

Topological Localized Molecular Orbitals

I. A Localization Method in the HMO Framework*

Juan Carlos Paniagua

Departament de Química Física, Facultat de Química, Universitat de Barcelona, Barcelona-28, Spain

Albert Moyano

Departament de Química Orgànica, Facultat de Química, Universitat de Barcelona, Barcelona-28, Spain

Luis María Tel

Departamento de Química Física, Facultad de Química, Universidad de Salamanca, Salamanca, Spain

Intrinsic and external π -orbital localization procedures which rely only on molecular topology are proposed and discussed. Localized molecular orbitals obtained by application of these procedures are referred to as “topological localized molecular orbitals”.

Key words: Topology – Localized molecular orbitals – Hückel method.

1. Introduction

Localization of molecular orbitals (MO's) in polycyclic conjugated hydrocarbons offers special interest due to the peculiar features of the resulting orbitals, particularly in connection with the concept of local aromaticity [1], and has deserved considerable attention [2]. Ruedenberg and coworkers [2a, b] first analyzed energy [3] localized MO's (LMO's) obtained from Hückel–Wheland canonical MO's (CMO's). The bielectronic integrals necessary for the localization, which are not available from the Hückel–Wheland calculation, were

* Presented in part at the European Symposium of Organic Chemistry II, Stresa, Italy, June 1981 and at the XII “Congrés dels químics teòrics d'expressió llatina”, Barcelona, Spain, September 1981.

approximated following an earlier suggestion made by Mulliken [4] which implies the evaluation of a Coulomb repulsion integral for every interatomic distance. Lipscomb et al. [2c] employed Boys [5] criterium to localize CMO's calculated in the partial retention of diatomic differential overlap approximation [6] for some conjugated polycyclic systems. The localization was carried out both with and without the constraint of σ - π separation, obtaining π -LMO's in good agreement with the calculations of Ruedenberg, in the first case, and pairs of banana type orbitals alternating with single bonds in the second. However, structures of this type are not always obtained when all the occupied MO's of a conjugated molecule are localized simultaneously: structures with σ - π separation result for some monocyclic aromatic ring systems [7] and even mixed structures have been found [2a]. On the other hand, the type of structure obtained often depends on the localization criterium being used [8]. Notwithstanding, σ - π structures always correspond to stationary points of the localization sum hypersurface [8]. Moreover, the information of interest in aromaticity studies can be drawn from π -LMO's, so that the utilization of simple π -type calculations for comparative localizations in related conjugated molecules is justified. Since in these systems topological methods for the calculation of π -CMO's have proved so successful [9], it seems reasonable to expect that LMO's can be obtained also by exclusively topological criteria. This would avoid having to resort to additional information concerning molecular geometry when topological CMO's are localized [2b], and would constitute the most coherent way to localize MO's obtained from Hückel calculations¹.

In the present work we develop a localization method for π -MO's which relies only on molecular topology, and we discuss several variants of it which can be related to previously proposed criteria. The method has been applied to a wide variety of conjugated compounds, giving results in good accordance with those obtained by using more sophisticated methods [2a, b, c]. Some of these results will be presented in the second publication of this series [11].

2. The Localization Sum in the Hückel Approximation

Let $\{\chi_r\}_{r=1,\dots,m}$ be the set of $2p_z$ atomic orbitals (AO's) centered on the carbon atoms of a conjugated hydrocarbon, and let ϕ_i be a MO built as a linear

¹ In reference 2d a modification of the Magnasco-Perico external localization method [10] is proposed in order to localize π -MO's without the inclusion of any data other than those obtained from a Hückel calculation. However the method is incorrectly settled since it is stated that localization is attained by minimization of the quantity

$$\sum_i \sum_{m,n} P_{mn} H_{mn}$$

(where index i refers to the prefixed bonds of the structure, and m, n run over the AO's contributing to bond i) which is clearly invariant under unitary transformations of the MO's. Nevertheless we feel that what the authors probably did was to apply the Magnasco-Perico method substituting the overlap matrix by the Hückel matrix \mathbf{H} in the definition of local orbital population (Eq. (8) of Ref. [10]).

combination of these AO's (LCAO):

$$\phi_i = \sum_r^m \chi_r C_{ri} = \boldsymbol{\chi} \mathbf{C}_i. \quad (1)$$

Hartree–Fock equations for the n CMO's corresponding to a $2n$ -electron closed-shell configuration reduce, in the Hückel formalism, to

$$\mathbf{TC} = \mathbf{C}\boldsymbol{\varepsilon} \quad (2)$$

where

\mathbf{T} is the topological matrix:

$$T_{rs} = 1 \text{ if atoms } r \text{ and } s \text{ are } \sigma\text{-bonded} \\ = 0 \text{ otherwise,}$$

\mathbf{C} is an $m \times n$ matrix whose columns are the \mathbf{C}_i 's, and

$\boldsymbol{\varepsilon}$ is a diagonal $n \times n$ matrix whose diagonal elements are the CMO's energies expressed in β units and taking $\alpha = 0$ as the origin.

An orthogonal transformation² of the CMO's leaves Eq. (2) invariant, i.e.

$$\mathbf{C}' = \mathbf{C}\mathbf{O},$$

$$\mathbf{O}'\mathbf{O} = \mathbf{O}\mathbf{O}' = \mathbf{1}$$

implies

$$\mathbf{TC}' = \mathbf{C}'\boldsymbol{\varepsilon}'$$

with

$$\boldsymbol{\varepsilon}' = \mathbf{O}'\boldsymbol{\varepsilon}\mathbf{O}$$

where $\boldsymbol{\varepsilon}'$ is, in general no longer diagonal, but still a symmetric matrix. This flexibility allows to select a set of MO's which exhibit a maximum degree of localization according to any predetermined criterium.

Intrinsic localization criteria can be expressed in a general form as the maximization of a localization sum

$$S = \sum_i^n [\phi_i \phi_i | \phi_i \phi_i] \quad (3)$$

where

$$[\phi_i \phi_j | \phi_k \phi_l] \equiv \int \phi_i(\mathbf{r}_1) \phi_j(\mathbf{r}_1) f(\mathbf{r}_{12}) \phi_k(\mathbf{r}_2) \phi_l(\mathbf{r}_2) d\tau_1 d\tau_2$$

$f(\mathbf{r}_{12})$ being a different function of \mathbf{r}_{12} for each criterium:

$$f(\mathbf{r}_{12}) = r_{12}^{-1} \text{ for the Edmiston–Ruedenberg criterium [3]}$$

$$= -r_{12}^2 \text{ for the Foster–Boys criterium [5]}$$

$$= \delta(\mathbf{r}_{12}) \text{ for the von Niessen criterium [12].}$$

² We can always choose the coefficients C_{ri} to be real.

Introducing the LCAO expansion (1) in the localization sum (3) we obtain

$$S = \sum_i^n \sum_{rstu}^m C_{ri} C_{si} [\chi_r \chi_s | \chi_t \chi_u] C_{ti} C_{ui}$$

and, making use of the zero differential overlap (ZDO) approximation

$$\chi_r(\mathbf{r}_1) \chi_s(\mathbf{r}_1) d\tau_1 = \delta_{rs} \chi_r^2(\mathbf{r}_1) d\tau_1$$

we can write

$$S = \sum_i^n \sum_{rt}^m C_{ri}^2 L_{rt} C_{it}^2 \quad (4)$$

where a matrix L has been introduced whose elements are defined as

$$L_{rt} \equiv [\chi_r \chi_r | \chi_t \chi_t] \quad (5)$$

and which will be referred to as the "localization matrix". Eq. (4) can be expressed in a more compact way by introducing a new matrix Q with elements

$$Q_{ri} \equiv (C_i C_i^t)_{rr} = C_{ri}^2$$

giving

$$S = \text{Tr} (Q^t L Q). \quad (6)$$

The calculation of the elements of matrix L requires the knowledge of molecular geometry in the case of Edmiston–Ruedenberg and Foster–Boys criteria and, in any case, the explicit functional form of the basis orbitals. Alternatively, those elements could be parametrized. Nevertheless, neither these parameters nor molecular geometry are defined within the HMO formalism; moreover, the functional form of AO's is not necessary for carrying out an HMO calculation. It is then clear that it would be desirable to dispose of topological expressions for the localization matrix.

3. Topological Localization Matrices

In the case of Edmiston–Ruedenberg criterium, the matrix element L_{rt} is the Coulomb repulsion integral between $2p_z$ AO's centered on atoms r and t . The behaviour of this integral as a function of the internuclear distance R_{rt} has been studied for different kinds of AO's, and several approximated analytical expressions have been proposed for that function [13]. In every case the following limiting behaviour is found:

$$\lim_{R_{rt} \rightarrow \infty} L_{rt} \propto 1/R_{rt}$$

as could be expected from a simple qualitative reasoning. Among the various kinds of topological matrices³ that have been introduced [14] the one most suited

³ The term "topological" refers here to any matrix used to describe molecular topology. We have also used it for matrix T , defined in the previous section, which is the most widely used topological matrix. Nevertheless its meaning will be clear from the context in every case.

to mimic this behaviour is the distance matrix \mathbf{D} [15], whose elements D_{rt} are equal to the minimum number of σ -bonds connecting atoms r and t . Expressed in terms of this matrix the topological localization matrix (TLM) corresponding to the Edmiston–Ruedenberg criterium could be:

$$\begin{aligned} L_{rt} &= 1 && \text{for } r = t \\ &= k/D_{rt} && \text{for } r \neq t \end{aligned} \quad (7)$$

where a constant proportionality factor expressing the units of \mathbf{L} is indifferent for our purposes, and k is an adjustable parameter which will be discussed later on.

Even though being purely topological⁴, definition (7) can be further simplified: since HMO calculations show that neglecting the matrix elements of the effective π -mono-electron hamiltonian between AO's on atoms not σ -bonded gives unexpectedly good results, it seems reasonable to try doing a similar approximation for the above localization matrix, i.e. to take

$$L_{rt} = \delta_{rt} + kT_{rt} \quad (8)$$

where δ_{rt} is the Kronecker delta.

In the Foster–Boys criterium, integrals (5) with $r \neq t$ are, barring their sign, a measure of the distance between AO's χ_r and χ_t (at least when atoms r and t are far enough apart), and the following limiting behaviour can easily be verified

$$\lim_{R_{rt} \rightarrow \infty} L_{rt} \propto -R_{rt}^2$$

while diagonal elements are related to the degree of concentration of the corresponding AO. Their topological version could then be

$$L_{rt} = -(\delta_{rt} + kD_{rt}^2)$$

where again parameter k should be given an adequate value.

In this case it is not clear that a simpler expression can be given in terms of matrix \mathbf{T} . On the other hand, maximization of the localization sum as expressed in Eq. (3) is not the most efficient way of effecting Boys localization [5c], so that we will not further consider this TLM. Notwithstanding, we will show later on that taking expression (8) for \mathbf{L} amounts to making a simple topological translation of Boys criterium.

Von Niessen integrals (5) reduce in the ZDO approximation to⁵

$$L_{rt} = \delta_{rt}[\chi_r\chi_r|\chi_r\chi_r]$$

and so, not even molecular topology is needed to get a simple parametrization

⁴ Actually, it can be shown [15] that matrix \mathbf{D} is a simple function of matrix \mathbf{T} .

⁵ For a recent localization method based on the application of the ZDO approximation to the von Niessen criterium, see Ref. 16.

for them:

$$L_{rt} = \delta_{rt}. \quad (9)$$

This can be viewed as a particular case of definition (8) with $k = 0$.

Let us analyze more deeply the TLM defined by Eq. (8), for which we will reserve henceforth the initials TLM. In order to verify that its maximization leads to LMO's, we shall consider the contribution made by a single MO to the localization sum, i.e.

$$S_i = \sum_r^m Q_{ri} L_{ri} Q_{ri} \quad (10)$$

for some special choices of coefficients $\{Q_{ri}\}_{r=1,\dots,m}$ satisfying the normalization condition

$$\sum_r^m Q_{ri} = 1. \quad (11)$$

Introducing Eq. (8), Eq. (10) becomes

$$S_i = \sum_r^m Q_{ri}^2 + 2k \sum_{r-t} Q_{ri} Q_{ti} \quad (12)$$

where “ $r-t$ ” means that the sum runs over every pair of σ -bonded atoms r, t .

Let us first consider the case $k = 1$. We define a p -centered MO as one having p and only p nonvanishing coefficients. A monocentric nonbonding MO centered on atom r would have $Q_{ri} = 1$ and $Q_{si} = 0$ for $s \neq r$ and its contribution to S will thus be $S_i = 1$. A bicentric MO ($Q_{ri} \neq 0$, $Q_{si} \neq 0$ and $Q_{ti} = 0$ for $t \neq r$ and s) centered on σ -bonded atoms r, s gives again a contribution of 1:

$$\begin{aligned} S_i &= Q_{ri}^2 + Q_{si}^2 + 2Q_{ri}Q_{si} \\ &= (Q_{ri} + Q_{si})^2 = 1, \end{aligned}$$

where use of Eq. (11) has been made. However, if atoms r and s were not σ -bonded we would get

$$\begin{aligned} S_i &= Q_{ri}^2 + Q_{si}^2 \\ &\leq (Q_{ri} + Q_{si})^2 = 1, \end{aligned}$$

where the equality only holds for $Q_{ri} = 0$ or $Q_{si} = 0$. For a tricentric MO (Q_{ri} , Q_{si} and $Q_{ti} \neq 0$, and $Q_{ui} = 0$ for $u \neq r, s$ and t), atoms s and t being σ -bonded to r ,

$$\begin{aligned} S_i &= Q_{ri}^2 + Q_{si}^2 + Q_{ti}^2 + 2Q_{ri}Q_{si} + 2Q_{ri}Q_{ti} \\ &\leq (Q_{ri} + Q_{si} + Q_{ti})^2 = 1, \end{aligned}$$

where the equality holds for $Q_{si} = 0$ and/or $Q_{ti} = 0$. Again the contribution would be smaller if atoms s and/or t were not σ -bonded to r . In general, it is clear that the contribution S_i will be < 1 for every p -centered MO with $p > 2$, and that

it will decrease as the MO spreads over the molecule, since the number of positive terms in the square of Eq. (11) not present in Eq. (12) will be larger. Hence, for $k = 1$ the localization procedure will tend to concentrate MO's on atoms σ -bonded between themselves and preferably over one or two centres.

Giving parameter k a value less than 1, similar results are obtained, but now a greater weight is given to a monocentric MO than to a bicentric one: in the first case the contribution takes the maximum value ($S_i = 1$) while in the second it depends on the asymmetry of the orbital, varying monotonically between $(1 + k)/2$ for a symmetric one and 1 for the monocentric limiting case. In the same way, tricentric MO's are given a comparatively minor weight, and the same happens to more widely-spread orbitals.

Inversely, a value greater than 1 for k would favour symmetric bicentric MO's as well as certain slightly peaked tricentric ones, as it can be seen in the limiting case $L = T$: the contribution of a bicentric MO goes from $1/2$ in the symmetric case to 0 in the most asymmetric one, that is, the monocentric MO; for tricentric orbitals, the maximal contribution ($S_i = 1/2$) corresponds to the case $Q_{ri} = 1/2$ and $Q_{si} = Q_{ti} = 1/4$, atom r being the central one.

Taking expression (9) for matrix L amounts to making $k = 0$ in expression (8), as we have already pointed, and therefore will tend to peak MO's over one of the contributing atoms. In this case it can be shown that a p -centered MO whose coefficients are all equal would make a contribution of $1/p$ to the localization sum⁶.

It is worth noting that, when we take for L the general expression (8), maximization of the localization sum amounts to minimizing the expression

$$n - S = \sum_i^n \left[2 \sum_{r \neq t} Q_{ri} Q_{ti} + 2(1 - k) \sum_{r=t} Q_{ri} Q_{ti} \right]$$

where " $r \neq t$ " means that the sum runs over every pair of not σ -bonded atoms, and use of Eq. (11) has been made. This can be viewed as a topological version of Boys criterium, since it corresponds to taking for the integrals

$$\int (\chi_r(\mathbf{r}_1))^2 r_{12}^2 (\chi_t(\mathbf{r}_2))^2 d\tau_1 d\tau_2$$

the following parametrization:

$$\begin{array}{ll} 0 & \text{if } r = t, \\ 1 - k & \text{if } r \text{ and } t \text{ are } \sigma\text{-bonded, and} \\ 1 & \text{otherwise.} \end{array}$$

⁶ The contribution of a MO of this kind to the localization sum is, in general,

$$S_i = 1/p \left\{ \sum_{\alpha}^p L_{i\alpha i\alpha} + 2 \sum_{\alpha > \beta}^p L_{i\alpha i\beta} \right\}$$

where indices α and β run over AO's upon which MO ϕ_i is centered.

The former considerations suggest that an external localization criterium can be introduced within the same formalism we are discussing. Let us build a localization matrix associated to a valence bond (VB) structure of the molecule under study in the following way:

$$\begin{aligned}
 L_{rt} &= 1 && \text{if atoms } r \text{ and } t \text{ are joined by a } \pi\text{-bond in the VB structure,} \\
 &= 1/2 && \text{if } r = t \text{ and corresponds to a } \pi\text{-lone pair, and} \\
 &= 0 && \text{otherwise.}
 \end{aligned} \tag{13}$$

With this choice for matrix L the localization sum (6) becomes

$$S = \sum_i^n \left\{ \frac{1}{2} \sum_r^{\text{l.p.}} Q_{ri}^2 + 2 \sum_{r,t}^{\pi\text{-b.}} Q_{ri} Q_{ti} \right\}$$

(where the meaning of the labels of the summation signs is clear), so that its maximization will tend to localize MO's where the associated VB structure has a π -bond or a π -lone pair. The reason for giving the value 1/2 to the non vanishing diagonal elements of L is that both types of LMO's are thus given the same weight in the localization procedure.

To end this discussion we will consider a simple example that illustrates how the choice of the localization matrix can condition the localization process. Let us localize the two degenerate highest occupied MO's (HOMO's) of cyclobutadiene. These two orbitals do not constitute any physical closed-shell system; however, they can be subjected to the mathematical process of localization, since this is applicable to any set of orthonormal MO's. Taking for matrix L the general expression (8) it is found that:

- (i) For $0 \leq k < 1/2$ localization leads to structure I (Fig. 1).
- (ii) For $k > 1/2$ localization leads to structure II (Fig. 1).
- (iii) For $k = 1/2$ localization is indeterminate: the localization sum is invariant under orthogonal transformations of the HOMO's.

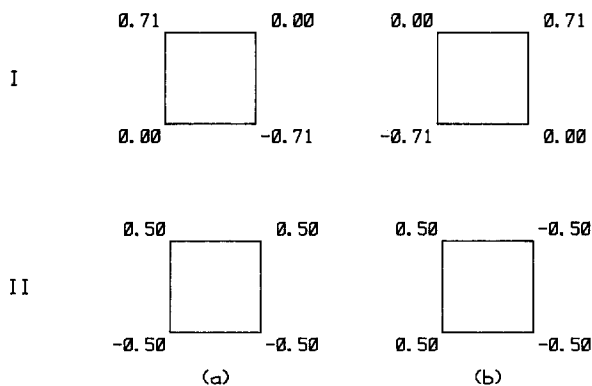


Fig. 1



Fig. 2

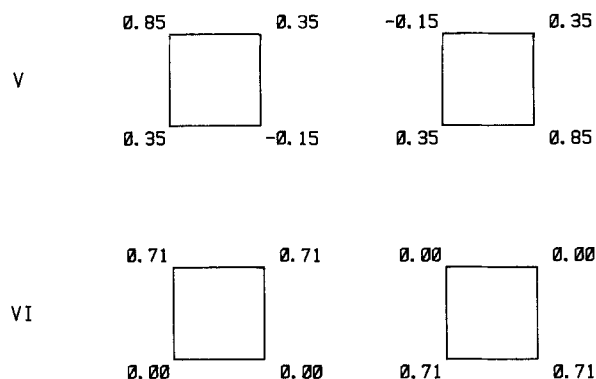


Fig. 3

On the other hand, taking for matrix L the external form (13) corresponding to the Kekulé structures III or IV (Fig. 2) we obtain the LMO's II(a) and II(b) respectively.

It is interesting to note that localizing one of the localized HOMO's with the remaining occupied CMO of cyclobutadiene, one obtains (see Fig. 3) a pair of equivalent trivalent (V) or bivalent (VI) LMO's, depending on whether k is $<1/2$ or $>1/2$ (the negative coefficients that appear in the first case are needed in order to keep LMO's orthogonal). These structures could be considered as the LMO's of the closed-shell systems obtained when the HOMO's degeneracy is broken by a slight distortion of the D_{4h} cyclobutadiene geometry to either rhomboidal (V) or rectangular (VI) D_{2h} . For the external localization matrix, Kekulé type LMO's as those depicted in VI are obtained, as could be expected.

From this example one might conclude that the TLM's we have proposed give rather different results for a given problem. However cyclobutadiene is a most peculiar case inasmuch as those differences appear abnormally exaggerated: in fact, it will be seen in the following paper [11] that, in most cases, different localization matrices lead to very similar LMO's.

4. Conclusions

The intrinsic and external localization procedures proposed in this paper constitute the topological version of well-established localization criteria, and can

be expected to possess the same limitations—but also the same virtues—as do Hückel calculations of CMO's.

As illustrated in the following paper [11], topological localization of π -systems is able to give results in good accordance with the ones obtained with previously-used versions of localization criteria.

Topological localization not only could be comparable, but preferable to more sophisticated procedures, when a refinement of the calculation beyond the topological level is an unnecessary complication that obscures the interpretation of the results (Möbius systems [17] and non pericyclic topologies [18]). Moreover, topological localization would be of obligatory use for the acyclic reference structures [19] used in the definition of topological resonance energy [20].

Topological localized molecular orbitals of these and related systems are currently under investigation.

References

1. England, W., Ruedenberg, K.: *J. Am. Chem. Soc.* **95**, 8769 (1973)
2. (a) England, W., Salmon, L. S., Ruedenberg, K.: *Fortschr. Chem. Forsch.* **23**, 31 (1971); (b) England, W., Ruedenberg, K.: *Theoret. Chim. Acta (Berl.)* **22**, 196 (1971); (c) Dixon, D. A., Kleier, D. A., Lipscomb, W. N.: *J. Am. Chem. Soc.* **100**, 5687 (1978); (d) Sahini, V. E., Savin, A.: *Rev. Roum. Chim.* **24**, 165 (1979)
3. Edmiston, C., Ruedenberg, K.: *Rev. Mod. Phys.* **35**, 457 (1963); *J. Chem. Phys.* **43**, S97 (1965)
4. Mulliken, R. S.: *J. Chim. Phys.* **46**, 500 (1949)
5. (a) Boys, S. F.: *Rev. Mod. Phys.* **32**, 296 (1960); (b) Foster, J. M., Boys, S. F.: *Rev. Mod. Phys.* **32**, 300 (1960); (c) Boys, S. F.: *Quantum theory of atoms, molecules and the solid state*. Ed. P.-O. Löwdin. New York: Academic Press 1966
6. Halgren, T. A., Lipscomb, W. N.: *Proc. Natl. Acad. Sci. U.S.A.* **69**, 652 (1972); *J. Chem. Phys.* **58**, 1569 (1973)
7. Kleier, D. A., Dixon, D. A., Lipscomb, W. N.: *Theoret. Chim. Acta (Berl.)* **40**, 33 (1975)
8. Kleier, D. A., Halgren, T. A., Hall, J. H., Lipscomb, W. N.: *J. Chem. Phys.* **61**, 3905 (1974)
9. Heilbronner, E., Bock, H.: *Das HMO-Modell und seine Anwendung*. Weinheim: Verlag Chemie 1968; Yates, K.: *Hückel molecular orbital theory*. New York: Academic Press 1978
10. Magnasco, V., Perico, A.: *J. Chem. Phys.* **47**, 971 (1967)
11. Moyano, A., Paniagua, J. C., Tel, L. M.: *Theoret. Chim. Acta (Berl.)* **62**, 277–291 (1983) accompanying paper
12. von Niessen, W.: *J. Chem. Phys.* **56**, 4290 (1972); *Theoret. Chim. Acta (Berl.)* **27**, 9 (1972); *Theoret. Chim. Acta (Berl.)* **29**, 29 (1973)
13. For a review see Murrell, J. N., Harget, A. J.: *Semi-empirical self-consistent-field molecular orbital theory of molecules*. London: Wiley-Interscience, § 2.2, 1972
14. Balaban, A. T., ed.: *Chemical applications of graph theory*. Chapter 7, § 2, New York: Academic Press 1976
15. Hakini, S. L., Yau, S. S.: *Quart. Appl. Math.* **22**, 305 (1965); Harary, F.: *Graph theory*. Reading, Massachusetts: Addison-Wesley 1969
16. Perkins, P. G., Stewart, J. P.: *J. Chem. Soc., Faraday Trans. 2*, **78**, 285 (1982)
17. Heilbronner, E.: *Tetrahedron Lett.* 1923 (1964); Zimmermann, H. E.: *J. Am. Chem. Soc.* **88**, 1564 (1966); Karadakov, P., Enchev, V., Fratev, F., Castaño, O.: *Chem. Phys. Letters* **83**, 529 (1981)
18. Goldstein, M. J., Hoffman, R.: *J. Am. Chem. Soc.* **93**, 6193 (1971)

19. Herndon, W. C., Párkányi, C.: *Tetrahedron* **34**, 3419 (1978); Aihara, J.: *Bull. Chem. Soc. Japan* **52**, 1529 (1979); Schaad, L. J., Hess, B. A., Nation, J. B., Trinajstić, N.: *Croat. Chem. Acta* **51**, 223 (1979); Graovac, A.: *Chem. Phys. Letters* **82**, 248 (1981)
20. Aihara, J.: *J. Am. Chem. Soc.* **98**, 2750 (1976); Gutman, I., Milum, M., Trinajstić, N.: *J. Am. Chem. Soc.* **99**, 1692 (1977)

Received July 15, 1982