Bull. Chem. Soc. Jpn., *71*, 1013—1017 (1998)

A Divide-and-Conquer Algorithm Use to Study Some Benzenoid Hydrocarbons. A Graphtheoretic and Computational Approach

Swarna M. Patra and Rama K. Mishra*,†

Molecular Biophysics Unit, Indian Institute of Science, Bangalore-560012, India

†Chemical Physics Group, Department of Chemistry, Sambalpur University, Jyoti Vihar 768019, India

(Received September 12, 1997)

A computer program based on a divide-and-conquer algorithm was used to numerically split a univariate polynomial into factors depicting the various parts of E_π (π: electronic energy) and K (total Kekulé structure count) of the benzenoid graphs. A multivariate regression equation is suggested, taking n, m, and K_i as the independent variables and E_π (total) as the dependent variable; n, m, and K_i are the number of vertices, edges, and the fraction of the total K, respectively. A satisfactory correlation coefficient ($R^2 > 0.99$) was obtained.

More involved quantum-chemical calculations, such as the various semiempirical and ab initio approaches on the large benzenoid hydrocarbons, are limited by the computational restrictions. Recently, Randić et al.[1)] have pointed out that the largest systems for which the MINDO-type calculation[2)] can be made have some six or less fused benzene rings.[3)] However, the simplest HMO calculations can be routinely applied to benzenoid hydrocarbons having a dozen fused benzene rings.[4)] It has been pointed out that with the currently available software based on HMO theory and using a high-speed computer the benzenoids possessing 500 or more fused benzene rings can be smoothly handled.[1)] Chemical graph theory has made it possible to extend the quantitative examination of very large benzenoid hydrocarbons. It has been further mentioned that the graphtheoretic computational methods tend to approach the SCF MO methods with suitable parameterization schemes.[1)]

The natural starting point of the various properties determination of the large benzenoid has been the Kekulé structure count (K) of the benzenoid graphs. This is because the structure-count information is always in qualitative agreement with the experimental manifestations of the stability and reactivity. Excellent reviews and books are available in the chemical literature to study the enumeration of the structure-count of large conjugated cyclic polyene systems.[5—14)] Owing to the presence of a number of fused benzene rings it is quite logical to study the enumeration of K by decomposing the large benzenoid systems. Gutman et al.[15)] and Randić[16)] have independently established a relationship between the Kekulé structures of the parent conjugated molecule and of its constituent fragments. This procedure consists of three different steps. By applying the above procedure the initial molecule has been decomposed into smaller parts, and the process continues until one arrives at parts with the known number of Kekulé structures. Hence, following the above process one obtains

$$K\,(\text{parent conjugated molecule}) = \sum_i K_i\,(\text{constituting fragments}).$$

Here, we do not broach the different methods used for determining K and ASC, but a less subtle computational method using a F-77 program is proposed. A possible linear relationship between the π-electron energy of benzenoid hydrocarbons and Kekulé structures (or a part there of) is proposed based on the computer program. A computational operation (extraction technique) is applied to the CPs of the molecular graphs and the quadratic and linear factors responsible for the K are determined. Earlier, we designed a computer program[17—19)] based on the Birge–Vieta–Bairstow method, where linear and quadratic factors (LFs and QFs) are extracted selectively from one nth-order polynomial following one iteration technique.

Two long-known properties (taking eigenvalues) of the benzenoid graphs are well cited in the chemical literature.[5,20)]

$$K = \prod_i^{n/2} X_i \quad \text{and} \quad E_\pi = 2\sum_{i=1}^{n/2} X_i. \tag{1}$$

Here, X_i are the non-negative eigenvalues. In particular, $K=0$ if and only if the molecular graph possesses a zero eigenvalue.

An Outline of the Method

The large benzenoid graphs possessing a C_{2v} or D_{2h} symmetry can be split into halves by applying a fragmentation technique given by Mc Clelland.[21)] The CPs of the different fragments are considered. It is well known that multiplication of the CPs of all the fragments equals to the CP of the parent conjugated molecule.[21—24)] Let the CP of a molecular fragment (A) be written as

Table 1. Fragments, K_i, $(E_\pi)_i$, K, and E_π of Some PAH6 Approximation $P_1=-5.0$, $Q_1=6$, $p'_1=-3.0$, and $p'_2=-2.0$

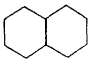
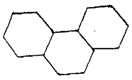
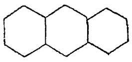
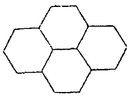
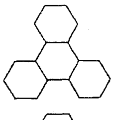
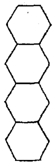
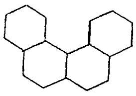
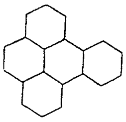
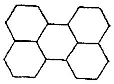
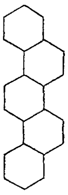
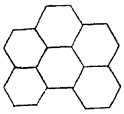
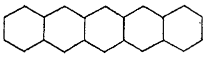
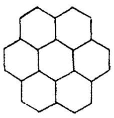
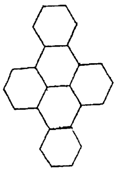
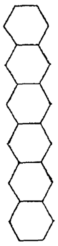
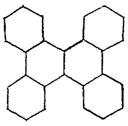
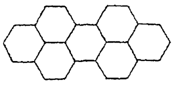
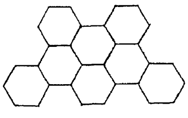
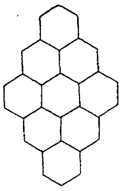
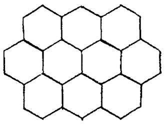
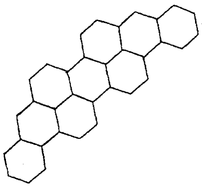
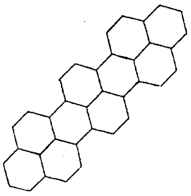
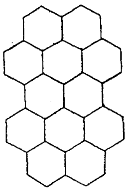
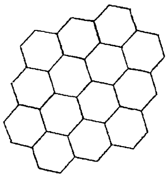
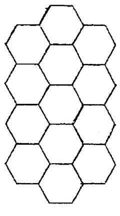
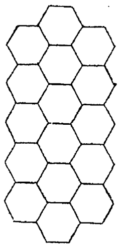
	Molecule	Fragment Right (A) & Left (B)	K_i	$(E_\pi)_i$ in $\lvert\beta\rvert$	K	E_π in $\lvert\beta\rvert$
1.		A	1	2.236	3	6.842
		B	3	4.606		
2.		A	2.9176	5.862	5	9.724
		B	1.7137	3.8620		
3.		A	2	4	4	9.656
		B	2	5.656		
4.		A	1	3.494	6	11.2528
		B	6	7.7588		
5.		A	5.9089	7.8186	9	12.63721
		B	1.53211	4.81861		
6.		A	1	4.977	5	12.4654
		B	5	7.4884		
7.		A	2	5.212146	8	12.593746
		B	4	7.3816		
8.		A	5.6270769	9.08395	11	14.167964
		B	1.954673	5.084014		
9.		A	1	4.75877	9	14.12267
		B	9	9.3639		
10.		A	4.6213072	8.735802	13	15.471607
		B	2.8130267	6.7358046		
11.		A	8.1325896	9.356272	14	15.712562
		B	1.72116714	6.35629		
12.		A	3	6.842	6	15.2723
		B	2	8.4303		
13.		A	2	6.4286	20	17.2859
		B	10	10.8573		
14.		A	3.6038645	5.8773	36	16.527191
		B	10	10.649891		
15.		A	1	7.7708	7	18.08302
		B	7	10.31222		
16.		A	4	8.191688	24	18.397738
		B	6	10.20605		
17.		A	1	5.742044	18	18.26014
		B	18	12.5181		
18.		A	13.982864	12.254253	45	21.520606
		B	3.1817157	9.2663526		
19.		A	1	7.549246	20	21.549246
		B	20	14		
20.		A	5	9.7241242	50	23.248724
		B	10	13.5246		
21.		A	6.1292972	13.134415	41	24.269076
		B	6.7052603	11.134661		
22.		A	1	8.566773	54	25.844279
		B	54	17.277506		

Table 1. (Continued)

	Molecule	Fragment Right (A) & Left (B)	K_i	$(E\pi)_i$ in $\|\beta\|$	K	$E\pi$ in $\|\beta\|$		Molecule	Fragment Right (A) & Left (B)	K_i	$(E\pi)_i$ in $\|\beta\|$	K	$E\pi$ in $\|\beta\|$
23.		A B	4 34	10.623356 16.132243	136	26.755599	25.		A B	5 35	12.151341 18.639171	175	30.790512
24.		A B	3 35	12.269338 16.918411	105	29.187749	26.		A B	14 14	15.71257 19.407155	196	35.119725

$$\mathrm{CP(A)} = X^n + A_1X^{n-1} + \cdots\cdots\cdots + A_{n-1}X + A_n = 0, \tag{2}$$

where A_1, A_2, $\cdots$, $\cdots$, A_n are the real coefficients, and may be associated with a positive or a negative sign. A quadratic factor of the type (X^2-PX+Q) or $(X-p_1)(X-p_2)$ is to be extracted from Eq. 2. If the polynomial Eq. 2 is divided by a trial factor, $(X^2+P_1X+Q_1)$, one obtains a quotient polynomial having degree $(n-2)$ and a remainder $(RX+S)$. On the other hand, if the polynomial Eq. 2 is divided by $(X+p'_1)$, a quotient polynomial of degree $(n-1)$ is obtained with a remainder R. Now, the problem is to solve the following equations with some iterative principle:

$$R(p'_1) = 0 \quad \text{or} \quad R(p'_2) = 0 \quad \text{(linear case)}, \tag{3}$$

$$R(P_1, Q_1) = 0, \quad S(P_1, Q_1) = 0 \quad \text{(quadratic case)}. \tag{4}$$

One can write

$$\begin{aligned}\mathrm{CP(A)} = {} & (X^2 + P_1X + Q_1)(B_nX^{n-2} + B_{n-1}X^{n-3} + \cdots\cdots \\ & + B_3X + B_2) + (RX + S),\end{aligned} \tag{5}$$

where the constants $B_n, B_{n-1}, \cdots, B_2, R$, and S have to be determined. Equating the coefficients of the like powers of X in Eqs. 2 and 5 we obtain:

$$\left.\begin{aligned} B_n &= A_n, \\ B_{n-1} &= A_{n-1} - P_1B_n, \\ B_{n-2} &= A_{n-2} - P_1B_{n-1} - QB_n, \\ &\rule{3em}{0.4pt} \\ R &= A_1 - P_1B_2 - Q_1B_3, \\ S &= A_0 - Q_1B_2. \end{aligned}\right] \tag{6}$$

From Eq. 6 it is clear that coefficients B of the factored polynomial (Deflated polynomial, DP) as well as the coefficients R and S are functions of P_1 and Q_1. Since, (X^2+PX+Q) is a factor of the given polynomial, it follows that

$$\left.\begin{aligned} & R(P,Q) = 0 \\ \text{and} \quad & S(P,Q) = 0 \end{aligned}\right] \tag{7}$$

$$\left.\begin{aligned} \text{Letting} \quad & P = P_1 + \Delta P_1 \\ \text{and} \quad & Q = Q_1 + \Delta Q_1 \end{aligned}\right] \tag{8}$$

Equation 8 can be expanded by Taylor's series as,

$$\left.\begin{aligned} & R(P,Q) = R(P_1,Q_1) + \Delta P_1\,\delta R/\delta P_1 + \Delta Q_1\,\delta R/\delta Q_1 = 0 \\ & \text{and} \\ & S(P,Q) = S(P_1,Q_1) + \Delta P_1\,\delta S/\delta P_1 + \Delta Q_1\,\delta S/\delta Q_1 = 0, \end{aligned}\right] \tag{9}$$

where the derivatives are computed at P_1 and Q_1. Equation 9 can be solved to obtain ΔP_1 and ΔQ_1. Using these values in Eq. 8 one can obtain a second approximation to P and Q. The process is repeated until successive values of P and Q show no significant change.[25] By this method a deflated polynomial (DP) and a quadratic factor, like (X^2-PX+Q) or a linear factor, like $(X-p_1)$, is obtained. To obtain the extracted factors correctly one must be very sure of the proper convergence of the factor and the coefficients of DP. Then, following the method a set of (P_1,Q_1) values is selected and given as initial approximations to run the program (A F-77 program has been given in one of our earlier papers[17]). The CPs of the two different fragments of a PAH6 are subjected to computational operation. Making use of the program, QFs $(X^2-P_iX+Q_i)$ and LFs $(X-p_i)$ are extracted from the CPs of the fragments. It may be mentioned that these QFs and LFs contain the Hückel positive eigenvalues, which are responsible for the K and $E\pi$ of the molecule under consideration.[5,20] Now, while extracting a QF, if another type of QF (which contains a negative eigenvalue), like (X^2+PX-Q) or (X^2-PX-Q) or (X^2-Q^2), is obtained, one has to go for one LF extraction in order to check the occurrence of the unwanted eigenvalue. Here, for the LF extraction the needed initial approximation (p'_1 or p'_2) is generated inside the program as $(p'_1+p'_2)=P_1$ and $p'_1\cdot p'_2=Q_1$. Through this extraction one must get the $(X-p_i)$-type factor; otherwise, the (P_1, Q_1) values so chosen are ignored and a fresh set of (P_1, Q_1) values are selected. It was pointed out earlier that a set of (P_1, Q_1) values (initial approximations) extracts a QF and (p'_1 or p'_2) extracts a LF through the iteration technique. The use of the initial approximations has been described in our previous papers.[17—19]

Example: Let us consider a representative (PAH6) molecule having 20 carbon atoms and applying the fragmentation technique, we have obtained 2 fragments; A and B fragments contain 10 carbon atoms each (Chart 1). Hence, two 10th-order polynomials were obtained. Now, considering a set of $(P_1=-5.0,\ Q_1=6.0)$ values, the above two polynomials were subjected to our computational technique.

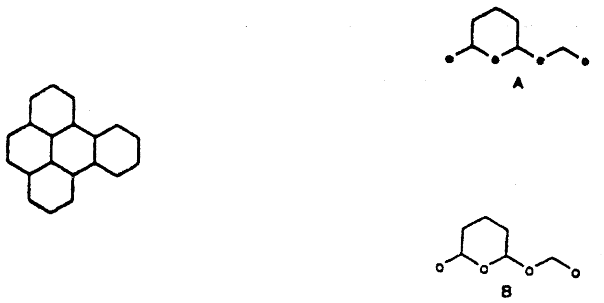


Chart 1.

Fragment A:

$$\begin{aligned}\mathrm{CP(A)} &= X^{10}+4X^9-4X^8-30X^7-8X^6+70X^5+41X^4\\ &\quad -62X^3-41X^2+18X+11\end{aligned}$$

$$\Big\downarrow\ \begin{array}{l} P_1=-5.0,\quad Q_1=6.0\\ (X^2-3.365933X+2.72208)\end{array}$$

$$\begin{aligned}\mathrm{DP}_1 &= X^8+7.365933X^7+18.07116X^6+10.77564X^5\\ &\quad -20.92105X^4-29.75101X^3-2.19112X^2\\ &\quad +11.60946X+4.041059\end{aligned}$$

$$\Big\downarrow\ \begin{array}{l} (P_1,Q_1)\\ (X^2-1.718081X+0.7180804)\end{array}$$

$$\begin{aligned}\mathrm{DP}_2 &= X^6+9.084013X^5+32.96014X^4+60.88076X^3\\ &\quad +60.00896X^2+29.6319X+5.627581\end{aligned}$$

(abandoned)

Note: No more $(X^2-P_iX+Q_i)$ or $(X-p_i)$ type of terms are extracted from DP_2; also, DP_2 does not contain a negative term. ($K_A=$ 1.954673 and $(E_\pi)A=5.084014\beta$)

Fragment B:

$$\begin{aligned}\mathrm{CP(B)} &= X^{10}-4X^9-4X^8+30X^7-8X^6-70X^5+41X^4\\ &\quad +62X^3-41X^2-18X+11\end{aligned}$$

$$\Big\downarrow\ \begin{array}{l} (P_1,Q_1)\\ (X^2-4.697174X+5.455858)\end{array}$$

$$\begin{aligned}\mathrm{DP}_1 &= X^8+0.6971741X^7-6.181111X^6-2.83744X^5\\ &\quad -12.39533X^4-3.703673X^3-9.230355X^2\\ &\quad -1.563303X+2.016412\end{aligned}$$

$$\Big\downarrow\ \begin{array}{l} (P_1,Q_1)\\ (X^2-2.889975X+2.075696)\end{array}$$

$$\begin{aligned}\mathrm{DP}_2 &= X^6+3.587149X^5+2.109967X^4-4.18552X^3\\ &\quad -4.080367X^2-0.5993796X+0.9714366\end{aligned}$$

$$\Big\downarrow\ \begin{array}{l} (P_1,Q_1)\\ (X^2-1.49688X+0.49696)\end{array}$$

$$\begin{aligned}\mathrm{DP}_3 &= X^4+5.083998X^3+9.22303X^2+7.093639X\\ &\quad +1.95460\end{aligned}$$

(abandoned)

$$K_B=5.6270769 \quad \text{and} \quad (E_\pi)B=9.08395\beta$$

By applying our method we are in a position to obtain two QFs of the type (X^2-PX+Q) from the CP of A and three QFs of the same type are obtained from B using $P_1=-5.0$ and $Q_1=6.0$. After extracting all of the factors the total K ($K=K_A\cdot K_B$) is calculated and for this molecule ($K=11$). Hence, we say that these (P_1, Q_1) values can give the factors responsible for the Kekulé structure count of the molecule under investigation.

Results

Here, we have selected a few cata-condensed and peri-condensed PAH6 having mirror-plane symmetry. By applying the Mc Clelland fragmentation technique, the large PAH6 molecules are divided into two fragments. It is well known that Kekuléan molecules have $K>0$.[5] For the two different parts, the K_is are calculated and found to be positive. These K_is are the parts of the original K of the PAH6 molecule under consideration. Our target is to find factors which are responsible for the Kekulé structure count of the large benzenoid system.

In order to verify that the computer operation ($P_1=-5.0$, $Q_1=6.0$) can give the factors responsible for the determination of the Kekulé structures for all of the molecules, we carried out the operation for both fragments of the different molecules under consideration. Table 1 shows the different components (A and B) of the molecules, K_i and $(E_\pi)_i$ values of the different components along with the total K and E_π of the benzenoid molecules.

Since the (P_1, Q_1) values give the desired factors, we are apt to say that the $P_1=-5.0$ and $Q_1=6.0$ operation[17—19] gives the Kekulé structure count for all of the 26 studied PAH6 molecules. It is true that a number of analytical methods are available for determination of the Kekulé structures. However, if the computational method is applied, one has to attempt the construction of the CP of a large benzenoid molecule and to compute all of the eigenvalues of the molecular graph. It is well known that CP construction requires high precision arithmetically,[26,27] which is not a demerit of the computational method, but requires a mainframe-type computer having the facilities of a quadruple precision. However, in this method smaller components are being used to determine the Kekulé structure, and easily factors like (X^2-PX+Q) and $(X-p_i)$ are obtained, which eventually give the total K and E_π of the benzenoid molecules.

Relationship between E_π (Total) and K_i (Fraction of the Total K) of the Benzenoid Molecules Studied through Regression Analysis. Using the Hückel theory, the total π-electronic energy of a benzenoid graph can be defined as twice the sum of the positive eigenvalues of the adjacency matrix.[28] Many workers[29—32] have investigated the relationship between E_π and K. Although Cioslowski[33] has arrived at a non-linear dependence of E_π on K, Hall[28] has shown that E_π largely depends on n, m, and K, where n is the number of vertices, m is the number of edges, and K is the number of the Kekulé structure of the benzenoid hydrocarbons. After that, Hall and Gutman[29] further advocated the linear dependence of E_π on K based on the condition that the highest occupied molecular orbitals (HOMOs) should be non-degenerate and well separated from the second occupied MOs for the benzenoid hydrocarbons.

As pointed out earlier, the large PAH6s contain a number of six-membered rings, and it is a stupendous task to determine the various properties of the molecules using different semiemperical calculations.[1] Hence, methods should be devised to break the PAH6 into various parts; also, the properties of the whole molecule have to be studied based on the different components. Here, in this work, after breaking the benzenoid graphs the K_i were calculated using the computer program. Now, making use of K_i (a part of the total K) the E_π (total) can be calculated, and thereby the linear dependence of K_i on E_π (total) can be established through the following regression equations with 26 data points, (16 fused benzene rings are present in one of the PAH6 under study):

$$E_\pi = an + bm + cK_i + d. \tag{10}$$

Here, a=0.7409 (±0.58747), b=0.38498 (±0.45338), c=0.0046197 (±0.045744), d=1.7691 (±1.5134), R^2=0.95, and F=133.4. It is well known that after fragmentation the different components contain two types of vertices and edges i.e. weighted and non-weighted. By separating the vertices to n_1 (weighted) and n_2 (non-weighted), and then the edges to m_1 (weighted) and m_2 (non-weighted), we obtained the best regression equation, as follows:

$$E_\pi = an_1 + bn_2 + cm_1 + dm_2 + eK_i + f, \tag{11}$$

where, a=2.2005 (±0.24817), b=2.4801 (±0.24904), c=−1.8933 (±0.24631), d=−0.77283 (±0.18422), e= 0.052714 (±0.016648), R^2=0.9939, and F=689.1.

This regression Eq. 11 clearly depicts a linear relationship between E_π and K_i. Further, we can say that if a large benzenoid molecule possesses C_{2v} or D_{2h} symmetry, a part of it (after breaking) is sufficient to give the total E_π of the molecule.

SMP thanks DBT, Govt. of India, for a Research Associate Position at Molecular Biophysics Unit, I. I. Sc, Bangalore-560012.

References

1) M. Randić, D. J. Klein, S. El Basil, and P. Calkins, *Croat. Chem. Acta*, **64**, 1639 (1996).

2) M. J. S. Dewar, "Molecular Orbital Theory of Organic Chemistry," McGraw-Hill, New York (1969).

3) M. J. S. Dewar and C. de Llano, *J. Am. Chem. Soc.*, **91**, 789 (1969).

4) C. A. Coulson and A. Streitwieser, Jr., "Dictionary of π-Electron Calculations," W. H. Freeman, San Francisco (1965).

5) N. Trinajstić, "Chemical Graph Theory," CRC Press, BR, Florida (1983), Vol. II.

6) S. J. Cyvin and I. Gutman, "Kekulé Structure in Benzenoid Hydrocarbons," Springer, Berlin (1988).

7) J. R. Dias, "Handbook of Polycyclic Hydrocarbons," Part A and B, Elsevier, New York (1987) and (1988).

8) G. G. Hall, *Chem. Phys. Lett.*, **145**, 168 (1988).

9) W. C. Herndon, *Tetrahedron*, **29**, 3 (1973); *J. Chem. Educ.*, **51**, 10 (1974).

10) C. F. Wilcox, *Tetrahedron Lett.*, **1968**, 795 ; *J. Am. Chem. Soc.*, **91**, 2732 (1969).

11) M. Randić, *Mol. Phys.*, **34**, 849 (1977).

12) I. Gutman, *Z. Naturforsch., A*, **39A**, 794 (1984).

13) J. R. Dias, *Z. Naturforsch., A*, **44A**, 761 (1989).

14) J. R. Dias, *J. Math. Chem.*, **9**, 253 (1992).

15) I. Gutman and N. Trinajstić, *Croat. Chem. Acta*, **45**, 423 (1973).

16) M. Randić, *J. Chem. Soc., Faraday Trans. 2*, **72**, 232 (1976).

17) S. M. Patra and R. K. Mishra, *Int. J. Quantum Chem.*, **53**, 361 (1995).

18) S. M. Patra and R. K. Mishra, *J. Mol. Struct. (THEOCHEM)*, **342**, 201 (1995).

19) S. M. Patra and R. K. Mishra, *Bull. Chem. Soc. Jpn.*, **70**, 997 (1997).

20) I. Gutman and O. E. Polansky, "Mathematical Concepts in Organic Chemistry," Springer Verlag, Berlin (1986).

21) B. J. McClelland, *J. Chem. Soc., Faraday Trans. 2*, **70**, 1453 (1974); *Mol. Phys.*, **45**, 189 (1982).

22) J. R. Dias, "Molecular Orbital Calculations Using Chemical Graph Theory," Springer Verlag, New York (1993).

23) J. R. Dias, *J. Chem. Educ.*, **66**, 1012 (1989).

24) J. R. Dias, *Mol. Phys.*, **88**, 407 (1996).

25) S. S. Sastry, "Introductory Methods of Numerical Analysis," Prentice- Hall of India, Pvt. Ltd., New Delhi (1995), p. 35.

26) K. Balasubramanian, *J. Comput. Chem.*, **5**, 387 (1984); **6**, 656 (1985).

27) T. P. Živkovic, *J. Comput. Chem.*, **11**, 217 (1990).

28) G. G. Hall, *Int. J. Quantum Chem.*, **39**, 605 (1991).

29) I. Gutman and G. G. Hall, *Int. J. Quantum Chem.*, **41**, 667 (1992), and references there in.

30) J. Cioslowski, *MATCH*, **20**, 95 (1986); *Int. J. Quantum Chem.*, **31**, 581 (1987).

31) I. Gutman and Marković, *MATCH*, **25**, 141 (1990).

32) I. Gutman, *MATCH*, **21**, 317 (1986).

33) J. Cioslowski, *MATCH (Mülheim)*, **20**, 95 (1986).