singlet ground state is

$$\chi_m = -\left(\frac{Ne^2}{6mc^2}\right)\left\{ j\sum_\mu \overline{r_\mu^2} + k\sum_\nu \overline{r_\nu^2} + l\sum_\eta \overline{r_\eta^2} \right\} \quad (11)$$

Evaluating $\overline{r_\mu^2}$, $\overline{r_\nu^2}$ etc. quantum mechanically, it is possible to write down the matrix χ. But such a matrix will be diagonal and not tridiagonal type as for a topological matrix. Instead, one can build up χ-matrix indirectly in tridiagonal form by spreading Pascal's constants[9] over atom-atom bond contributions so as to match the final value of diamagnetic susceptibility of the molecule in a homologous series of compounds. A specific example of χ-matrix for n-hexane will be

$$\chi_{(molar)} \text{ [Hexane]} =$$

$$-10^{-6} \begin{pmatrix} 12.6 & 1.56 & 0.0 & 0.0 & 0.0 & 0.0 \\ 1.56 & 8.43 & 1.56 & 0.0 & 0.0 & 0.0 \\ 0.0 & 1.56 & 8.43 & 1.56 & 0.0 & 0.0 \\ 0.0 & 0.0 & 1.56 & 8.43 & 1.56 & 0.0 \\ 0.0 & 0.0 & 0.0 & 1.56 & 8.43 & 1.56 \\ 0.0 & 0.0 & 0.0 & 0.0 & 1.56 & 12.6 \end{pmatrix} (12)$$

Hence χ (for a general member of the homologous series of alkanes)

$$= a' T + b' 1 + M \quad (13)$$

where, $a' \approx -1.56 \times 10^{-6}$ and $b' \approx -8.43 \times 10^{-6}$. M is a matrix where, $(M)_{11} = (M)_{NN} = -4.213 \times 10^{-6}$ and all other $(M)_{ij}$'s are zero.

Hence, all the observed linear relations are explicable following the reasonings employed for E_{ex} and $\sum S_{ab}$ previously. Similarly, the matrices of the thermodynamic functions can be built up in tridiagonal forms only indirectly like χ-matrix.

(B) Discovery of new inter-relationships :

Let us revert to the equations for the c.d.e.i. matrix $D^{(cond)} = a T + b 1 + M$, where the M matrix is diagonal with $(M)_{11} = (M)_{NN} = (b/2)$ and others, zero. If the eigenvector matrix for T be U, the unitary transformation leads to

$$U^\dagger [D^{(cond)} - M] U = U^\dagger D_c U =$$

$$a U^\dagger T U + U^\dagger b 1 U,$$

i.e., $U^\dagger D_c U = D'_c = a T' + b 1 \quad (14)$

whereT' is diagonal and contains the topological eigenvalues as the diagonal elements. Since 1 is a unit matrix, it follows that D'_c should necessary be diagonal containing its eigenvalues, though these eigenvalues may not connote anything of physical or chemical relevance.

If $\lambda_i^{D'c}$ $(= d'_{11})$, $\lambda_i^{T'}(= t'_{11})$ and λ_i^1 $(=\delta_{11})$ be the i-th. eigenvalues of D'_c matrix, T' matrix and 1 matrix respectively, then by McClelland's relation[7],

$$\left| \sum_i^N (d'_{11} - N b) \right| = \left| a \sum_i^N t'_{11} \right| \simeq 0.91 \, a\sqrt{2MN} \quad (15)$$

$$\text{or, } d'_{11} = a \, t'_{11} + b \, \delta_{11} \quad (16)$$

where N is the dimensionality of the matrices. Since b (exchange energy component) is intrinsically negative, it follows that

$$\left| \sum_i^N d'_{11} \right| = 0.91 \, a\sqrt{2MN} - N |b| \quad (17)$$

$$= 2 \, Z \, a - N |b|$$

While equation (16) demands a linear relation between the eigenvalues of D'_c and the corresponding topological eigenvalues, equation (17) predicts a linear plot for the modular sum of the eigenvalues of D'_c vs Z. These predictions are borne out by actual calculations and plots (Table 2, Fig. 1). A close analysis reveals that the eigenvalues of D'_c matrix are those of the tridiagonal c.d.e.i. matrix of a polymethylene $-(CH_2)_n-$ corresponding to the normal alkane C_nH_{2n+2} from which two terminal H-atoms have been clipped.

Thus it is found that not only are the E_{ex}'s of normal alkanes but the E'_{ex}'s also of the derivable polymethylenes topologically vary with N and Z in excellent linear relationships. Extension of such argument to the overlap matrix equation, viz. $S^{(cond)} = c T + d 1 + P$ (eq. 8) enables one to predict the corresponding interrelationships covering both the normal alkanes and the polymethylenes.

Treatments meted out to the normal alkanes under (A) and (B) can at once be extended to the homologous series of primary alcohols and of carboxylic acids wherein the oxygen atoms in –O–H and >C=O

Fig. 1.

sation : For a homologous series, quantum mechanical quantities and properties of molecules, which are solely or predominantly the sum of two bonded-atom contributions, can be expressed in suitable tridiagonal condensed matrices which are dimensionally and structurally similar to the tridiagonal topological matrices. Such quantities/properties are topology-dependent and exhibit excellent linear variations with N, Z and RBI.

References

1. M. RANDIC, *J. Am. Chem. Soc.*, 1975, **97**, 6609.

2. D. BONCHEV and N. TRINAJSTIC, *J. Chem. Phys*, 1977 **67** 4517; D. V. S. JAIN, S. SINGH and V. GOMBAR, *Indian J Chem.*, *Sect. A*, 1985, **24**, 545; S. SINGH, B. K MISHRA and R. K. MISHRA, *Indian J. Chem.*, *Sect. A*, 1988, **27**, 188; H HOSOYA, *Bull. Chem. Soc. Jpn.*, 1971, **44**, 2332.

3. S. C. RAKSHIT, M. BANERJEE and B HAZRA, *Indian J Chem.*, *Sect. A*, 1988, **27**, 183.

4. S. C. RAKSHIT and B. HAZRA, *J. Indian Chem. Soc*, 1990, **67**, 887.

5. S. C. RAKSHIT and B. MUKHERJEE, *Indian J. Chem.*, *Sect A* 1983, **22**, 273; S. C. RAKSHIT and B HAZRA, *Indian J Chem.*, *Sect. A*, 1988, **27**, 145.

6. R. S. MULLIKEN, C. A. RICKE, D ORLOFF and H ORLOH, *J. Chem. Phys.*, 1949, **17**, 1248.

7. B. J. MCCLELLAND, *J. Chem. Phys*, 1971, **54**, 640, 1 GUTMAN and O. E. POLANSKY, "Mathematical Concepts in Organic-Chemistry", Springer-Verlag, Berlin 1986, 1. TURKER, *Math. Chem.*, 1984, **16**, 83; I. GUTMAN / *Phys Chem. (Leipzig)*, 1985, **266**, 59; I. GUTMAN, A V TEODOROVIC and LJ NEDELJKOVIC, *Theor. Chim Acta*, 1984, **65**, 23.

8. J. A. POPLE, *J. Chem. Phys.*, 1962, **37**, 60.

9. P. W. SELWOOD, "Magnetochemistry", Interscience. New York, 1956.

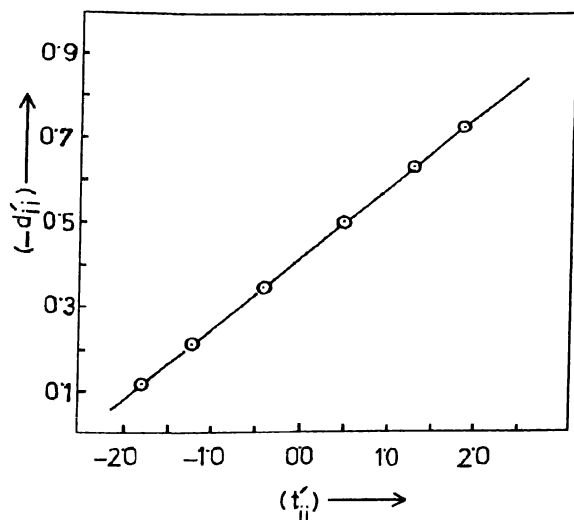10 R. L. DUTTA and A. S. SYAMAL, "Elements of Magneto-chemistry", S. Chand and Co., New Delhi, 1982.

groups can be deemed as pseudo-carbon atoms. The latter consideration is justified on the ground that the diatomic exchange integrals of the pairs [-OH, -CH] and [>C =O, -C-C-] are close enough as Table 3 demonstrates. This justifies the linear topological correlationships[3,4] with the quantum mechanical quan-

TABLE 3 – DIATOMIC EXCHANGE INTEGRALS OF SOME ATOM PAIRS IN HOMOLOGOUS SERIES (CNDO/2)

Homologous series	Diatomic exchange integral (a.u.)			
	O–H	C–II	O–C	C–C
Alkanes	–	–0.21	–	–0.17
Alcohols	–0.24	–0.21	–0.18	–0.17
Carboxylic acids	–0.24	–0.21	–0.19	–0.17

tities and molecular properties of the molecules of these homologous series. Expectedly and in reality such properties of silanes and primary amines are likewise related to topology.

In fine, we conclude with the following generali-