QUANTUM CHEMICAL STUDIES OF ACTIVITY-STRUCTURE RELATIONSHIPS

A THESIS

SUBMITTED TO

THE SCHOOL OF GRADUATE STUDIES

ADDIS ABABA UNIVERSITY

IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE IN CHEMISTRY

BY

AHMED MUSTEFA

SEPTEMBER, 1993

TO MY PARENTS

CONTENTS

ACKNOWLEDGMENTS

I offer my deepest gratitude and my respect first and foremost to my instructor and advisor Dr. Habtamu Zewdie for his consistent supervision, limitless effort and dedication in guiding this work. **His knowledge, interest and experience in the field facilitated my progress in the thesis work and aroused my interest in the field for my further research work.**

I express my deepest sense of gratitude to my friends and colleagues for their continued encouragement in my efforts to produce this research.

Finally, I would like to express my thanks to my institute, Arba Minch Technology Institute (AWTI) for sponsoring in doing my graduate studies.

AHMED MUSTEFA
SEPTEMBER 1993

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND SYMBOLS

A_μ	electron affinity
A,B	indices of atoms
AH	aromatic hydrocarbon
AO	atomic orbital
a.u.	atomic unit
CNDO	complete neglect of differential overlap
eV	electron volt
HF	Hartree-Fock
HMO	Hückel molecular orbital
HOMO	highest occupied molecular orbital
I_μ	ionization potential
LCAO	linear combination of atomic orbital
PAH	polycyclic aromatic hydrocarbons
SD	superdelocalizability
SAMO	simulated ab initio molecular orbital
SCF	self-consistent field
STO	Slater-type orbital
T	kinetic energy
V	potential energy
ZDO	zero differential overlap
α	coulomb integral
α_A	coulomb integral of rth atomic orbital or Atom
β	bond integral
$\beta_A^\circ, \beta_B^\circ$	atomic bonding parameters
β_{AB}	bond integral between atomic orbitals A and B

β_{AB}°	bond parameters of AB bond
b_{AB}	Mayer's bond order
c_{iA}	coefficient of A^{th} AO in the i^{th} MO
δ_{ij}	Kronecker delta
∇	Laplace operator
e	electronic charge
ϵ_i	one electron energy , energy of i^{th} molecular orbital
E	total energy
E_A	one-centre contribution of the total energy
E_{AB}	two-centre contribution of the total energy
E_o	ground state energy
E_{π}	π-energy
ESCA	Electron Spectroscopy for Chemical Analysis
F	matrix of the Hartree-Fock Hamiltonian operator
G	free energy
$\Delta G_{\neq}$	free energy of activation
γ_{AB}	an average electrostatic repulsion between any electrons on A and B
γ_{AA}, γ_{BB}	an average electrostatic repulsion between any two electrons on A or any two electrons on B
h	Planck's constant
$\hbar$	reduced Planck's constant
i,j,k	indices of molecular orbitals
I_K	K-region superdelocalizability index
I_L	L-region superdelocalizability index
$H_{\mu\nu}$	Hamiltonian integral between orbtials μ and ν
$\mathcal{H}$	Hamiltonian operator

k_i	a rate constant
K_i	equilibrium constant
l	angular momentum quantum number
L_b	bond localization energy
L_p	para localization energy
L_A	localization energy of atom A
m	mass, magnetic quantum number
n_k	occupation number of closed shell orbitals
p	atomic orbital with $l = 1$
p_{AB}	bond order between atoms A and B
P_{AB}^{π}	π-bond order
P_{AB}^{σ}	σ-bond order
P_{AB}^{ov}	bond overlap population
P	density matrix
s	atomic orbital with $l = 0$
S	overlap integral
$\mathbf{S}$	overlap matrix
S_A	superdelocalizability of atom A
$S_{\mu\nu}$	overlap integral between atomic orbitals μ and ν
μ, ν	indices of atomic orbitals
R_{AB}	bond reactivity
ρ	charge density
φ	atomic wave function
ψ_j	j^{th} molecular orbital
Ψ	total product wave function
Y_{lm}	spherical harmonics

ABSTRACT

The semiempirical quantum chemical method CNDO/2 was proposed to calculate the reactivity indices, atomic superdelocalizability, bond reactivity, total bond orders, π-bond order and bond overlap population. The atomic superdelocalizability was defined to include the total electron density (π- and σ-electrons). The bond reactivity was defined to include the total bond overlap density. The π-bond order was calculated using atomic orbital with π-symmetry. Test calculations were performed to establish correlations between our reactivity parameters and those which are based on the HMO method. A computer program was developed for practical implementation of the reactivity parameters. The proposed approach is applicable for studying the reactivity of a wide range of molecules, i.e., planar, non-planar, conjugated, non-conjugated, etc.

In this study, geometry optimization was performed for all molecules before the actual computation of the reactivity parameters. Parameters of reactivity for the optimum structure of 22 hydrocarbons have been calculated using CNDO/2 method within the framework of the proposed approach. This resulted in a good description of reactivity parameters. The result of the superdelocalizability indices calculation indicates that carcinogenicity of the molecules is related to the K-region superdelocalizability index ≤ -10.70 (1/a.u.) and L-region superdelocalizability ≥ -10.98 (1/a.u.). The carcinogenic molecules are found to have K-region π-bond order ≥ 0.80, K-region bond overlap population ≥ 1.026, and K-region bond reactivity ≤ -1.123.

Correlations are demonstrated relating the calculated indices with the indices obtained using other methods. The correlation of the calculated indices with carcinogenic activity is also shown. All correlation coefficients are found to be significant at 5% level.

1.0. INTRODUCTION

The main objective for undertaking this study was to develop an all-valence quantum chemical method for screening carcinogenic polycyclic hydrocarbons. Many polycyclic hydrocarbons exhibit high toxicity and are implicated as carcinogens.

The majority of human cancer is generally considered to be traceable, either directly or indirectly, to extrinsic causes and is therefore, at least in theory, preventable[1]. In fact, there is good reason to believe that a large proportion of human cancer may be attributed to environmental carcinogens, the majority of these being chemicals. Most of the recognized carcinogens or carcinogen precursors consist of polycyclic aromatics, aromatic amines and azo dyes, biological alkylating agents, or antibiotics[2].

Avoiding cancer risk from these sources requires identification of those chemicals that are potential initiators of carcinogenesis. Extensive experimental testing of potentially active chemicals may be necessary. Traditionally, such identification requires life-time feeding studies in various animal models. As a result, it is both expensive and time consuming. Accordingly, a great deal of effort is being invested in research (experimental and theoretical) for faster and more economical methods of identification of these potentially hazardous materials.

The discovery that certain aromatic polycyclic hydrocarbons can induce cancerous growth was perhaps one of the most unexpected observation made in the field of the correlation of molecular structure and biological activity[3]. After a large number of compounds were tested, it was postulated that geometrical features of aromatic carcinogens could be factors in carcinogenicity[4,5]. In addition, suggestions that metabolism of carcinogens was a probable first step in carcinogenesis[6] caused attention to be drawn to chemical reactions of aromatic hydrocarbons. Following this idea, attempts to develop a parallelism between the chemical reactivity *in vitro* and carcinogenicity have been pursued[7-10].

Schmidt[11-13] and Svartholm[14] were the first to try to establish a relationship between the electronic distribution of conjugated molecules and their carcinogenic strength. Both deduced that certain regions of aromatic hydrocarbons have high electron densities and the compounds should therefore be most susceptible to electrophilic reactions at those positions. Svartholm[14] showed that reactivity could be associated with the phenanthrenic site, i.e., the exposed corner on a phenanthrene segment, and with the mesoanthracenic site, i.e., the two central atoms in an anthracene segment. These regions were later termed as the K- and L-regions, respectively. Figure 1 illustrates the location of these regions in benz[a]anthracene.

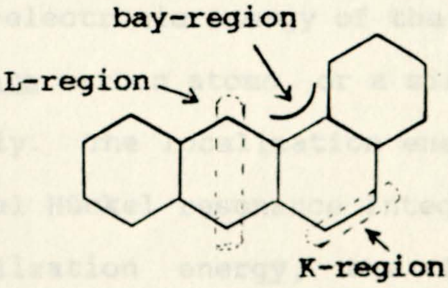


Fig. 1. K-, L-, and bay regions in benz[a]anthracene.

The first complete development of a theory of carcinogenesis is due to A. Pullman[15] and was developed into a more extended theory along with B. Pullman[16]. For a long time, this theory dominated the quantum chemical approach to understanding the carcinogenicity of PAHs. In this theory, it is demonstrated that a correlation exists between an aromatic molecule's carcinogenicity and a quantum chemical parameter of the K- and L-regions. Further, it is shown that the critical cancer-inducing reaction takes place at a bond of the K-region type, and reactions at L-region do not induce cancer, i.e., these reactions bypass the development of cancer.

Electronic indices for the π-electrons of K- and L-regions, calculated mostly by the highly empirical Hückel Molecular Orbital (HMO) method, were then correlated with experimental parameters. The K- and L-region indices are computed from the π-electron energy changes resulting from an imagined electro-philic reaction at the K- and the L-region. But, no simple index was found to be sufficient to relate the relative chemical reactivity of the two regions, and a composite index of bond localization energy (BLE), para localization energy (PLE), and carbon atom localization energy (CLE) was used. Each localiza-tion energy is the difference in π-electronic energy of the fragments remaining after a bond, two para carbon atoms, or a single carbon atom are removed, respectively. The localization energies are expressed in terms of the usual Hückel resonance integral, β. The larger a particular localization energy, the less reactive that particular region or atom will be.

The K-region index put forward by A. Pullman[15] is calculated as follows:

(i) Computed the pi-energy lost when the K-region bond is isolated from the remainder of the PAH, i.e., the bond localization energy, (BLE).

(ii) Compute the energies lost when each atom of the K-region is separated from the molecule, i.e., the carbon atom localization energy, (CLE).

(iii) Pick the smaller of the two energies from (ii) and add it to the energy from (i), [BLE + CLE(min)].

This gives the value of the K-region index. The larger this number, the less reactive the K-region bond is.

Carcinogenicity requires an I_K value of 3.31ß or less. The L-region index is calculated in a similar manner except that now the para localization energy is calculate instead of the BLE. Carcinogenicity requires an I_L value of 5.66ß or greater. That is, an aromatic hydrocarbon will be carcinogenic if:

(i) $[\mathrm{BLE} + \mathrm{CLE(min)}]_{K\text{-region}} \leq 3.31\beta$

(ii) $[\mathrm{BLE} + \mathrm{CLE(min)}]_{L\text{-region}} \geq 5.66\beta$

Mainster and Memory[17] proposed a related but computationally more convenient index which is obtained using HMO theory. The proposed indices are based upon the use of superdelocalizability (SD) to estimate relative chemical reactivity. Using the HMO method, the SD of a particular atom is described by the following formula,

$$S_\lambda = 2\Sigma_{j=1}^{occ} c_{j\lambda}^2/\lambda_j \qquad (1.1)$$

where c_{jA} is the coefficient of A^{th} atomic π orbital in the j^{th} MO, and λ_j is the coefficient in the equation $E_j = \alpha + \lambda_j\beta$ for which E_j is the energy of the j^{th} MO, and α and β are Coulomb and resonance integrals, respectively.

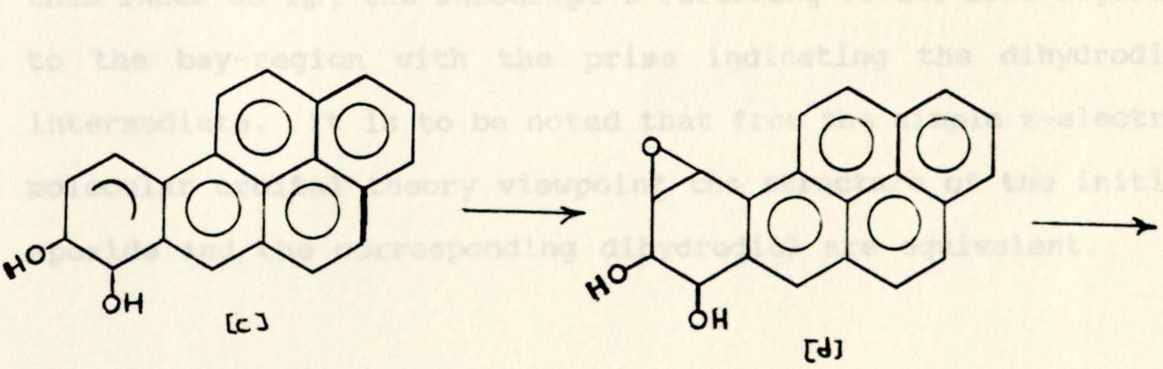
The proposed reactivity indices[17] for the K- and L-regions, I_K and I_L, are the sum of the superdelocalizabilities of the two atoms which respectively compose the K- and L-regions of a molecule, i.e. if A and B are atomic sites which compose the K-region and C and D are the atomic sties that compose the L-region then $I_K = S_A + S_B$, and $I_L = S_C + S_D$. By means of these indices, a reformulated criterion for chemical carcinogenicity was determined. For an aromatic hydrocarbon to be carcinogenic: (i) it must possess a K-region which is more reactive than 2.05(1/β);

(ii) if the molecule also possesses an L-region, this region must be less reactive than 2.30(1/β).

Duke and his coworkers[18] used the K-region π-bond overlap population index for the chemical reactivity of the region. This was obtained by the simulated ab initio molecular orbital (SAMO)[19] method, which closely approximates the results of complete ab initio MO method. The SAMO method entails transfer-ring ab initio Hartree-Fock matrix elements from calculations on small but similarly patterned molecules, thereby avoiding calculation and repeated use of extremely long lists of two-electron repulsion integrals. The basis set consisted of three sp^2 hybrid orbitals, one $2p_\pi$ orbital, and 1s core orbital for each trigonal planar carbon. Four sp^3 hybrids plus a 1s orbital were used for tetrahedral carbons, and a single 1s orbital was used at each hydrogen. The results obtained by this method indicate that the carcinogenicity of the aromatic hydrocarbons is related to K-region π-bond overlap population greater than 0.340, but there is no correlation with the σ-bond overlap population.

These chemical positions (K- and L-region) were suggested at a time when very little was known about what actually occurs between the exposure of an organism to a PAH and the occurrence of cancer. Since then, much progress has been made[20].

Later, a different picture, known as the "bay-region" theory emerged[21,22]. As Figure 1 indicates, a bay-region is the inner or unexposed corner in a phenanthrene segment. One of the end rings in the phenanthrene segment must also be a terminal ring in the parent PAH. Benzo[a]pyrene has been the most thoroughly studied molecule, and the sequence of enzyme-assisted chemical transformations thought to lead to the induction of tumorigenesis is shown in Figure 2.



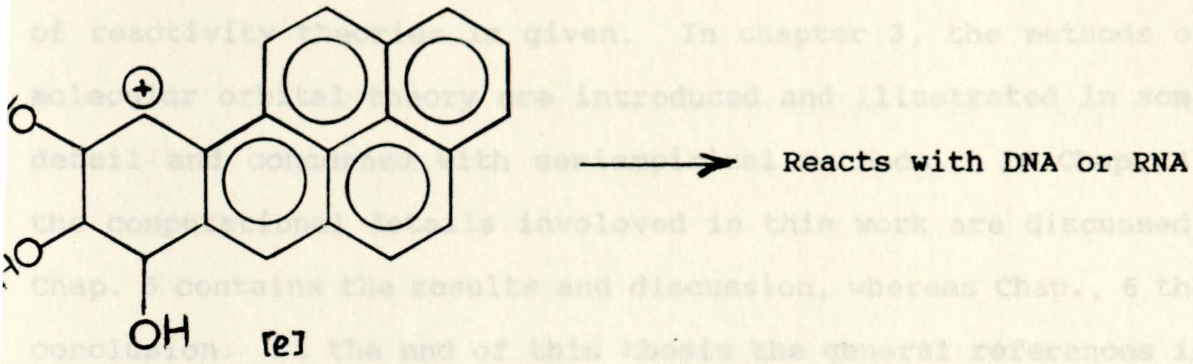


Fig. 2. Benzo[a]pyrene forms in succession (a) an epoxide (b) a dihydrodiol (c) ("proximate carcinogen"), a dihydrodiol-epoxide (d) ("ultimate carcinogen"), and an unstable triol carbocation intermediate which (e) reacts with DNA or RNA.

Recognition of the diol-epoxide (d) of benzo[a]pyrene as the "ultimate carcinogen" of this PAH provided important direction for understanding the wide range of tumour initiating activity of different PAH. In particular, Jerina and co-workers[21,22] have shown that a correlation exists between observed PAH carcinogenicity and computed ease of going from the diol-epoxide (d) to the triol carbocation form (e).

Berger and co-workers[23] suggested that the formation of the dihydrodiol (c) (Fig. 2) causes the nearby 9,10-region bond to become highly activated, thereby preparing it for epoxidation. Reactivity at this latter bond was represented using SD indices of the type employed by Mainster and Memory[17]. They designated this index as $I_{B'}$, the subscript B referring to the bond adjacent to the bay-region with the prime indicating the dihydrodiol intermediate. It is to be noted that from the simple π-electron molecular orbital theory viewpoint the structure of the initial epoxide and the corresponding dihydrodiol are equivalent.

In the following chapter, a brief description of the basis of reactivity theories is given. In chapter 3, the methods of molecular orbital theory are introduced and illustrated in some detail and concerned with semiempirical methods. In Chap. 4, the computational details involoved in this work are discussed. Chap. 5 contains the results and discussion, whereas Chap., 6 the conclusion. At the end of this thesis the general references is given.

2.0. CHEMICAL REACTIVITY

2.1. REACTIVITY THEORIES:

In the theory of absolute reaction rates[24], the rate, v, of a reaction of the type $A + B \rightarrow C + D$ is given by

$$v = k_f[A][B] \qquad (2.1)$$

k_f is the forward specific rate constant given by

$$k_f = \kappa(kT)/h \exp(-\Delta G^{\ddagger}/RT) \qquad (2.2)$$

where k is Boltzmann's constant, κ is a transmission coefficient, h is Planck's constant, R is the molar gas constant, T is the temperature, and $\Delta G^{\ddagger}$ is the difference in free energy between the transition state and the reactants, all referring to their standard states. [A] and [B] in eqn. (2.1) are the concentrations of reactants A and B, respectively. The transition state is the most unstable configuration through which the molecules of reactants must pass on their way to becoming products (Fig. 3). Thus, at a given temperature, the free energy of activation, $\Delta G^{\ddagger}$, determines indirectly the rate of a reaction[24], v, through eqn. (2.2). For an irreversible reaction $\Delta G^{\ddagger}$ is given by:

$$-RT\log k_f = \Delta G^{\ddagger} \qquad (2.3)$$

The accurate calculation of a free energy change ΔG is clearly impossible, for, since $G = U - TS + PV$, ΔG includes an entropy change, (and hence requires detailed knowledge of vibrational partition functions etc.), as well as the energy of all the electrons, together with solvent interactions etc. In many cases, however, we study not absolute reaction rates but the rate of one reaction relative to that of another rather similar reaction. If we are prepared to assume that similar changes of entropy, solvation energy, etc. occur, then the difference between the free-energy changes, ΔG for one reaction and ΔG° for a 'standard' reaction, may be assumed to arise principally from the difference between ΔU and ΔU°. Since ΔU is simply the total electronic energy change ΔE (i.e. the relevant part of the internal energy change) for one mole we may then write, for a typical irreversible reaction,

$$-RT\log(k/k_\circ) = \Delta E - \Delta E^\circ \qquad (2.4)$$

For conjugated systems it is usual (when comparing similar conjugated systems in a given type of reaction) to assume that the change in the σ-electron change is the same in both systems, and as a result the difference $\Delta E - \Delta E_\circ$ arises from the difference of π-electron energy changes.

The π-energy variation, as a given substituent approaches two alternative reaction sites can be represented, by the two curves in Fig. 3.

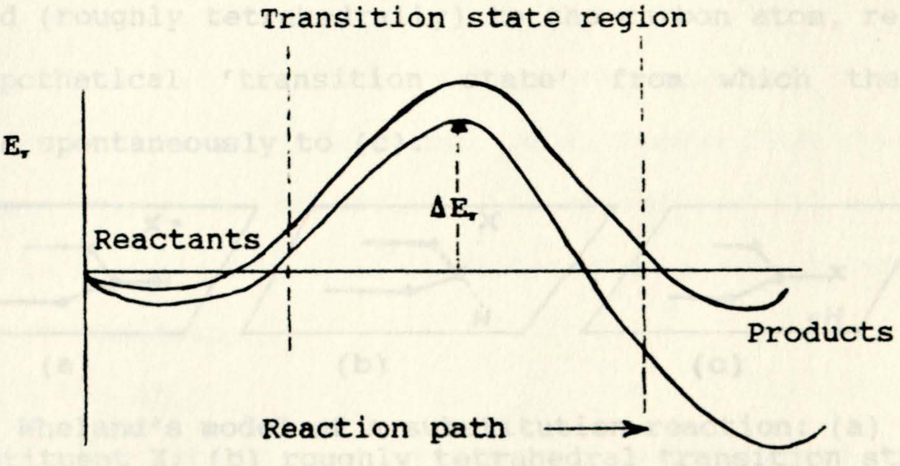


Fig. 3. Variation of π-electron energy for two substitution positions.

The aim of all theoretical discussions of aromatic substitution is to estimate the forms of the reaction path for alternative positions of substitution. There are two main approaches, depending on whether we look at the initial parts of the curves (the 'isolated molecule approximation') or at the transition state region (the 'localization approximation'), and in both approaches we shall find that the energy changes are related to the charges and bond orders.

2.2. THE LOCALIZATION APPROXIMATION

In this approach, one assumes that activation energies can be compared by considering an intermediate or "activated" state complex to transition state. For aromatic substitution reactions which proceed by attack at one atom, the most widely adopted model of the transition state is that due to Wheland[25]. The substitution proceeds through the three proposed stages indicated in Fig. 4. The conformation (b), in which both X and H are attached (roughly tetrahedrally) to the carbon atom, represents the hypothetical 'transition state' from which the system proceeds spontaneously to (c).

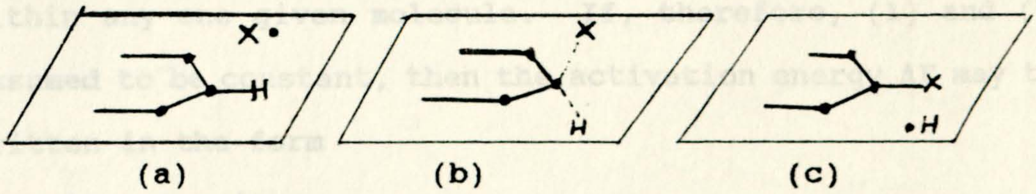


Fig. 4. Wheland's model of a substitution reaction: (a) approach of substituent X; (b) roughly tetrahedral transition state; (c) expulsion of hydrogen.

X may be a radical or an ion, but in any case the carbon is removed from the conjugation and 0, 1, or 2 π-electrons go into the localized C-X bond. When the same substitution occurs at different positions the changes in σ-electron energy are assumed to be similar (involving similar local changes in the molecular framework), but the changes in π-electron energy may be very different, depending on the shape and size of the conjugated system. In such cases, therefore, the interpretation of each ΔE in eqn. (2.4) as a change of π-electron energy is reasonable; the other terms cancel.

In order to achieve this arrangement, it is evident that three energy changes are needed:

(1) Energy would be required to change carbon-atom r from a trigonal state to tetrahedral one.

(2) Energy would be required to establish a C_r-X bond.

(3) Some π-electron energy would have been lost as a result of what might be called a limited π "volume", ΔE_r, which is the extent of the network available to the π-electrons.

The first two of the above are not expected to vary very much, either from molecule to molecule, indeed from atom to atom within any one given molecule. If, therefore, (1) and (2) are assumed to be constant, then the activation energy ΔE may then be written in the form

$$\Delta E = C + \Delta E_r \tag{2.5}$$

where C is a constant which arises as a result of the assumption made to the energies in step (1) and (2) above.

For a given type of substitution reaction, the activation energy, ΔE′, and the rate of reaction will be determined by differences in ΔE_r. This was proposed by Wheland as a react-ivity index which has since become known as a localization energy[26]. The energy of reaction, ΔE, is usually calculated by the HMO method by taking the difference in π-electron binding energies, and appear as multiple of the parameter β, i.e.,

$$\Delta E_r = E_{r(\text{residual molecule})} - E_{r(\text{original molecule})} \qquad (2.6)$$

For example, benzene has HMO π-energy of 6α + 8β. Localization of an electron pair at position 1 leaves a pentadienyl cation whose HMO π-energy is 4α + 5.464β. ΔE_r = 5.464 - 8 = -2.536β. And in the case of semiempirical SCF methods (CNDO/2) the energies are calculated unambiguously, in a.u., electron volt or kilocalorie/mole, and allowance is made for the changes in geometry on passing from the neutral hydrocarbon to the corresponding ion.

The localization energy method is equally useful for reactions other than substitution reactions such as relatively complicated addition reactions. The best example is the Diels-Alder addition, Fig. 5.

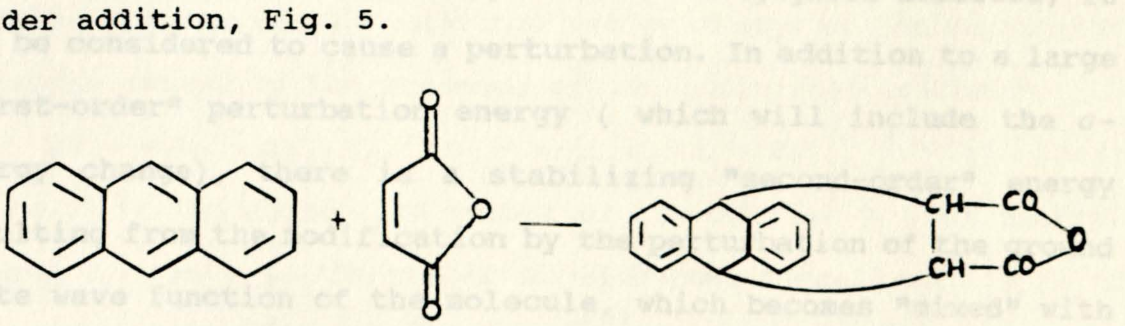


Fig. 5. Diels Alder addition on anthracene.

Brown[27] pointed out that the energy difference between the initial and transition state should then be treated as the difference in π-electron energy between the parent aromatic hydrocarbon and aromatic or pair of aromatic hydrocarbons, obtained by removing the two carbon atoms at which addition takes place. This energy difference is called the para-localization energy, L_{p},.

In a similar fashion, ortho localization energies[27] or bond localization energies have been defined for reactions such as the addition of osmium tetroxide to aromatics, leading to the formation of dihydrodiols.

2.3. ISOLATED MOLECULES APPROXIMATION[28]

In isolated molecules approximation the model of the transition state is generally one in which the aromatic π-system is perturbed to a relatively small degree by the attacking reagent. The tendency for reaction to occur at different centres is determined by one of the several indices defined in terms of the MO's of the isolated hydrocarbon. One of these indices is atomic superdelocalizability, which was derived by Fukui and his co-workers[29]. As a reagent approaches a conjugated molecule, it can be considered to cause a perturbation. In addition to a large "first-order" perturbation energy (which will include the σ-energy change), there is a stabilizing "second-order" energy resulting from the modification by the perturbation of the ground state wave function of the molecule, which becomes "mixed" with excited states. The largest contribution to the second-order perturbation energy is obtained by mixing the ground state wave function with the lowest possible lying excited state, usually

that which involves the lowest antibonding π orbital φ_{m+1} (for closed-shell even molecule). The mixing will be greatest, and hence the total activation energy lowest (1) for those molecules for which the "excitation energy" to φ_{m+1} is smallest; (2) if the orbital φ_{m+1} "offers" a large amplitude at the location of the perturbing reagent.

In the localization energy theory hyperconjugation between the pseudo-π-orbital XH (Fig. 6) and the carbon π-orbtials on the ring is neglected. But, this hyperconjugation is found to be important in decreasing the activation energy[30]. The corresponding stabilization energy can be measured as reactivity index, the "superdelocalizability" S_A as follows.

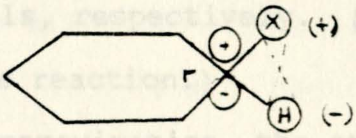


Fig. 6. Pseudo-π orbital in the σ complex.

Let α_h be the Coulomb integral for the pseduo-π orbital which can hyperconjugate with the system, and γ the resonance integral for the interaction of this orbital with the adjacent atomic orbital φ_A. **The total π-electron energy change of the conjugated molecule caused by the presence of the additional orbital[29]**

$$\Delta E = \nu(\alpha_h - \alpha) - \Sigma_{j=1}^{2m}(\nu_j - \nu)c_{jA}^2\Delta^2/(\alpha_h - E_j), \qquad (2.7)$$

where ν_j is the occupation number of the orbital ϕ_j (of orbital **energy E_j) and ν is the number of electrons "localized" on XH (0** for and electrophilic reagent, 1 for radical reagent, 2 for a **nucleophilic reagent). The quantity c_{jA} is the A^th AO coefficient** in ϕ_j. The energy change (2.7) is chosen as measure of the **activation energy in the reaction, and only the negative second-**

order term varies if we consider a single type of reagent acting on different molecules. The reactivity is then measured by the coefficients of Δ^2 in (2.7), and, furthermore, α_b is assumed to be equal to α, a rather drastic assumption. For an even hydrocarbon with occupied, levels 1, 2, ..., m and unoccupied levels m + 1, ..., 2m, the superdelocalizabilities for the different type of reaction are then

$$S_A^{\text{electrophilic}} = 2\Sigma_{j=1}^{m} C_{Aj}^2/\lambda_j, \tag{2.8}$$

$$S_A^{\text{radical}} = \Sigma_{j=1}^{m} C_{Aj}^2/\lambda_j + \Sigma_{j=m+1}^{2m} C_{Aj}^2/(-\lambda_j), \tag{2.9}$$

$$S_A^{\text{nucleophilic}} = 2\Sigma_{j=m+1}^{2m} C_{Aj}^2/(-\lambda_j). \tag{2.10}$$

where λ_j is the coefficient in the equation $\epsilon_j = \alpha + \lambda_j\beta$, ϵ_j is the energy of the j^th^ molecular orbital, and α and β are coulomb and resonance integrals, respectively. (The larger the magnitude of S_A, the easier the reaction.)

Under this approximation, the other reactivity index is the π-bond reactivity which is obtained by the following expression:[31]

$$R_{AB} = \Sigma_k^{occ} \rho_{AB}^{\pi}(k)/\epsilon(k) \tag{2.11}$$

where $\rho_{AB}(k)$ is the π-bond overlap density between atoms A and B in k^th^ molecular orbital, given by

$$\rho_{AB}^{\pi}(k) = c_{kA}c_{kB} \tag{2.12}$$

and $\epsilon(k)$ is the total energy of MO k in units of a.u. or eV. Where a.u. is the atomic unit of energy (called Hartree or Rydberg, Ry). c_{kA} and c_{kB} in eqn. (2.9) are the coefficients of atomic π-orbitals of A and B in the k^th^ MO, respectively.

3.0. MOLECULAR ORBITAL THEORY

The main objective of any theory of molecular structure is to provide some insight into the various physical laws governing the chemical constitution of molecules in terms of some fundamental universal physical laws governing the motions and the interaction of the constituents.

Let us recall the most important features of molecular quantum theory which are assumed to be known. Molecules consist of electrons and nuclei which are considered as points. Electrons and nuclei are characterized by their mass, charge and intrinsic angular momentum, the spin. The charges give rise to the Coulomb potential which is the only primary interaction considered in the theory of isolated molecules. The electronic charge is -e, the charge of a nucleus with atomic number Z is +Ze. The spin of an electron may be ±ħ/2 corresponding to the "spin up" and "spin down" states, α and β, respectively.

The physical state of a molecule is represented by the wave function, $\Phi(x,t)$, which depends on the spatial and spin coordinates of all particles and on the time (t). The space of all coordinates (dynamic variables) is called configuration space, a point in this space is denoted by x. The wave function, Φ, has some special mathematical properties which are characteristic of the elements of an (abstract) Hilbert Space. This is space of the square-integrable functions, for which the integral $\int\Phi^*\Phi dx$ is finite. In addition, it is assumed that Φ is differentiable and regular. The above conditions are valid only for bound states.

Physical quantities are represented by Hermitian operators over the Hilbert space. Expectation values of a given physical quantities corresponding to operator A, are given by the expression,

$$\langle A\rangle = \int\Phi^{*}A\Phi dx/\int\Phi^{*}\Phi dx = \langle\Phi|A|\Phi\rangle/\langle\Phi|\Phi\rangle \qquad (3.1)$$

The operator of the energy is called the Hamiltonian. By neglecting the weak magnetic interactions (spin-spin, spin-orbit), one obtains the non-relativistic molecular Hamiltonian, $\mathcal{H}$, which has the simple form analogous to that in classical mechanics,

$$\mathcal{H} = \hat{T} + \hat{V} \qquad (3.2)$$

where T is the sum of the kinetic energy operators of all particles and $\hat{V}$ is the potential energy operator. For an isolated molecule we can write in the usual coordinate representation

$$\hat{V} = \Sigma_{i<k}Q_iQ_k/|r_i - r_k| \qquad (3.3)$$

and

$$\hat{T} = -\hbar^2/2(\Sigma_i\nabla_i^2/m_i) \qquad (3.4)$$

where Q_i is the charge of particle; i, r_i, and r_k are the position vectors of particles i and k, respectively; m_i is the mass of particle i, and ∇_i^2 is the Laplace operator corresponding to the coordinates of particle i. In the presence of an external field the operator $\hat{V}$ in eqn. (3.3) has to be modified appropriately. The time evaluation of the state is given by the Schrődinger equation,

$$H\Phi = -i\hbar\partial\Phi/\partial t. \qquad (3.5)$$

Stationary states of the molecules, Φ_n, can be obtained from the time-independent Schrödinger equation,

$$\mathcal{H}\Phi = E_n\Phi_n \qquad (3.6)$$

The n^{th} eigenvalue of H, E_n is the energy of the molecule in the n^{th} eigenstate.

The Schrodinger equation, eqn. (3.6), for a molecule consisting of n electrons and N nuclei interacting with each other is written as

$$\mathcal{H}^{total}(1,\ldots,N;1,\ldots,n)\Psi(1,\ldots,N;1,\ldots,n) = E\Psi(1,\ldots,N;1,\ldots,n) \qquad (3.7)$$

where Ψ is the complete wave function of all particles and E is the total energy of the molecule; 1,2,..N and 1,2,...,n are spatial coordinates of the nuclei and electrons, respectively, $\mathcal{H}^{total}$ is the sum of the kinetic energy operators for the nuclei and for the electrons together with the potential energy terms representing the various Columbic interactions. These are repulsive for electron-electron and nucleus-nucleus pairs, but attractive between electron-nucleus pairs. For such molecules, the total molecular Hamiltonian operator $\mathcal{H}^{total}$ is

$$\mathcal{H}^{total}(1,\ldots,N;1,\ldots,n) = -h^2/8\pi^2\Sigma_A{}^N M_A{}^{-1}\nabla_A{}^2 + \Sigma_{A<B}e^2Z_AZ_Br_{AB}{}^{-1} - h^2/8\pi m^2\Sigma_p{}^n\nabla_p{}^2 - \Sigma_A\Sigma_p e^2Z_Ar_{Ap}{}^{-1} + \Sigma_{p<q}e^2r_{pq}{}^{-1} \qquad (3.8)$$

Here M_A is the mass of nucleus A; m and e are the electronic mass and charge, respectively; Z_Ae is the charge on nucleus A; and r_{ij} is the distance between particles i and j. Summations involving indices A and B are over nuclei and those involving p and q are over electrons. Since each particle is described by three cartesian coordinates, eqn. (3.7) will be a differential equation in 3N + 3n variables.

The full Schrödinger equation for any molecular system will have an infinite number of solutions, only certain ones of which are acceptable. In quantum chemistry we are mostly interested in solving eqn. (3.6). The solutions in this case can be made unique (up to an irrelevant phase factor) by the following boundary and normalization conditions:

$$\Psi_n(\pm\infty) = 0 \qquad (3.9a)$$

$$\langle\Psi_n|\Psi_n\rangle \equiv \int\Psi^*_n(x)\Psi_n(x)dx = 1 \qquad (3.9b)$$

This means that the wave function must vanish asymptotically with increasing (at least one) of the spatial variables. Equation (3.9a) is consistent with the probability interpretation. The fulfilment of eqn. (3.9b) is necessary to ensure square integrability of Ψ_n.

In practice, rather than attempting to find a wave function describing both electronic and nuclear motion together, it is usually sufficient to break the problem into two parts and consider first the motion of electrons in the field of stationary nuclei. There is then a separate, purely electronic problem for each set of nuclear positions. This is a reasonable procedure because the masses of the nuclei are several thousands times larger than the masses of the electrons. The nuclei move much more slowly, and we may reasonably suppose the electrons to adjust themselves to new nuclear positions so rapidly that at any one instant their motion is just as it would be if the nuclei were at rest at the positions they occupy at the same instant. This simplification is referred to as the Born- Oppenheimer approximation[32]. In more quantitative terms, the Born-Oppenheimer approximation amounts to separating the nuclear K.E. and nuclear repulsion terms from $\mathcal{H}^{total}$, and considering only the part of the Hamiltonian which depends on the positions but not the momenta of the nuclei. This is the electronic Hamiltonian operator H^{el},

$$H^{el} = -h^2/8\pi m^2 \Sigma_p \nabla_p^2 - \Sigma_A \Sigma_p e^2 Z_A r_{Ap}^{-1} + \Sigma_{p<q} e^2 r_{pq}^{-1}. \tag{3.10}$$

The electronic Hamiltonian may be used in the modified Schrödinger equation,

$$H^{el}(1,2,...,n)\Psi^{el}(1,2,...,n) = \epsilon\Psi^{el}(1,2,...,n) \tag{3.11}$$

the solutions of which are purely electronic wave function Ψ^{el}, describing the motion of the electrons in the field of fixed nuclei. Under the Born-Oppenheimer approximation the total energy E of the system of a given internuclear distance is given as

$$E = \epsilon + \Sigma_{A<B} e^2 Z_A Z_B r_{AB}^{-1} \tag{3.12}$$

where ϵ is the electronic energy and the second term is the electrostatic internuclear repulsion energy. Molecular orbital theory is concerned with electronic wave functions only, and henceforth the superscript el on the Hamiltonian operator and wave function is dropped without ambiguity. Equation (3.11) is a fundamental equation of quantum chemistry and various approaches that are used to solve it led to the emergence of an immense variety of quantum chemical methods.

In dealing with the equations of quantum mechanics, it is convenient to introduce new units which are appropriate to atomic dimensions and which eliminate some of the constants, e, m, and ℏ (the charge and mass of electron, and Planck's constant divided by 2π, respectively) from the wave function. These are called atomic units (a.u.). By an appropriate choice of the units of physical quantities one can make the values of the above quantities equal to unity, $e = 1$, $m = 1$ and $\hbar = 1$. By this choice the unit of length is the so called Bohr radius (or Bohr), denoted by a_0: $a_0 = \hbar^2/me^2 = 0.529107 \times 10^{-8}$cm while the atomic unit of energy (called Hartree or Rydberg, Ry) is $\epsilon_0 = e^2/a_0 = 4.3598 \times 10^{-18}$J. The atomic unit of mass is the electron mass, $m = 9.0191 \times 10^{-28}$g. The atomic unit of electronic charge is the protonic charge, $e = 4.80298 \times 10^{-10}$esu.

The electronic Hamiltonian operator of eqn. (3.10) reduces in a.u. to

$$H = -\Sigma_p 1/2\nabla_p^2 - \Sigma_A \Sigma_p Z_A r_{Ap}^{-1} + \Sigma_{p<q} r_{pq}^{-1} \tag{3.13}$$

The use of a.u. has two advantages. First the expression becomes simpler. Secondly, if the results of calculations are expressed in a.u. units, then they are independent of the actual values of these universal constants which are determined experimentally and may be changed occasionally when the measurement techniques become more precise. For this reason, we shall use these units here unless otherwise specified.

There are two general approaches to obtain approximate solutions of eqn. (3.11). The first approach is known as orbital approximation which attempts to construct a satisfactory approximate molecular (many electron) wave function from a combination of functions each dependent upon the coordinates of one electron only. For an n-electron system, the simplest way to do this is to associate the n-electrons with n one-electron wave functions $\varphi_1, \varphi_2, \ldots, \varphi_n$ and write the total wave function

$\psi(1,2,\ldots,n)$ as a product of the one-electron wave functions,

$$\psi(1,2\ldots,n) = \varphi_1(1)\varphi_2(2)\ldots\varphi_n(n) \tag{3.14}$$

where φ_i is a one-electron wave function and is called an orbital. The total electronic wave function of the products of one-electron wave functions as such is known as a Hartree product[33].

In this approach, it is assumed that the motion of a single electron is governed by the average field created by the nuclei and the other electrons. Thus, for each electron it is possible to write a one-electron Schrödinger equation as:

$$[-1/2\nabla^2_i + \Sigma_A Z_A r_{iA}^{-1} + \Sigma_{i(\neq j)} r_{ij}^{-1}]\varphi_i = \epsilon_i\varphi_i \tag{3.15}$$

where r_{ij} is average interelectronic distance between i and j, and φ_i is the one-electron wave function for electron i.

Although an orbital as such gives a complete specification of the spatial distribution on an electron, it is still incomplete in that it does not specify the state of electron spin. Thus, the complete wave function for a single electron is a product of a spatial function and a spin function, $\varphi_i(r)\eta(\zeta)$, called a spin orbital. A given spatial orbital φ_i may be associated with an α or β spin function, giving rise to the two spin orbitals $\varphi_i(r)\alpha(\zeta)$ and $\beta(\zeta)$, where ζ is spin coordinate. A product wave function including electron spin is obtained directly as a Hartree product of spin orbitals:

$$\psi(1,2,\ldots,n) = \varphi_1(1)\alpha(1)\varphi_2(2)\beta(2)\ldots\varphi_n(n)\beta(n) \tag{3.16}$$

and often written in the contracted notation

$$\psi(1,2,\ldots,n) = \varphi_1(1)\varphi_2(2)\ldots\varphi_{n-1}(n-1)\varphi_n(n) \tag{3.17}$$

where 1 signifies the appropriate spatial and spin coordinates of electron 1, etc., unbarred orbitals φ_i have α spin functions, and barred orbitals φ_1 have β spin functions (defined as ±1/2 in units of $\hbar$, respectively).

Further, the electronic wave function must satisfy the anti-symmetry condition. The anti-symmetry principle states that the electronic wave function (spin orbital) must be anti-symmetric with respect to mutual permutation of a pair of electrons (interchange of their spatial and spin coordinates) in accordance with Pauli Principle.

For the n-electron system instead of a simple Hartree product $\varphi_1, \varphi_2, \ldots, \varphi_n$ one should have the following determinant:

$$\Phi_n = 1/\sqrt{n!} \begin{vmatrix} \varphi_1(1)\alpha(1) & \varphi_1(1)\beta(1) & \varphi_2(1)\alpha(1) & \ldots & \varphi_n(1)\beta(1) \\ \varphi_1(2)\alpha(2) & \varphi_1(2)\beta(2) & \varphi_2(2)\alpha(2) & \ldots & \varphi_n(2)\beta(2) \\ \vdots & & & & \\ \varphi_{n-1}(n-1)\alpha(n-1). & & . & & .\ \varphi_{n-1}(n-1)\beta(n-1) \\ \varphi_n(n)\alpha(n) & . & . & . & \varphi_n(n)\beta(n) \end{vmatrix} \tag{3.18}$$

where the factor $1/\sqrt{n!}$ arises from normalization. The above determinantal wave functions are called Slater-determinants[34].

Equation. (3.18) is often written in short-hand notation as

$$\Phi_n = 1/\sqrt{n!}\ |\varphi_1\varphi_1\varphi_2\varphi_2 \ldots \varphi_n\varphi_n| \tag{3.19}$$

where the spin components α and β are defined as ± 1/2 in units of $\hbar$, respectively and $\varphi = \varphi\alpha$ and $\varphi = \varphi\beta$.

Slater determinants give only the means to express the total wave function as a product of spin orbitals. To obtain the best approximate solution of eqn. (3.11) by this method, variational principle is used. The variational principle is based on the variational theorem which states that the average energy, E, of a system calculated using a well-behaved trial wave function, Ψ, is always greater than or equal to the true energy E_0 obtained from exact the wave function Ψ_o.

Mathematically, for an arbitrary function Ψ, the variational theorem is expressed as

$$E = \langle\Psi|\mathcal{H}|\Psi\rangle / \langle\Psi|\Psi\rangle \geq E_o \qquad (3.20)$$

where equality holds when the trial function is the same as the exact wave function (ground state). A very common use of the variation method is with a linear combination of fixed functions $\varphi_1, \varphi_2, \ldots$

$$\psi(c_1, c_2, \ldots) = c_1\varphi_1 + c_2\varphi_2, \ldots \qquad (3.21)$$

The given functions φ_i are often referred to as a basis or basis functions. If the basis functions are linearly independent (i.e., if none can be written as a linear combination of the others), the variational method then leads to a set of approximate energies and wave functions.

The second approach, most rewarding to date, is to seek combinations of atomic orbitals which will be good approximations of the molecular orbitals of the system, the simplest being a simple sum with appropriate linear weighting coefficients. Considering a set of atomic functions φ_μ, $\mu = 1, 2, \ldots$, associated with the various atoms of the molecule, one can try to represent any particular molecular orbital ψ_i as

$$\psi_i = c_{11}\varphi_1 + c_{21}\varphi_2 + c_{31}\varphi_3 + \ldots \qquad (3.22)$$

where the $c_{\mu i}$ are numerical coefficients which may be of either sign and may be real or complex numbers. This type of expansion is known as a linear combination of atomic orbitals[35], LCAO. Expansions of the LCAO type thus provide a mathematical framework for detailed calculations, with the actual computation of the molecular wave function for the system reduced to the determination of the linear expansion coefficients, $c_{\mu i}$, for each of the orbitals.

Intuitively, the nature of chemical problems makes it profitable to relate MOs to the corresponding AOs of the constituent atoms. However, to carry out such MO calculations, a convenient analytical form for the AO of each type of atom in the molecule is required. The solutions of the Schrodinger equation for one-electron atomic systems can be written in the form[35]

$$\psi(r,\theta,\varphi) = R_{nl}(r)Y_{lm}(\theta,\varphi) \qquad (3.23)$$

with

$$l = 0,1,\ldots,n-1 \qquad (3.24a)$$

$$|m| \leq l \qquad (3.24b)$$

where r,θ and φ are the spherical polar coordinates centred on the atom. The angular parts $Y_{lm}(\theta,\varphi)$ are the well-known spherical harmonic, defined as

$$Y_{lm}(\theta,\varphi) = \theta_{lm}(\theta)\Phi_m(\varphi). \qquad (3.25)$$

The radial part of the atomic functions $R_{nl}(r)$ are polynomials in the radial distance r multiplied by decaying exponent $e^{-\zeta r}$, where ζ is the orbital exponent.

Some general aspects of the orbital description of electronic structure have been considered above. Following a more detailed discussion of the actual calculation of orbitals for many electron system.

The variational approach to approximate solution of the Schrödinger equation involves working with the energy expectation alue $\langle\Psi|\mathcal{H}|\Psi\rangle$, and for algebraic manipulation it is necessary to ave a convenient expression for this quantity in terms of the rbitals involved. Generalizing eqn. (3.19) to a closed shell orm with 2n electrons, the orbital wave function may be written n the form

$$\Phi_n = 1/\sqrt{2n!}\,|\varphi_1,\varphi_2,\varphi_3,\ldots,\varphi_{2n}| \tag{3.26}$$

he orbitals φ_i may be considered orthonormal without loss of enerality, i.e.,

$$S_{ij} = \int\varphi_i(1)\varphi_j(1)d\tau_1 = \delta_{ij} \tag{3.27}$$

Now we proceed to the evaluation of the energy expectation value $\langle\Psi|\mathcal{H}|\Psi\rangle$, where Ψ is a determinantal wave function. The Hamiltonian operator may be separated into two parts,

$$\mathcal{H} = \mathcal{H}_1 + \mathcal{H}_2 \tag{3.28}$$

where

$$\mathcal{H}_1 = \Sigma_p H^{core}(p) \tag{3.29}$$

with

$$\mathcal{H}^{core}(p) = -1/2\nabla_p^2 - \Sigma_A Z_A r_{pA}^{-1} \tag{3.30}$$

and

$$\mathcal{H}_2 = \Sigma_{p<q} r_{pq}^{-1} \tag{3.31}$$

The quantity $\mathcal{H}^{core}$ is the one-electron Hamiltonian corresponding to motion of an electron in the field of the bare nuclei, the charge of nucleus A being Z_A.

Substituting eqn. (3.28) into the energy expectation value allows separation of the ϵ^{el} into one- and two-electron parts,

$$\langle\Psi|\mathcal{H}|\Psi\rangle = \langle\Psi|\mathcal{H}_1|\Psi\rangle + \langle\Psi|\mathcal{H}_2|\Psi\rangle \tag{3.32}$$

which are conveniently treated separately. For the one-electron part, using eqn. (3.29),

$$\langle\Psi|\mathcal{H}_1|\Psi\rangle = \Sigma_p^{2n}\langle\Psi|\mathcal{H}^{core}(p)|\Psi\rangle. \tag{3.33}$$

Now since electrons are indistinguishable and are treated on an equal footing in Ψ, the expectation value of $\mathcal{H}^{core}(p)$ must be the same for all 2n values of p. Thus, we need only to consider $\mathcal{H}^{core}(1)$, noting that

$$\langle\Psi|\mathcal{H}_1|\Psi\rangle = 2n\langle\Psi|\mathcal{H}^{core}(1)|\Psi\rangle. \tag{3.34}$$

After substituting the full expansion for Ψ and performing a series of calculations the final result becomes

$$\langle\Psi|\mathcal{H}_1|\Psi\rangle = 2\Sigma^n_{i=1}\mathcal{H}_{ii} \tag{3.35}$$

where $\mathcal{H}_{ii}$ is the expectation value of the one-electron core Hamiltonian corresponding to the molecular orbital

$$\mathcal{H}_{ii} = \langle\psi_i(1)|\mathcal{H}^{core}|\psi_i(1)\rangle \tag{3.36}$$

The factor two in eqn. (3.35) corresponds to the fact that there are two electrons in each MO ψ_i.

The expectation value of the two-electron Hamiltonian, $\mathcal{H}_2$ can be evaluated in a similar manner. Here if electrons are assigned to two different spatial MO's, ψ_i and ψ_j, both may have α and β spin and there will be four conditions each equal to $1/2J_{ij}$, where

$$J_{ij} = \int\int\psi_i^*(1)\psi_j^*(2)1/r_{12}\psi_i(1)\psi_j(2)d\tau_1 d\tau_2 \tag{3.37}$$

this being a six-dimensional integral over space coordinates only. If electrons 1 and 2 are assigned to the same MO ψ, they must have opposite spins and there are only two terms $1/2J_{ii}$. The total contribution is thus

$$2\Sigma_i\Sigma_{j(\neq i)}J_{ij} + \Sigma J_{ii} \tag{3.38}$$

If 1 and 2 are assigned to the different spatial orbitals ψ_i and ψ_j, there are two terms equal to $-1/2K_{ij}$, where

$$K_{ij} = \int\int\psi_i^*(1)\psi_j^*(2)1/r_{12}\psi_j(1)\psi_i(2)d\tau_1 d\tau_2 \tag{3.39}$$

If electrons 1 and 2 are assigned to the same spatial orbital, they must have different spin and the corresponding integral vanishes by integration over spin coordinates.

By collecting terms, the final expression for the electronic energy is

$$\epsilon = 2\Sigma_i^n \mathcal{H}_{ii} + \Sigma_i^n J_{ii} + \Sigma_i^n \Sigma_{j(\neq i)} (2J_{ij} - K_{ij}). \qquad (3.40)$$

Alternatively, noting that $K_{ii} = J_{ii}$, this may be rearranged into the more compact form

$$\epsilon = 2\Sigma_i^n \mathcal{H}_{ii} + \Sigma_i^n \Sigma_j^n (2J_{ij} - K_{ij}). \qquad (3.41)$$

This important formula has reduced the many electron integration to the set of three- and six-dimensional integrals $\mathcal{H}_{ii}$, J_{ij} and K_{ij}. J_{ij} and K_{ij} are known as Coulomb and exchange integrals, respectively.

It is useful to define a set of one-electron orbital energies ϵ_i,

$$\epsilon_i = \mathcal{H}_{ii} + \Sigma_j^n (2J_{ij} - K_{ij}). \qquad (3.42)$$

This is essentially the energy of an electron in ϕ_i interacting with the core and the other 2n-1 electrons. With the assumption that there is no reorganization of the other 2n-1 electrons on ionization, $-\epsilon_i$ may be associated with the ionization potential of an electron in ϕ_i. This is sometimes referred to as a Koopmans[37], or vertical, ionization potential. Using the orbital energies, the total electronic energy can then be written in the useful alternative forms

$$\epsilon = 2\Sigma_i^n \epsilon_i - \Sigma_i^n \Sigma_j^n (2J_{ij} - K_{ij}) \qquad (3.43)$$

or

$$\epsilon = \Sigma_i (\epsilon_i + \mathcal{H}_{ii}). \qquad (3.44)$$

It should be noted that the total electronic energy ϵ is not equal to the sum of the one-electron energies. This is because the sum of one-electron energies includes each electron-electron interaction twice (the repulsion between 1 and 2 contributes to the one-electron energies associated with both electrons). The second term in eqn. (3.43) corrects for this. Having established the proper form for the many-electron wave function for closed shells as a single determinant of spin orbitals and developed a convenient expression for the electronic energy, the details of the actual determination of the spatial orbitals φ_i for a closed shell system will be discussed next. If no restriction (other than orthonormality) is imposed on these functions (i.e., if they are completely flexible functions of the coordinates of one electron), then we can deduce differential equations for the optimum forms of the molecular orbitals by using the variational method. These differential equations were first derived by Fock[38] based on earlier work by Hartree[33], and are now generally known as the Hartree-Fock equations.

According to the variation principle, if we adjust an approximate many-electron wave function such as eqn. (3.26) to lower the energy, then the accurate solution of the many-electron wave equation will be approached. The best molecular orbitals, therefore, are obtained by varying all the contribut-ing one-electron functions φ_1, φ_2, ... φ_n in the determinant until the energy achieves its minimum value. This will not, of course, give the correct many-electron Ψ for a closed-shell system, but rather the closest possible approach in the form of a single determinant of orbitals. Such orbitals are known as self-consistent, or Hartree-Fock, molecular orbitals. Thus the central mathematical problem is the determination of the orbitals giving a stationary value of $\langle\Psi|\mathcal{H}|\Psi\rangle$, with Ψ being a many electron orbital wave function. In addition, the constraint that the orbitals remain orthonormal, that is, eqn. (3.27) should be satisfied throughout. If this stationary point does in fact correspond to the energy minimum, the corresponding wave function Ψ is the self-consistent solution for the electronic ground state.

Constrained variational problems of this type are handled mathematically by the calculus of variations, using the method of undetermined multipliers. This leads directly to the differential equations

$$[\mathcal{H}^{core} + \Sigma_j(2J_j - K_j)]\varphi_i = \Sigma_j\epsilon_{ij}\varphi_j \qquad i = 1, \ . \ . \ , \ n. \qquad (3.45)$$

These are n one-electron wave equations for the orbitals φ_1, φ_2, . . ., φ_n. The quantity in square brackets is known as the Fock Hamiltonian operator **F**, and the wave equations may be written in the form

$$\mathbf{F}\varphi_i = \Sigma_j\epsilon_{ij}\varphi_j \qquad i = 1, \ . \ . \ . \ , \ n. \qquad (3.46)$$

Here **F** may be considered an effective one-electron Hamiltonian for the electron in the molecular environment, and its various terms have a simple physical interpretation. $\mathcal{H}^{core}$ is the one-electron Hamiltonian for an electron moving in the field of bare nuclei. J_i ($=K_i$) is the potential due to the other electron occupying the same molecular orbital φ_i. Similarly $2J_j$, where j is not equal to i, is the average electrostatic potential of the two electrons in the orbital φ_j. The exchange potential $\mathbf{K}_j$ is somewhat more complicated, but it arises from the effect of the anti-symmetry of the total wave function on the correlation between electrons of parallel spin. The differential equation of eqn. (3.46) differs from ordinary one-electron wave equations in that they each have a whole set of constants ϵ_{ij} on the right-hand side instead of a single eigenvalue, and this arises because the solutions to the set of the wave equations are not unique.

It is desirable to remove this indeterminacy from the problem and to fix the molecular orbitals uniquely. Since the ϵ_{ij} form a Hermitian matrix, there exists a unitary transforma-tion of the form $\varphi_i' = \Sigma_j T_{ij}\varphi_j$, which will bring the matrix of Lagrangian multipliers to diagonal form, that is, all $\epsilon_{ij} = 0$ unless $i = j$. Applying the transformation to the orbitals, the differential equations are brought into the form analogous to a standard eigenvalue equation

$$F\varphi_i = \epsilon_i \varphi_i \qquad i = 1, n. \tag{3.47}$$

These are commonly known as Hartree-Fock equations and state that the best molecular orbitals are all eigenfunctions of the Hartree-Fock Hamiltonian operator $\mathbf{F}$, which in turn is defined in terms of these orbitals through the Coulomb and exchange operators $\mathbf{J}_j$ and $\mathbf{K}_j$. The general expression for the eigenvalues of the Hartree-Fock Hamiltonian operator is

$$\epsilon_i = \mathcal{H}_{ii}^{core} + \Sigma_j(2J_{ij} - K_{ij}) \tag{3.48}$$

which are just those quantities associated with the energy of an electron in orbital φ_i and are thus known as orbital energies.

Although optimum molecular orbitals may be defined as solutions of a set of coupled non-linear differential equations, (for molecular system of any size) direct solution of these equations is impractical and more approximate methods are required. The most rewarding approach to date has been to approximate Hartree-Fock orbitals with linear combination of atomic orbitals (LCAO). **This method has the further advantage that it aids the interpretability of the results, since the nature of chemical problems frequently involves relating properties of molecules to those of the constituent atoms.**

In this approach, each molecular orbital is considered in the form

$$\psi_i = \Sigma_\nu C_{\nu i}\varphi_\nu \qquad (3.49)$$

where the φ_ν are real functions. This form is used within the determinantal wave function, eqn. (3.18). Again, it is required that the orbitals ψ_i form an orthonormal set, and for this to be possible it is necessary that the number of AOs in the basis is greater than or equal to the number of occupied molecular orbitals. The requirement that the molecular orbitals be orthonormal in the LCAO approximation demands that

$$\Sigma_{\mu\nu} c_{\mu i} c_{\nu j} S_{\mu\nu} = \delta_{\mu\nu} \qquad (3.50)$$

where δ_{ij} is the Kronecker delta and $S_{\mu\nu}$ is the overlap integral for atomic functions φ_μ and φ_ν,

$$S_{\mu\nu} = \int\varphi_\mu(1)\varphi_\nu(1)d\tau_1. \qquad (3.51)$$

Molecular orbitals may be obtained to essentially any accuracy desired by appropriate adjustment of the number of basis functions employed in the LCAO expansion. Here we distinguish hree types of basis sets commonly encountered: (1) minimal basis ets, comprising of those atomic orbitals up to and including the rbitals of the valence shell of each atom of the system; (2) xtended basis sets, amounting to a minimal basis set plus any umber of atomic orbitals lying outside the valence shell for ach atom; (3) valence basis sets, comprising just those orbitals of the valence shell of each atom in the system. For example, the valence basis set for the NaH molecule would be 2s, $2p_x$, $2p_y$, and $2p_z$ sodium atomic functions plus the hydrogen 1s function. Adding the 1s sodium orbital brings the valence basis set to a minimal basis set. Adding 3s, 3p, 3d, . . . functions on sodium and 2s, 2p, 3s, . . . functions to hydrogen would give an extended basis set.

At this point it is useful to write down the expression for the electron charge density in the LCAO approximation. The charge density ρ at position **R** is obtained by working out the expectation value of the charge density operator $\varrho(R)$, i.e.,

$$\rho(\mathbf{R}) = \langle\Psi|\varrho(\mathbf{R})|\Psi\rangle = 2\Sigma_i^{occ}\psi_i^*(R)\psi_i(R) \qquad (3.52)$$

Using eqn. (3.49),

$$\rho(\mathbf{R}) = \Sigma_{\mu\nu}P_{\mu\nu}\varphi_\mu(R)\varphi_\nu(R) \qquad (3.53)$$

where

$$P_{\mu\nu} = \Sigma_i^{occ}c^*_{\mu i}c_{\nu i} \qquad (3.54)$$

The integral $\rho(\mathbf{R})$ over all **R** should be equivalent to the total number of electrons in the system, i.e.,

$$2n = \int\rho(\mathbf{R})\,d\mathbf{R} = \Sigma_{\mu\nu}P_{\mu\nu}\int\varphi_\mu(R)\varphi_\nu(\mathbf{R})d\mathbf{R} = \Sigma_{\mu\nu}P_{\mu\nu}S_{\mu\nu}. \qquad (3.55)$$

By means of eqn. (3.55), the electronic charge distribution may be decomposed into contributions associated with the various basis functions of the LCAO expansion. This provides a convenient interpretation of the wave function in terms of constituent atoms and their orbitals. A quantity $P_{\mu\nu}S_{\mu\nu}$ may be considered the electronic population of the atomic overlap distribution $\varphi_\mu\varphi_\nu$, and the diagonal terms such as $P_{\mu\mu}S_{\mu\mu}$ may be associated with the net electronic charges residing in orbital φ_μ. An indication of contributions to chemical binding is given by off-diagonal terms $P_{\mu\nu}S_{\mu\nu}$ with φ_μ and φ_ν centred on different atoms. The matrix element $P_{\mu\nu}$ is thus known as the density matrix.

The total electronic energy can also be written in terms of integrals over atomic orbitals if we substitute the linear expansion of eqn. (3.49) in the molecular orbital integrals. Thus

$$\mathcal{H}_{ii} = \Sigma_{\mu\nu} c^{*}_{\mu i} c_{\nu i} \mathcal{H}_{\mu\nu} \qquad (3.56)$$

where $\mathcal{H}_{\mu\nu}$ is the matrix of the core Hamiltonian with respect to atomic orbitals

$$\mathcal{H}_{\mu\nu} = \int \varphi_\mu(1) \mathcal{H}^{\mathrm{core}} \varphi_\nu(1)\, d\tau_1 \qquad (3.57)$$

similarly we may write

$$J_{ij} = \Sigma_{\mu\lambda\nu\sigma}\ c^{*}_{\mu i} c^{*}_{\lambda j} c_{\nu i} c_{\sigma j} \langle \mu\nu | \lambda\sigma \rangle \qquad (3.58)$$

$$K_{ij} = \Sigma_{\mu\lambda\nu\sigma}\ c^{*}_{\mu i} c^{*}_{\lambda j} c_{\nu i} c_{\sigma j} \langle \mu\lambda | \nu\sigma \rangle \qquad (3.59)$$

where $\langle \mu\nu | \lambda\sigma \rangle$ is the general two-electron interaction integral over atomic orbitals,

$$\langle \mu\nu | \lambda\sigma \rangle = \int\int \varphi_\mu(1) \varphi_\nu(1) 1/r_{12} \varphi_\lambda(2) \varphi_\sigma(2)\, d\tau_1 d\tau_2. \qquad (3.60)$$

This six-dimensional integral gives the Coulomb interaction between two local product densities $\varphi_\mu\varphi_\nu$ and $\varphi_\lambda\varphi_\sigma$.

If these expressions are substituted in eqn. (3.41) for the total electronic energy, we obtain

$$\epsilon = \Sigma_{\mu\nu} P_{\mu\nu} \mathcal{H}_{\mu\nu} + 1/2\ \Sigma_{\mu\nu\lambda\sigma}\ P_{\mu\nu} P_{\lambda\sigma} [<\mu\nu|\lambda\sigma> - 1/2<\mu\lambda|\nu\sigma>. \qquad (3.61)$$

The next important step is to find the optimum values of the coefficients $c_{\mu i}$ leading to a set of SC LCAO or LCAO-SCF MOs. Using the criterion of lowest calculated total energy, such orbitals will be the best for any particular set of basis functions φ_μ. This can be carried out by similar methods to the Hartree-Fock procedure. As a result, we have the equations

$$\Sigma(F_{\mu\nu} - \epsilon_i S_{\mu\nu}) c_{\nu i} = 0 \qquad (3.62)$$

where the elements of the matrix representation of the Hartree-Fock Hamiltonian operator F are

$$F_{\mu\nu} = \mathcal{H}_{\mu\nu} + \Sigma_{\lambda\sigma} P_{\lambda\sigma} <\mu\nu|\lambda\sigma> - 1/2<\mu\lambda|\nu\sigma>. \qquad (3.63)$$

It is seen that the equations of the LCAO SCF MOs, eqn. (3.63) differs from the HF equations in that they are algebraic equations rather than differential equations. They were originally set forth independently by Hall[39], and by Roothaan[35], and are now generally known as the Roothaan equations.

The Roothaan equations for the LCAO-SCF coefficients are cubic since the Fock matrix $F_{\mu\nu}$ is itself a quadratic function of $c_{\mu i}$. This is the mathematical consequence of the fact that the potential experienced by one electron will depend on the number and distribution of other electrons in the system. As a result, the equations have to be solved by an iterative procedure.

Up to this point molecular orbital theory from an <u>ab initio</u> view point has been considered, with the calculation of a wave function involving the evaluation of a number of integrals followed by an algebraic self consistent procedure. A more approximate method which avoids the evaluation of many difficult integrals and which makes some use of experimental data in selecting values of others shall be seen next. Approximate molecular orbital theories are by nature semiempirical, in that one no longer attempts to derive molecular properties directly from the principles of quantum mechanics, but rather seeks to interpret correlations within experimental data.

3.1. SEMIEMPIRICAL METHODS

Computer time and storage requirements of ab initio procedures make computations on large molecules prohibitive. Computation on larger systems are possible but very expensive. The other important factor is the speed of the computer, but since the computational time increases with the fourth power of the number of basis orbitals, if the speed increases by as much as an order of magnitude, this allows only the treatment of a roughly two times larger molecule.

Thus, many molecules of chemical and biological interest are not appropriate for ab initio computations. This makes it necessary that new approximations be incorporated which allow to reduce the computer time by several orders of magnitude. Such approximations concern the neglect of overlap between some AOs, the neglect of certain types of one- and two-electron integrals, approximate evaluation of integrals, and the restriction that only certain groups of electrons (valence or π-type) be treated explicitly. These approximations could easily destroy the reliability of the procedures, since they are, from the numerical point of view, rather crude or sometimes even quite unrealistic.

within the given approximation scheme, however, some quantities appear which can be estimated by making use of experimental data. For example, one-centre integrals in the diagonal elements of the core Hamiltonian can be treated as the atomic ionization potential, which can be taken from experiment instead of being computed analytically (section 3.3.1). In this manner the computational work decreases further, and what is much more important, the incorporation of empirical data may compensate for some of the errors due to the crude approximations. Such procedures are called semiempirical. A semiempirical calculation presupposes the specification of the empirical parameters as well. The parameters can be taken not only from experiments but also from some model calculations on other (e.g. smaller) molecules. It has been observed that some semiempirical methods, despite the inherent rough approxima-tions, work unexpectedly well, providing sometimes finer results than *ab initio* procedures using small basis sets[40].

Semiempirical quantum chemical methods have a wide range of chemical problems. They have been used to provide a potential tool for learning about drug-receptor site interactions, such as the energetics of the reactions, conformations and charge distributions necessary for maximum interaction, molecular sites of reactivity, and mechanism of reactions itself[40].

One of the useful applications of the methods is in the study of activity-structure relationship[30]. It has provided insight into the close relationship between carcinogenicity and electronic structure (electronic energy, bond order, charge density, etc.), and has also made possible prediction of carcinogenicity of aromatic compounds.

In most cases, semiempirical quantum chemical calculations have been used to generate reactivity indices for polycyclic aromatic hydrocarbons (PAH) or related molecules and correlations have been sought between these indices and carcinogenicity.

3.2. THE ZERO DIFFERENTIAL OVERLAP APPROXIMATION (ZDO) [41]

Before reviewing a specific semiempirical method, it is worthwhile to discuss a special feature of many semiempirical procedures.

The most difficult and time-consuming part of LCAO-SCF MO calculations is the evaluation and handling of a large number of electron repulsion integrals. It is known that many of these electron repulsion integrals have values near zero, especially those involving the overlap distribution $\varphi_\mu(1)\varphi_\nu(1)$, with $\mu \neq \nu$. Thus, in developing approximate self-consistent field molecular orbital schemes, a useful approach is the systematic neglect of electron repulsion integrals having uniformly small values. This is effected by means of the zero-differential overlap approximation[41](ZDO).

The scalar product of the atomic orbital $\varphi_\mu(r)$ and $\varphi_\nu(r)$ is defined as their overlap integral:

$$S_{\mu\nu} = \int \varphi_\mu(r)\varphi_\nu(r)d\tau = \int f_{\mu\nu}(r)d\tau \qquad (3.64)$$

where, if $\mu \neq \nu$, the function $f_{\mu\nu}(r)$ equal to the product of basis function $\varphi_\mu(r)\varphi_\nu(r)$ is usually called the differential overlap of the orbitals φ_μ and φ_ν. If φ_μ and φ_ν are centred on far-lying atoms or they are of quite different orientations in space, in their differential overlap $f_{\mu\nu}(r)$ is nearly zero for all r. This feature is often utilized in semiempirical methods by neglecting all or most of the integrals where $f_{\mu\nu} (\mu \neq \nu)$ is present. If all such integrals are neglected, we arrive at the so-called ZDO approximation. The ZDO assumption implies that the overlap integral matrix, S, becomes the unitary matrix, as we obtain:

$$S_{\mu\nu} = \int f_{\mu\nu}(r)\, d\tau = \delta_{\mu\nu} \tag{3.65}$$

where δ_{ij} is Kronecker delta.

The ZDO approximation represents a drastic simplification. Not only the orthogonality of the basis orbitals follows immediately from it, but also most multicentre two-electron integrals vanish:

$$\begin{aligned} \langle\mu\nu|\lambda\sigma\rangle &= \int\varphi_\mu(r_1)\varphi_\nu(r_1)r_{12}^{-1}\varphi_\lambda(r_2)\varphi_\sigma(r_2)\, d\tau_1 d\tau_2 \\ &= \int f_{\mu\nu}(r_1)r_{12}^{-1}f_{\lambda\sigma}(r_2)\, d\tau_1 d\tau_2 \\ &= \delta_{\mu\nu}\delta_{\lambda\sigma}\langle\mu\mu|\nu\nu\rangle. \end{aligned} \tag{3.66}$$

That is, only the one-centre and some two-centre integrals survive. The two-centre integrals $\langle\mu\mu|\nu\nu\rangle$ which are conserved in the ZDO theory describe the Coulomb repulsion between electrons on AOs φ_μ and φ_ν.

If the ZDO approximation is used for all atomic orbital pairs, the Roothaan equations (3.62) for the LCAO coefficients for a closed-shell molecule simplify to

$$\Sigma_\nu F_{\mu\nu}C_{\nu i} = \epsilon_i C_{\mu i} \tag{3.67}$$

where the elements of the Fock matrix $F_{\mu\nu}$ are now given by

$$F_{\mu\mu} = \mathcal{H}_{\mu\nu} - 1/2P_{\mu\mu}\langle\mu\mu|\mu\mu\rangle + \Sigma_\lambda P_{\lambda\lambda}\langle\mu\mu|\lambda\lambda\rangle \tag{3.68}$$

and

$$F_{\mu\nu} = \mathcal{H}_{\mu\nu} - 1/2P_{\mu\nu}\langle\mu\mu|\nu\nu\rangle \qquad \mu \neq \nu. \tag{3.69}$$

3.3. ALL-VALENCE ELECTRON SCHEMES

The other most important feature of semiempirical methods concerns the restriction that only certain groups of electrons (valence or π-type) be treated explicitly.

It has been shown that the role of inner shell electrons in chemistry is limited[40]. This feature can be formulated in several ways. One can say for example that the bonding situation in a molecule can be described merely by appropriate valence shell hybrids without mixing the core orbitals, or a completely filled inner shell is chemically rather inactive (cf. chemical periodicity). The less phenomenological statement is to say that the core AOs are of very low energy relative to the valence orbitals, and usual chemical processes can hardly stimulate a virtual core → valence electronic excitation which would make the role of core orbitals important. There are, exceptions which the behaviour of the core electron reflects chemically important effects (cf. X-ray spectroscopy or ESCA).

A physically acceptable model for most of chemistry arises if one considers only valence electrons explicitly. In such a theory the inner shells are considered as rigid (independent of chemical changes).

Semiempirical methods usually do not include inner shell orbitals, only valence electrons are treated explicitly, (Pseudopotentials are often parameterized semiempirically, but then the valence shell is treated largely at the ab initio level). The so-called all-valence electron scheme emerges if one compresses the core electrons into the nuclei, and works with an effective core charge of atom A:

$$Z_A^c = Z_A - n_A^c \tag{3.70}$$

where Z_A is the atomic number and n_A^c is the number of the inner shell electrons. Accordingly, one deals explicitly only with the valence orbitals, e.g. with 1s AO for hydrogens and with the 2s, $2p_x$, $2p_y$ and $2p_z$ AOs for the first row atoms.

There are several kinds of semiempirical all-valence electron methods which are simple enough to be applied to a wide range of chemical problems without computational effort. One of the most successful all-valence electron methods is referred to as Complete Neglect of Differential Overlap, CNDO[42,43]. Other all-valence electron methods are the intermediate neglect of differential overlap (INDO)[42,43] and the neglect of diatomic differential overlap (NDDO)[42,43] methods.

The CNDO method is the method that was used in this study. As a result the method is discussed in brief in the following section.

3.3.1. COMPLETE NEGLECT OF DIFFERENTIAL OVERLAP (CNDO)[42,43]

The most elementary theory retaining the main features of electron repulsion is the complete neglect of differential overlap method (CNDO) introduced by Pople, Santry, and Segal[42,43]. The method makes a consequent use of the ZDO approximation, complemented by the assumption that the surviving two-electron integrals depend only on the kinds of atoms in question, and are independent of the azimuthal quantum number of the AOs. For the main group elements in which all the valence electrons have the same principal quantum numbers one simply has:

$$<\mu\mu|\nu\nu> = \gamma_{AB} \qquad (3.71)$$

where $\mu \epsilon A$ and $\nu \epsilon B$. The use of eqn. (3.71) is necessary for preserving **the rotation and hybridization invariance of the ZDO, which requires that the results of the calculations should be invariant if one performs a unitary transformation mixing basis orbitals centred on the same atom. The γ_{AB} integrals are** evaluated for s-type Slater orbitals either analytically or by using some approximate formulae.

Using eqn. (3.71), the CNDO expressions for the Fock Hamiltonian matrix elements given in eqns. (3.68) and (3.69) simplify to

$$F_{\mu\mu} = \mathcal{H}_{\mu\mu} - 1/2P_{\mu\mu}\gamma_{AA} + \Sigma_B P_{BB}\gamma_{AB} \qquad \varphi_\mu \text{ on A} \qquad (3.72)$$

and

$$F_{\mu\nu} = \mathcal{H}_{\mu\nu} - 1/2P_{\mu\nu}\gamma_{AB} \qquad \varphi_\mu \text{ on A, } \varphi_\nu \text{ on B.} \qquad (3.73)$$

Here symbol P_{BB} is used for the total electron density associated with atom B,

$$P_{BB} = \Sigma_\lambda^B P_{\lambda\lambda} \qquad (3.74)$$

where the summation is over all atomic orbitals on B.

The next step is to apply a related series of approximations to the matrix elements $\mathcal{H}_{\mu\nu}$ of the core Hamiltonian operator,

$$\mathcal{H} = 1/2\nabla^2 - \Sigma_B V_B \qquad (3.75)$$

where $-V_B$ is the potential due to the nucleus and inner shells of atom B. The diagonal matrix elements $\mathcal{H}_{\mu\mu}$ are conveniently separated in one- and two-centre contributions. If φ_μ is on atom A, we write

$$\mathcal{H}_{\mu\mu} = U_{\mu\mu} - \Sigma_{B(\neq A)} <\mu|V_B|\mu> \tag{3.76}$$

where $U_{\mu\mu}$ is the one-centre term

$$U_{\mu\mu} = <\mu|1/2\nabla^2 - V_A|\mu> \tag{3.77}$$

and is essentially an atomic quantity (the energy of φ_μ in the bare field of the core of its own atom). $U_{\mu\mu}$ is obtained semi-empirically from atomic data. The remaining terms in eqn. (3.76) give the electrostatic interaction of an electron in φ_μ with cores of other atoms B. Its role is to keep valence electrons from penetrating into the inner shell regions. The off-diagonal core matrix elements $\mathcal{H}_{\mu\nu}$ between different AOs φ_μ and φ_ν on the same atom A. This may again be separated into two parts analogously to eqn. (3.76),

$$\mathcal{H}_{\mu\nu} = U_{\mu\nu} - \Sigma_{B(\neq A)} <\mu|V_B|\nu> \quad \varphi_\mu,\varphi_\nu \text{ on } A \tag{3.78}$$

where again $U_{\mu\nu}$ is the one-electron matrix element using only the local core Hamiltonian. If φ_μ, φ_ν are functions of the s, p, d, . . . type, $U_{\mu\nu}$ is zero by symmetry. On the other hand, if a hybrid basis is used, this is no longer so. However, we shall restrict ourselves to s, p, d, . . . sets. The remaining terms in eqn. (3.78) represents the interaction of the distribution $\varphi_\mu\varphi_\nu$ with the cores of other atoms.

In the CNDO method, the two-centre terms $<\mu|V_B|\mu>$ and $<\mu|V_B|\nu>$ in eqns. (3.76) and (3.78) have to be approximated in a manner which is consistent with the way the two-electron integrals are treated. Thus, neglect of monatomic differential overlap $\varphi_\mu\varphi_\nu$ $(\mu \neq \nu)$ on atom A means that $<\mu|V_B|\nu>$ is taken to zero. Further, the invariance conditions also require that the diagonal elements $<\mu|V_B|\mu>$ are the same for all φ_μ on A (for reasons comparable to those already given for replacing $<\mu\mu|\nu\nu>$ by γ_{AB}). Consequently, we shall write

$$<\mu|V_B|\mu> = V_{AB} \tag{3.79}$$

where $-V_{AB}$ is the interaction of any valence electron on atom A with core of atom B. For large internuclear distances R_{AB}, it is approximately equal to R_{AB}^{-1}.

As a result of these approximations, we have

$$\mathcal{H}_{\mu\mu} = U_{\mu\mu} - \Sigma_{B(\neq A)} V_{AB} \qquad \varphi_\mu \text{ on A} \tag{3.80}$$

$$\mathcal{H}_{\mu\nu} = 0 \qquad \varphi_\mu \neq \varphi_\nu, \text{ both on A.} \tag{3.81}$$

To complete the specification of the calculation, we need the off-diagonal core matrix elements $\mathcal{H}_{\mu\nu}$, where φ_μ and φ_ν are on different atoms A and B. Here we do not neglect differential overlap, since these elements take account of the basic chemical bonding capacity of the overlap between the orbitals. However, we may separate the cores of atoms A and B and write

$$\mathcal{H}_{\mu\nu} = <\mu|-1/2\nabla^2 - V_A - V_B|\nu> - \Sigma_{C(\neq A,B)} <\mu|V_C|\nu> \tag{3.82}$$

where the second part gives the interaction of the distribution with the cores of third atom C. These integrals will be neglected, since they are comparable to three-centre, two-electron integrals which have already been omitted. The first part of eqn. (3.82) then depends only on the local environment and is a measure of the possible lowering of energy levels by an electron being in the electrostatic field of two atoms simultaneously. It is commonly referred to as a resonance integral (beta integrals) and denoted by the symbol $\beta_{\mu\nu}$. It is characteristic of the strength of the binding between orbitals φ_μ and φ_ν.

In the CNDO method, the resonance integral $\beta_{\mu\nu}$ are handled in a semiempirical manner. However, this has to be done in a manner which satisfies the required invariance conditions. This will be done by assuming that $\beta_{\mu\nu}$ is proportional to the overlap integral

$$\mathcal{H}_{\mu\nu} = \beta_{\mu\nu} = \beta_{AB}°S_{\mu\nu} \qquad (3.83)$$

This assumption is not reasonable since the bonding capacity of the overlap will increase as the overlap increases. For the calculations to be invariant under transformations of the atomic basis sets, it is required that the proportionality factor between $\mathcal{H}_{\mu\nu}$ and $S_{\mu\nu}$ is the same for all atomic orbitals. This is necessary since $S_{\mu\nu}$ itself transforms correctly. $S_{\mu\nu}$ is evaluated over STOs analytically[44]. (For aforementioned reason, the ZDO approximation must not be applied here. We shall note once again that the overlap integrals $S_{\mu\nu}$ ($\mu \neq \nu$) are absent in the formalism due to the ZDO assumption; they are, however, used to determine the values of the core parameters $\mathcal{H}_{\mu\nu}$.) Using all these approximations, the matrix elements of the Fock Hamiltonian reduce to the following simple form (φ_μ belonging to atom A and φ_ν to atom B):

$$F_{\mu\mu} = U_{\mu\mu} + P_{AA} - 1/2P_{\mu\mu})\gamma_{AA} + \Sigma_{B(\neq A)}\,(P_{BB}\gamma_{AB} - V_{AB}) \qquad (3.84)$$

$$F_{\mu\nu} = \beta_{AB}°S_{\mu\nu} - 1/2P_{\mu\nu}\gamma_{AB} \qquad \mu \neq \nu. \qquad (3.85)$$

The off-diagonal expression, eqn. (3.85), applies even if φ_μ and φ_ν are both on the same atom A, when $S_{\mu\nu} = 0$ and γ_{AB} is replaced by γ_{AA}.

For the expression given in eqn. (3.84) the diagonal matrix element can be rearranged in the form

$$F_{\mu\mu} = U_{\mu\mu} + (P_{AA} -1/2P_{\mu\mu})\gamma_{AA} + \Sigma_{B(\neq A)} [-Q_B\gamma_{AB} + (Z_B\gamma_{AB} -V_{AB}) \quad (3.86)$$

where Q_B is the net charge on atom B,

$$Q_B = Z_A - P_{BB}. \quad (3.87)$$

The two-centre terms in eqn. (3.86) are then easily interpreted. The term $-Q_A\gamma_{AB}$ represents the effect of the potential due to the total charge on atom B (and will vanish if this atom is neutral in the molecular environment). The quantity $Z_B\gamma_{AB} - V_{AB}$ represents the difference between the potentials due to the valence electrons and core of the neutral atom B, and is termed as penetration integral[45].

Once a set of CNDO coefficients $c_{\mu i}$, and a corresponding density matrix $P_{\mu\nu}$ have been obtained, the total energy can be found from

$$\epsilon^{total} = 1/2\Sigma_{\mu\nu} P_{\mu\nu}(\mathcal{H}_{\mu\nu} + F_{\mu\nu}) + \Sigma_{A<B}Z_AZ_BR_{AB}^{-1} \quad (3.88)$$

using appropriate expressions for $\mathcal{H}_{\mu\nu}$ and $F_{\mu\nu}$.

One useful feature of a CNDO calculation is that every term in the total energy expression is associated with one or two atoms, so that an energy breakdown into monatomic and diatomic contributions is possible:

$$\epsilon^{total} = \mathcal{H}_{\mu\nu} + \Sigma_{A<B}\epsilon_{AB} \quad (3.89)$$

The detailed expressions for ϵ_A and ϵ_{AB} are

$$\epsilon_A = \Sigma\mu^A P_{\mu\mu}U_{\mu\mu} + 1/2\Sigma_\mu{}^A\Sigma_\nu{}^A (P_{\mu\mu}P_{\nu\nu} - 1/2P_{\mu\nu}{}^2) \quad (3.90)$$

and

$$\begin{aligned}\epsilon_{AB} = \Sigma_\mu\Sigma_\nu\ (2P_{\mu\nu}\beta_{\mu\nu} - 1/2P_{\mu\nu}{}^2\gamma_{AB}) + Z_AZ_BR_{AB}{}^{-1} \\ - P_{AA}V_{AB} - P_{BB}V_{BA} + P_{AA}P_{BB}\gamma_{AB}). \quad (3.91)\end{aligned}$$

For large intermolecular separations, the potential integrals V_{AB},

V_{BA}, and γ_{AB} all approximate to R_{AB}^{-1} so that the last group of terms in eqn. (3.91) becomes $Q_A Q_B R_{AB}^{-1}$. This shows that the theory takes proper account of the electrostatic interaction between charged atoms in a molecule.

Two methods have been described for choosing parameters. They differ in the way both V_{AB} and U are treated. In the original method[41], which is termed as CNDO/1, V_{AB} and γ_{AB} were both calculated exactly using 2s Slater orbitals; integrals obtained in this way are close to the average values for 2s, $2p_x$, $2p_y$, $2p_z$ orbitals. In the second approach, CNDO/2[46] (the most widely used version of CNDO-type methods), all "neutral penetration integrals" are neglected, which is equivalent to setting $V_{AB} = Z_B\gamma_{AB}$ in eqn. (3.86). Penetration integrals, $Z_B\gamma_{AB} - V_{AB}$, give rise to calculated bonding energies even when the bond orders connecting two atoms are zero. The second approach works considerably better than the first. It should be noted, however, that the method used to obtain $U_{\mu\mu}$, in CNDO/1 may be better than that later used in CNDO/2. In the original method, the authors set

$$U_{\mu\mu} = -I_\mu - (Z_A - 1)\gamma_{AA}. \tag{3.92}$$

In CNDO/2, I_μ is replaced by $1/2(I_\mu + A_\mu)$, so that

$$U_{\mu\mu} = -1/2\ (I_\mu + A_\mu) + (-1/2 - Z_A)\gamma_{AA}. \tag{3.93}$$

The reasoning, clearly somewhat faulty in the case of CNDO/2, is based on the assumption that Koopman's theorem[36] should apply to both the atom and the negative ion as follows:

$$-I_\mu = U_{\mu\mu} + (Z_A - 1)\gamma_{AA} \tag{3.94}$$

$$-A_\mu = U_{\mu\mu} + Z_A\gamma_{AA}. \tag{3.95}$$

Equation (3.93) is then obtained as the average of the two formulas above.

Both methods make the extra approximation

$$\beta_{AB}^{\circ} = 1/2(\beta_A^{\circ} + \beta_B^{\circ}) \qquad (3.96)$$

where the β°'s are the bonding parameters determined semi-empirically to give the optimal fit to the results of <u>ab initio</u> calculations on diatomics. **For CNDO/2 these result in the diagonal element becoming**

$$F_{\mu\mu} = -1/2(I_\mu + A_\mu) - Q_A\gamma_{AA} + (1 - P_{\mu\mu})\gamma_{AA}/2 + \Sigma_{B\neq A}\, Q_B\gamma_{AB}. \qquad (3.97)$$

Equation (3.97) has a nice interpretation. The energy of an electron in atomic orbital φ_μ is equal to the negative of its Mulliken electronegativity[47], $(I + A)/2$, corrected for charge build-up on atom A, $Q_A\gamma_{AA}$, to which it "belongs," corrected in turn for the number of electrons in the φ_μ orbital itself $(1 - P_{\mu\mu})$, and then corrected by the Madelung terms (Coulomb attractions and repulsion) located throughout the molecule, $Q_B\gamma_{AB}$.

The CNDO/2 method is very simple (it requires computations proportional to the third power of the number of basis orbitals), yet accounts for many important tendencies and often leads to qualitative or semiquantitative agreement with experimental results for ground state electronic properties such as geometry and charge distribution.

The higher sensitivity of calculations to a selected system of atomic parameters caused the development of the whole series of different parameterizations which differ in the way the atomic parameters are selected and in the computation of β_{AB}° to correctly reproduce the definite character of the system under study:

binding energy, CNDO/SW[48], geometry, CNDO/BW[49], and electronic spectra, CNDO/S[50].

3.4. π-ELECTRON MODEL

In certain type of molecules, the valence electrons can be grouped by distinguishing between σ- and π-electrons. As a matter of fact, in planar unsaturated compounds, the molecular symmetry enforces that each canonical MO be either symmetric or antisymmetric with respect to reflection in the molecular plane. Consistent with the nomenclature for AOs, the MOs symmetric with respect to the molecular plane are expanded over symmetric AOs, and they are said to be σ-type MOs while antisymmetric MOs are π-type orbitals. The π-type MOs consist of only p_z AOs; the mixing between σ- and π-orbitals is symmetry forbidden. The antisymmetric nature of the π-type MOs means that the π-electrons have a zero probability density in the molecular plane. This implies that they are not so strongly bonded to the molecule as the sigma electrons which have maximum density in the molecular plane. Thus the π-electrons are most easily excitable in an unsaturated planar molecule. In such a case, it is a physically acceptable model to handle only the π-electrons explicitly while the σ-electrons are considered as a rigid core having only an electrostatic influence on the π-electron system.

3.4.1. THE HüCKEL (HMO) METHOD

Although the calculations involved in the CNDO method are much simpler than those in the Roothaan method, they still require the use of computers. In the past, at a time when computers were not generally available, it was necessary to use still simpler and crude treatments if one wished to undertake numerical calculations for typical organic molecules.

For these reasons most of the quantum chemical work in this field has been carried out by the Hückel (HMO) method, introduced originally by Hückel in 1931[51].

The Hückel molecular orbital (HMO) method is certainly the simplest model: only one orbital per atom is considered and no electrostatic interaction is treated explicitly. The Hückel MOs are expanded in terms of the finite set of available π-AOs belonging to the atoms, i.e,

$$\psi_j = \Sigma c_{jA}\varphi_A \qquad (3.98)$$

where φ_A is the π-AO of atom A in the jth and the summation is over all available AOs. In orbital representation, the individual MOs of a molecule are eigenfunctions of the corresponding one-electron HF operator H;

$$H\psi_\mu = \epsilon_\mu\psi_\mu . \qquad (3.99)$$

Instead of solving this equation, the MOs are approximated by using the variational method, with an LCAO trial eigenfunction. Assuming neglect of differential overlap, the corresponding variation equations are:

$$|H_{\mu\nu} - \epsilon\delta_{\mu\nu}| = 0 \qquad (3.100)$$

$$\Sigma_\nu c_{i\nu}\ |H_{\mu\nu} - \epsilon_\mu\delta_{\mu\nu}| = 0 \qquad i = 1, 2, \ldots N \qquad (3.101)$$

where

$$H_{\mu\nu} = \int\varphi_\mu\varphi_\nu d\tau . \qquad (3.102)$$

Instead of calculating the integrals, in HMO theory they are treated as parameters. The diagonal element $H_{\mu\mu}$, called a Coulomb integral is supposed to have a value α_μ characteristic of the AO φ_μ, which is independent of the rest of the molecule, while the off-diagonal element $H_{\mu\nu}$, written as $\beta_{\mu\nu}$ and called a resonance integral, is determined by the nature of the chemical bond between atoms having the orbitals μ and ν. The β parameter of the C-C bond can be used as the energy unit.

It is usually assumed that

$$\beta_{\mu\nu} = 0 \quad \text{unless } \varphi_\mu, \varphi_\nu \text{ are AOs of two atoms that are first neighbours.} \tag{3.103}$$

The eigenvalues of the Hamiltonian are interpreted as π-orbital energies. For the ground state of a molecule possessing 2n electrons, the n lowest energy orbitals are all considered to be occupied by two electrons. Note that the orbitals and orbital energies are completely determined by the set of parameters α and β, and do not depend on the actual number of the π-electrons in the system. This is, however, far from being true in reality and is to be considered as a typical limitation of the Hückel model.

The Hückel total energy is the sum of the energies of all π-electrons; that is

$$E = \Sigma_k n_k \epsilon_k \tag{3.104}$$

where n_k is the occupation number and ϵ_k is the π-energy of the kth MO, and given by $\epsilon_k = \alpha + \lambda_j \beta$. For closed-shell system, the total energy obviously reduces to

$$E = 2\Sigma_k^{occ} \epsilon_k . \tag{3.105}$$

At the present era of modern computers, much more sophisticated quantum chemical calculations can be performed, even for quite large molecules. As a result, the HMO method is not used as it was used before.

4.0. COMPUTATIONAL DETAILS

In this chapter, we consider the calculation of the atomic superdelocalizability, bond reactivity, total bond order, π-bond order, and bond overlap population by means of the methods discussed in the preceding chapter. In addition, a flow chart which shows the computational details involved in calculating the reactivity parameters is given.

4.1. ATOMIC SUPERDELOCALIZABILITY

The HMO definition of atomic superdelocalizability of a particular atom in a given molecule has the following formula[29]

$$S_A = 2\Sigma_{j=1}^{occ} c_{jA}^2/\lambda_j \qquad (4.1)$$

where c_{jA} is the coefficient of A^{th} atomic π-orbital in the j^{th} MO, λ_j is the coefficient in the equation $E_j = \alpha + \lambda_j \beta$, E_j is the energy of the j^{th} MO, and α and β are Coulomb and resonance integrals, respectively. S_A here is given in units of (1/β).
The atomic π superdelocalizability is also defined by[31]

$$S_A = \Sigma_k^{occ} q_A^\pi(k)/\epsilon(k) \qquad (4.2)$$

where $q_A^\pi(k)$ is the π-electron density of atom A in k^{th} MO and $\epsilon(k)$ is the energy of k^{th} MO in units of a.u. or eV. The higher the absolute value of S_A the more likely a reaction will occur at atom A.

In this study we defined atomic superdelocalizability which differs from the atomic superdelocalizabilities given above. The atomic superdelocalizability we calculated using the eigenvectors and eigenvalues obtained from the CNDO/2 method is given by the following expression:

$$S_A' = \Sigma_k^{occ} q_A(k)/\epsilon(k) \tag{4.3}$$

where $q_A(k)$ is the total electron density of atom A in k^th MO and $\epsilon(k)$ is the energy of k^th MO in units of a.u. or eV. The a.u. is the atomic unit of energy (Hartree or Rydberg, Ry).

The difference between the atomic superdelocalizability we defined and the atomic superdelocalizability obtained by eqns. (4.1) and (4.2) is **that in calculating the atomic superdelocalizability we consider the** total electron density by taking into account the σ-electron and π-electron contribution. In eqns. (4.1) and (4.2)[29,31] **only the π-electron density is taken** into account, the role of the σ-electrons is neglected. There are certain molecules for which the σ-framework may play the determining role. As a result we considered the σ-electrons in our calculation.

4.2. BOND REACTIVITY

The π-bond reactivity of a bond between two atoms A and B is obtained by the following expression:[31]

$$R_{AB} = \Sigma_k^{occ} \rho_{AB}^{\pi}(k)/\epsilon(k) \tag{4.4}$$

where $\rho_{AB}(k)$ is the π-bond overlap density between atoms A and B in k^th MO, given by

$$\rho_{AB}^{\pi}(k) = c_{kA}c_{kB} \tag{4.5}$$

and $\epsilon(k)$ is the total energy of k^th MO in units of a.u. or eV. The c_{kA} and c_{kB} in eqn. (4.5) are the coefficients of the π-atomic orbitals (p_z) of A and B, respectively in k^th molecular orbital.

In this work we defined the bond reactivity by the following relation:

$$R_{AB}' = \Sigma_k^{occ}\ \rho_{AB}(k)/\epsilon(k) \qquad (4.6)$$

where $\rho_{AB}(k)$ is the total bond overlap density between atoms A and B in k^{th} MO

$$\rho_{AB}(k) = c_{kA}c_{kB} \qquad (4.7)$$

and $\epsilon(k)$ is the total energy of k^{th} MO and is calculated using the CNDO/2 method. c_{kA} and c_{kB} in eqn. (4.7) are the coefficients of the atomic orbitals (valence orbitals) of A and B, respectively in the k^{th} molecular orbital. The bond reactivity we defined differs from what is used in eqn. (4.4) in that we considered the total bond overlap density in calculating the bond reactivity but eqn. (4.4) takes only π-bond overlap density into account, the σ-bond overlap density is neglected.

4.3. TOTAL BOND ORDER

For a model where the LCAO basis is orthonormalized, i.e. $S_{\mu\nu} = \delta_{\mu\nu}$ and only one orbital per atom is considered, the bond order between atoms A and B is defined by the off-diagonal element of the density matrix[52]:

$$P_{AB} = P_{\mu\nu} \quad ; \ \mu\epsilon A,\ \nu\epsilon B. \qquad (4.8)$$

Whenever one considers more than one AO on one atom, the simple bond order definition, eqn. (4.8) can not be maintained. For such a case, if the AO basis is orthogonal, Wiberg[53] proposed an index characterizing the AB bond order:

$$W_{AB} = \Sigma_{\mu\epsilon A}\Sigma_{\nu\epsilon B}(P_{\mu\nu})^2 \qquad (4.9)$$

where $P_{\mu\nu}$ is the off-diagonal element of the density matrix.

Sophisticated quantum chemical HF methods work within nonorthogonal (overlapping) AO basis sets[40]. In such cases neither the Wiberg indices nor simple bond order, P_{AB}, can serve as an appropriate measure of the chemical bond strength. Mayer[54], proposed an index characterizing the AB bond order to the general multiorbital and nonorthogonal case. According to Mayer's definition the bond order between atoms A and B in a closed-shell molecule can be calculated as:

$$b_{AB} = \Sigma_{\mu\in A}\Sigma_{\nu\in B}(PS)_{\mu\nu}(PS)_{\nu\mu} \tag{4.10}$$

where P is the density matrix and S is the atomic overlap matrix. Mayer's bond order reduces to Wiberg bond order in the special case of $S_{\mu\nu} = \delta_{\mu\nu}$, that is if $S_{\mu\nu} = \delta_{\mu\nu}$, then $b_{AB} = W_{AB}$. In our case, we have used all-valence orbitals for each atom. However, the basis set is orthonormalized, i.e., $S_{\mu\nu} = \delta_{\mu\nu}$. As a result, we obtain the same value for the two bond orders, i.e., for the Wiberg and Mayer bond orders. Hence, we calculated only the Wiberg bond order for our purpose.

4.4. π-BOND ORDER (COULSON)[52]

Whenever one considers one AO on one atom, the simple bond order definition eqn. (4.8) can be used. In Hückel Molecular Orbital (HMO) method, only one orbital per atom is considered. The Hückel MOs are expanded in terms of the finite set of available π-AOs belonging to the atoms, i.e,

$$\psi_j = \Sigma c_{jA}\varphi_A \tag{4.11}$$

where φ_A is the π-AO of atom A and the summation is overall available AOs.

The π-bond order in HMO method between two adjacent atoms A and B (sometimes called the mobile bond order) then is given by:

$$P_{AB}^{\pi} = \Sigma_j^k n_j c_{jA} c_{jB} \quad (4.12)$$

where c_{jA} and c_{jB} are the coefficients of atomic π-orbital (p_z) of atoms A and B, respectively in the j^{th} molecular orbital. The variable n_j (=0,1 or 2) is the number of electrons in j^{th} MO.

In our calculations we have considered all valence orbitals in each atom, e.g. 1s AO for hydrogens and 2s, $2p_x$, $2p_y$ and $2p_z$ AOs for the first row atoms. **Hence in calculating the π-bond order, P_{AB}^{π}, we have identified the atomic orbitals with π-symmetry for each carbon atom and used them for calculating the π-bond order.**

4.5. BOND OVERLAP POPULATION

We used the following expression to calculate the overlap population[40] between orbitals μ and ν:

$$P_{\mu\nu}^{ov} = P_{\mu\nu} S_{\mu\nu} \quad (4.13)$$

where $P_{\mu\nu}$ is the off-diagonal density matrix and $S_{\mu\nu}$ is the off-diagonal elements of the overlap matrix. The calculated bond overlap population characterize the binding between the two orbitals μ and ν.

The SAMO method mentioned above (Chapter 1) uses the Mulliken bond overlap population[18] to calculate the overlap population between orbitals μ and ν. Here the molecular orbitals ψ_i are expressed as linear combinations of the members of hybrid sets φ_μ.

$$\psi_i = \Sigma \mu C_{i\mu} \varphi_\mu \quad (4.14)$$

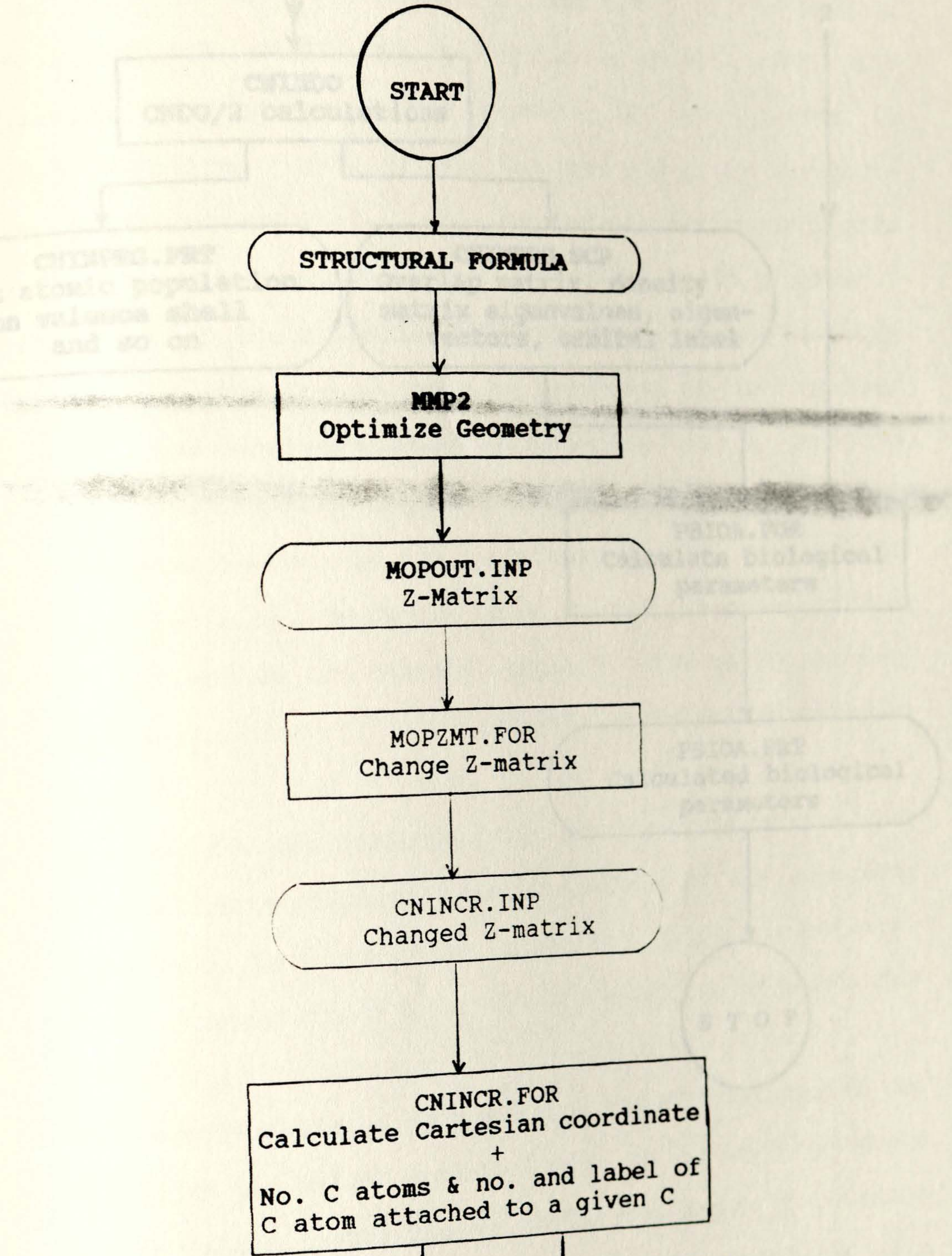
The Mulliken bond overlap population is then given by the following relation:

$$\beta_{\mu\nu} = 4\Sigma_i C_{i\mu} C_{i\nu} S_{\mu\nu} \quad (4.15)$$

where the sum is over the occupied orbitals and $S_{\mu\nu}$ is an element of the overlap $S_{\mu\nu} = \langle\varphi_\mu|\varphi_\nu\rangle$.

In order to calculate the reactivity parameters defined above for a given aromatic hydrocarbon we need to have eigenvalues, eigenvectors, orbital label, S-matrix, and density matrix of the atoms which compose the system. We should also know the number of carbon atoms in the compound, and the number of carbon atoms attached to a given carbon atom in the compound. All these quantities are determined using computer programs.

A flow chart of the programs is given below. The flow chart is a summary of the computations involved in this work. It shows the hierarchy of the programs and flow of data among the programs.


START
STRUCTURAL FORMULA
MMP2
Optimize Geometry
MOPOUT.INP
Z-Matrix
MOPZMT.FOR
Change Z-matrix
CNINCR.INP
Changed Z-matrix
CNINCR.FOR
Calculate Cartesian coordinate
+
No. C atoms & no. and label of C atom attached to a given C
CNINCR.OTP
Calculated Cartesian Coordinate
CNINCR.CNN
No. C atoms & no. and label of C atom attached to a given C

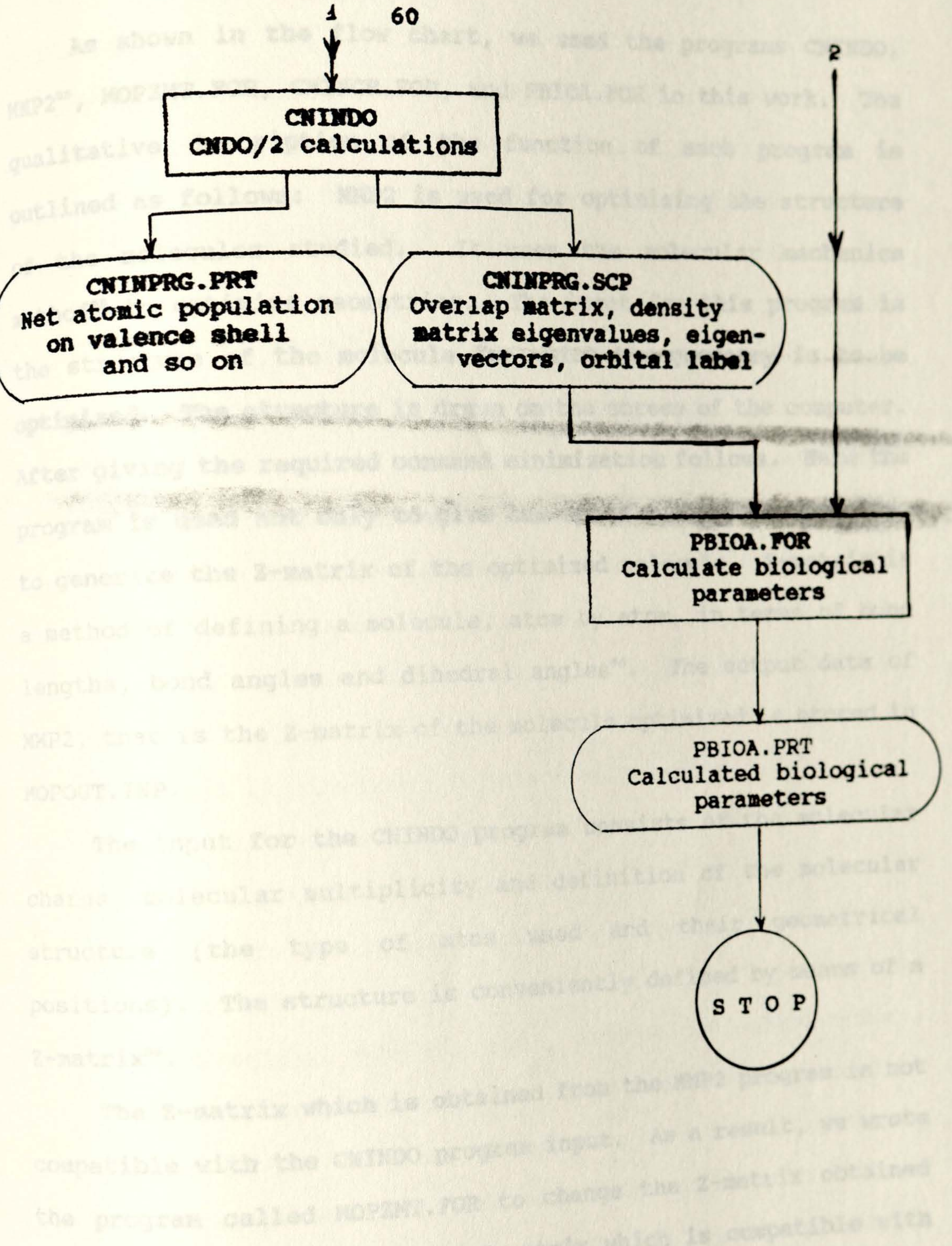


Fig. 7. Flow chart for calculating the reactivity parameters.

As shown in the flow chart, we used the programs CNINDO, MMP2[55], MOPZMT.FOR, CNINCR.FOR, and PBIOA.FOR in this work. The qualitative description of the function of each program is outlined as follows: MMP2 is used for optimizing the structure of the molecules studied. It uses the molecular mechanics method[55] to optimize geometries. The input for this program is the structure of the molecule for which the geometry is to be optimized. The structure is drawn on the screen of the computer. After giving the required command minimization follows. Here the program is used not only to give the optimum structure but also to generate the Z-matrix of the optimized molecule. Z-matrix is a method of defining a molecule, atom by atom, in terms of bond lengths, bond angles and dihedral angles[56]. The output data of MMP2, that is the Z-matrix of the molecule optimized is stored in MOPOUT.INP.

The input for the CNINDO program consists of the molecular charge, molecular multiplicity and definition of the molecular structure (the type of atom used and their geometrical positions). The structure is conveniently defined by means of a Z-matrix[56].

The Z-matrix which is obtained from the MMP2 program is not compatible with the CNINDO program input. As a result, we wrote the program called MOPZMT.FOR to change the Z-matrix obtained from MMP2 program to another Z-matrix which is compatible with the CNINDO program input. The input data for MOPZMT.FOR is obtained from MOPOUT.INP. The output of MOPZMT.FOR, the Z-matrix which is compatible with the CNINDO program, is stored in

CNINCR.INP.

The program CNINCR.FOR calculates the Cartesian coordinates of each atom in a molecule (geometrical position) from the Z-matrix. The program CNINCR.FOR obtains the required input data with the proper formats for identification and comments; the number of atoms, molecular charge, multiplicity, bond angles, bond lengths, and dihedral angles (Z-matrix) from the file CNINCR.INP. The calculated Cartesian coordinates are stored in CNINCR.OUT. The original CNINCR.FOR program which was written to calculate the Cartesian coordinate was modified to determine the number of carbon atoms in a given molecule, and assign the number and label of carbon atoms attached to a given carbon atom in a molecule being studied. These results are stored in CNINCR.CNN.

The CNINDO program is used for the CNDO/2 calculations. It calculates the eigenvalues, eigenvectors, density matrix, overlap matrix, net atomic population on valence shell and so on. The required input data is obtained from CNINCR.OUT, i.e., the Cartesian coordinates of each atom. The output of the CNINDO is stored in CNINCR.PRT. The CNINDO is modified to store the calculated S-matrix, orbital label, eigenvalues, eigenvectors, and density matrix separately in CNINPRG.SCP. This values are arranged in such a way that they can be used to calculate the reactivity parameters. This program is also modified to identify the orbitals with π-symmetry for calculating the π-bond order.

The program PBIOA.FOR which we wrote is used to calculate the reactivity parameters, the atomic superdelocalizability, bond reactivity, total bond order of Wiberg, π-bond order, and overlap population. It takes input data from CNINCR.CNN and CNINPRG.SCP and stores the output data in PBIOA.PRT.

5.0. RESULTS AND DISCUSSION

As mentioned in the previous chapter, the reactivity parameters, the atomic superdelocalizability, and the bond reactivity are defined to include all valence electrons. That is, in calculating the atomic SD and bond reactivity both the σ- and π-electrons and bonds, respectively, were considered. The π-bond order is calculated using atomic orbitals with π-symmetry. In the HMO method the atomic superdelocalizability and the bond reactivity are calculated by considering only the π-electrons and π-bond, respectively. The π-bond order, in HMO method, is calculated using p_z orbitals. Hence our first task was to establish a correlation between our reactivity parameters and those based on the HMO method. To establish the correlation test calculations were performed.

5.1. TEST CALCULATIONS

As a test of the calculated atomic superdelocalizability, S_A', we calculated and compared the atomic superdelocalizabilities of benzene, nitrobenzene, phenol, aniline, naphthalene and phenanthrene.

First we considered benzene, phenol, aniline and nitrobenzene. In benzene all the carbon atoms are equivalent, $S_A' = -5.137$. If it has a substituent, the remaining five carbon atoms are no longer equal. With respect to the substituent, two ortho, two meta, and one para positions are distinguishable. The -OH and $-NH_2$ groups in phenol and aniline, respectively, are electron donating in nature and increase the electron density of the benzene ring[57]. This in turn, causes the redistribution of the electron density in the ring which creates a surplus of electron density in the ortho and para positions. As a result, these positions are more reactive towards electrophilic attack than the meta position[57].

The absolute value of the calculated atomic superdelocalizability (eqn. 4.1) increases with increasing electron density if the orbital energy is constant or the difference in the orbital energy is small. So surplus electron density in the ortho and para positions (in phenol and aniline) means these positions have large negative atomic superdelocalizabilities.

The atomic superdelocalizabilities of the two ortho, two meta and one para positions of phenol are -5.35, -5.06, and -5.30, respectively and of aniline are -5.55, -5.14, and -5.47, respectively. This result shows that the ortho and para positions of phenol and aniline have more electron density than their meta positions. Hence, as stated above the ortho and para positions are more reactive than their meta positions. Whereas in the case of nitrobenzene, the nitro group being electron accepting, pulls the electron cloud from the benzene ring in its direction. As a result, the electron density in the benzene ring is lower somewhat, especially in the ortho and para positions and the reactivity with an electrophile decreases at the ortho and para positions. The atomic superdelocalizability of the two ortho, two meta, and one para positions of nitrobenzene are -4.83, -4.88, and 4.81, respectively. Note that the absolute value of the atomic superdelocalizabilities of the ortho and para expected that the bond length decreases with increasing bond order. In Fig. 7 are plotted bond lengths of ethane (1.54Å), ethylene (1.34Å), acetylene (1.2Å), butadiene (outer bond 1.35Å and inner bond 1.46Å) and benzene (1.40Å) with the corresponding calculated bond order P_{AB}. As expected the bond length decreases with increasing calculated bond order (Fig. 7). The correlation coefficient relating the experimental bond length with the calculated bond order is -0.988. The negative sign in the correlation coefficient indicates that the trend the higher the total bond order the lower the bond length holds.

Since bond reactivity is a measure of bond overlap density (eqn. 4.6), its absolute value increases with increasing bond overlap density if the orbital energy is constant or the difference in the orbital energy is small. The bond reactivity of ethane (single bond), ethene (double bond) and acetylene (triple bond) is -0.70, -1.12, and -1.77, respectively. This shows that as the bond order increases the absolute value of the bond reactivity increases. This is what is observed in Fig. 8 which is a plot of the calculated bond reactivity versus calculated total bond order. Fig. 9, the plot of bond reactivity versus calculated π-bond order, again shows that absolute value of the bond reactivity increases with increasing π-bond order. The correlation coefficients relating the bond reactivity with the calculated total bond order and π-bond order are 0.988 and 0.999, respectively. Note that the coefficient relating the bond reactivity with π-bond order is greater than the one relating the bond reactivity with total bond order. This shows indirectly that the role played by π-bond is more than the σ-bond in bond reactivity.

Figure 10 shows the relation between the two π-bond orders, i.e. the π-bond order we calculated and the one calculated by HMO method. The two π-bond orders are in strong agreement (correlation coefficient = 0.985).

The plot of the calculated overlap population versus the calculated total bond order and π-bond orders are shown in Fig. 11 and 12, respectively. Generally an increase in the calculated bond orders result in increase in the calculated overlap population. The correlation coefficient relating the calculated bond overlap population to the calculated total bond order and, to π-bond order are 0.906 and 0.998, respectively. This relationship shows the contribution of the π-bond in the calculated overlap population to be higher than the σ-bond.

Table 1. Relative reactivities (nitration)[a], the calculated atomic superdelocalizabilities, S_A' and the atomic superdelocaliabilities calculated by HMO method, S_A in phenanthrene.

Position of reaction	log $k_A/k_{\alpha\text{-napth}}$[a]	S_A' (unit of 1/a.u.)	S_A (unit of 1/β)[a]
9	0.02	-5.329	0.997
1	-0.11	-5.309	0.977
3	-0.19	-5.205	0.893
2	-0.71	-5.191	0.860
4	-0.77	-5.294	0.939

[a] from Ref. 30

Table 2. Calculated bond order b_{AB}, π-bond orders (P_{AB}^{π}), overlap population (P_{AB}^{ov}), and bond reactivity (R_{AB}).

Compound	Bond	b_{AB}	P_{AB}^{π}		P_{AB}^{ov}	R_{AB}
			CNDO/2	HMO[c]		
Ethane	c – c	1.06	0.18		0.74	-0.70
Ethylene	c = c	2.06	1.00	1.00	1.13	-1.12
Butadiene	c – c c = c	1.06 1.96	0.29 0.96	0.45 0.89	0.78 1.11	-0.74 -1.18
Benzene	c:::c	1.44	0.67	0.67	0.95	-0.99

[c] from Ref. 30

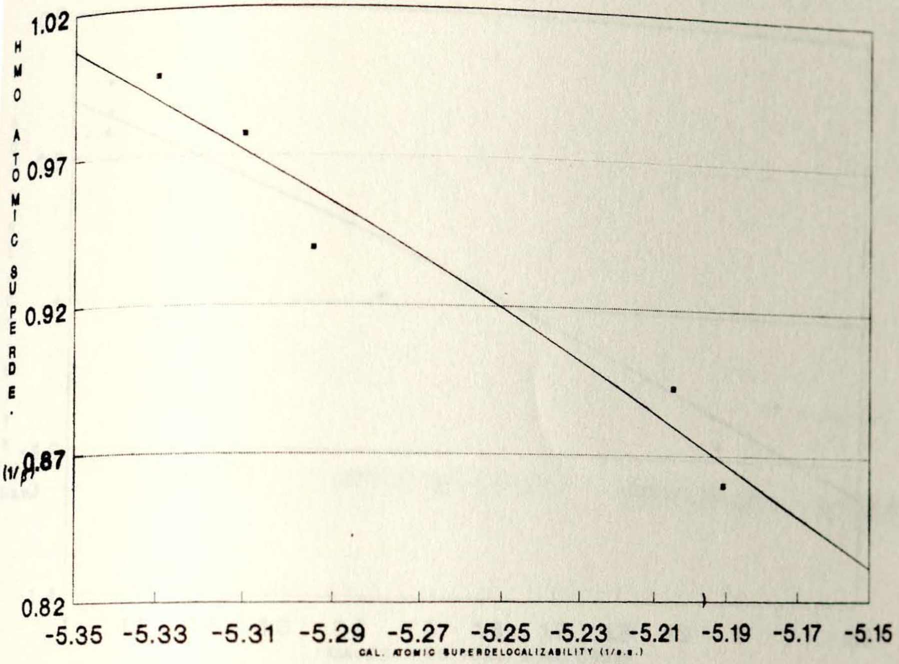


Fig. 8. Calculated atomic superdelocalizabilities versus HMO atomic superdelocalizabilities in phenanthrene.

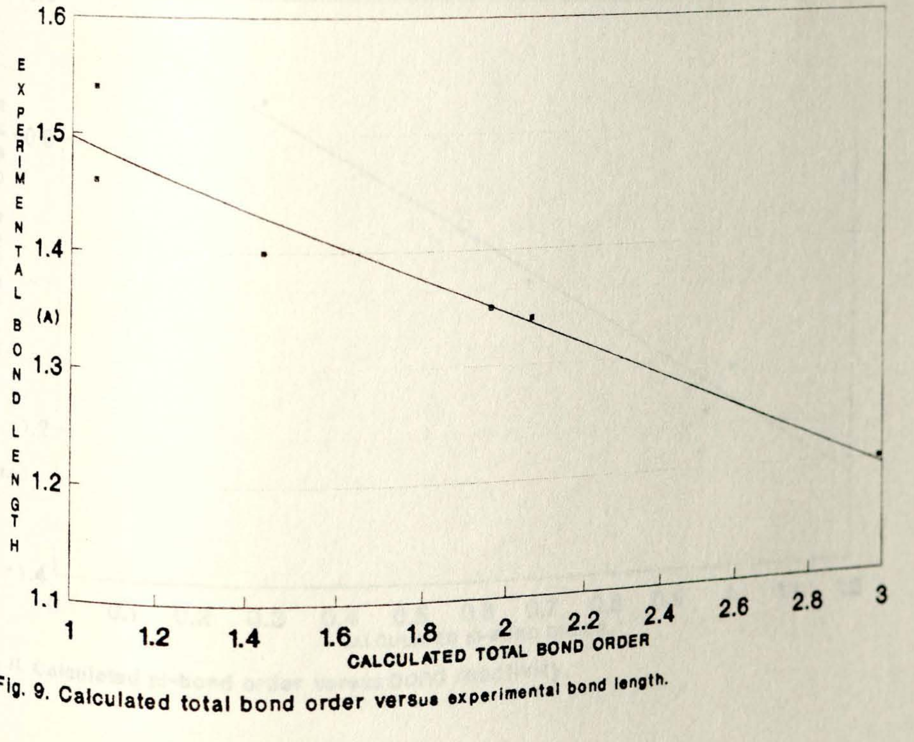


Fig. 9. Calculated total bond order versus experimental bond length.

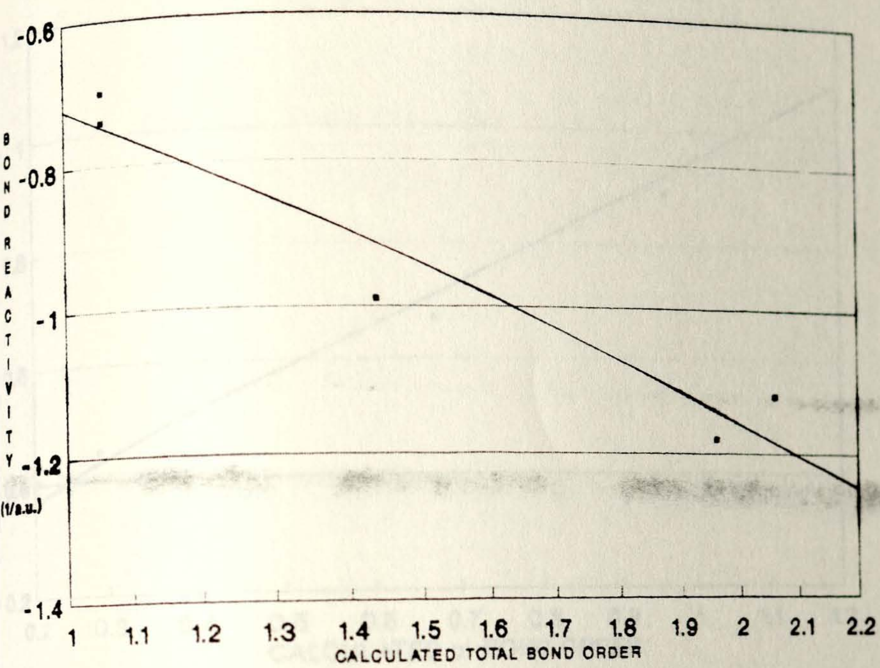
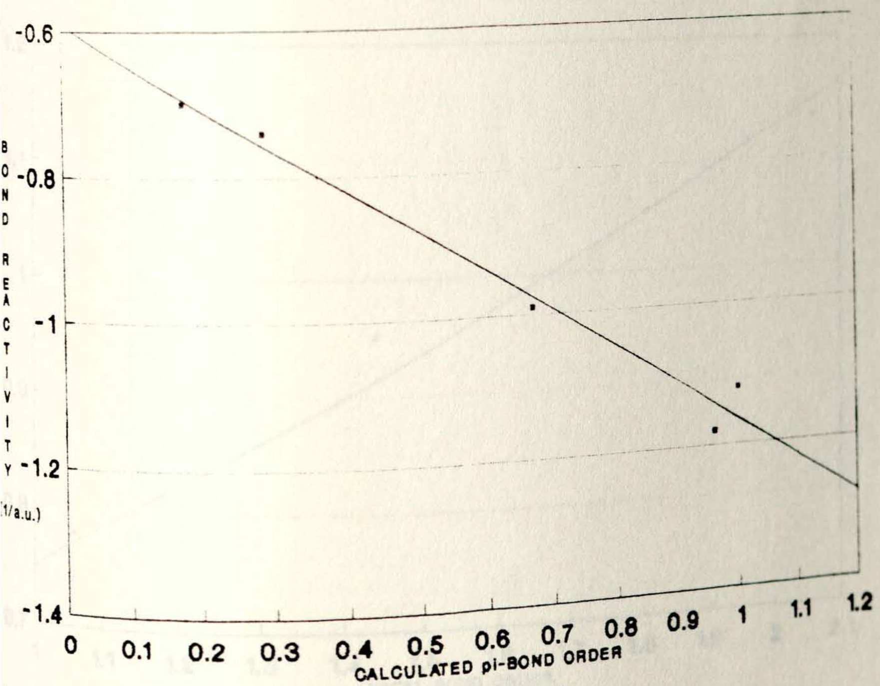
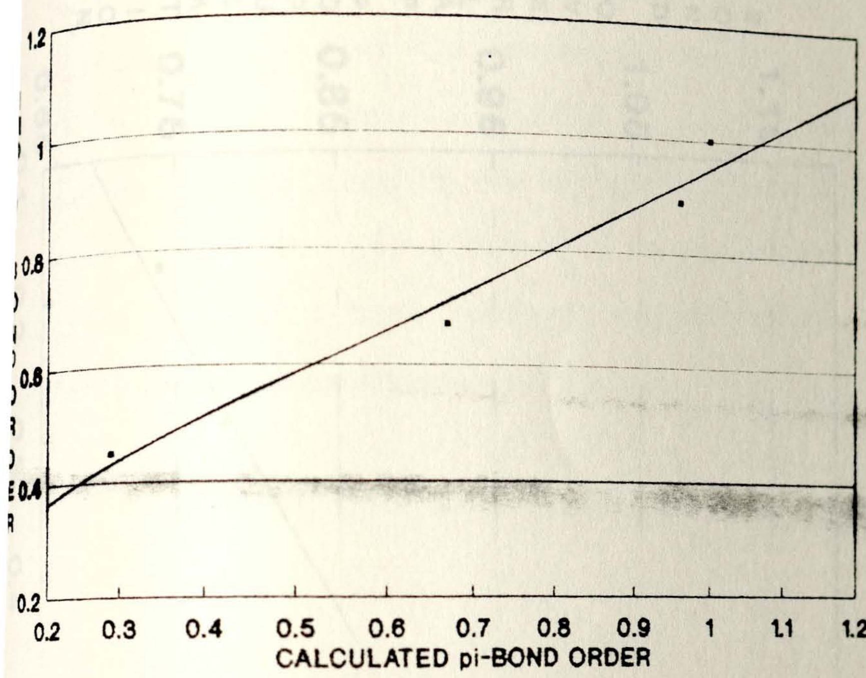


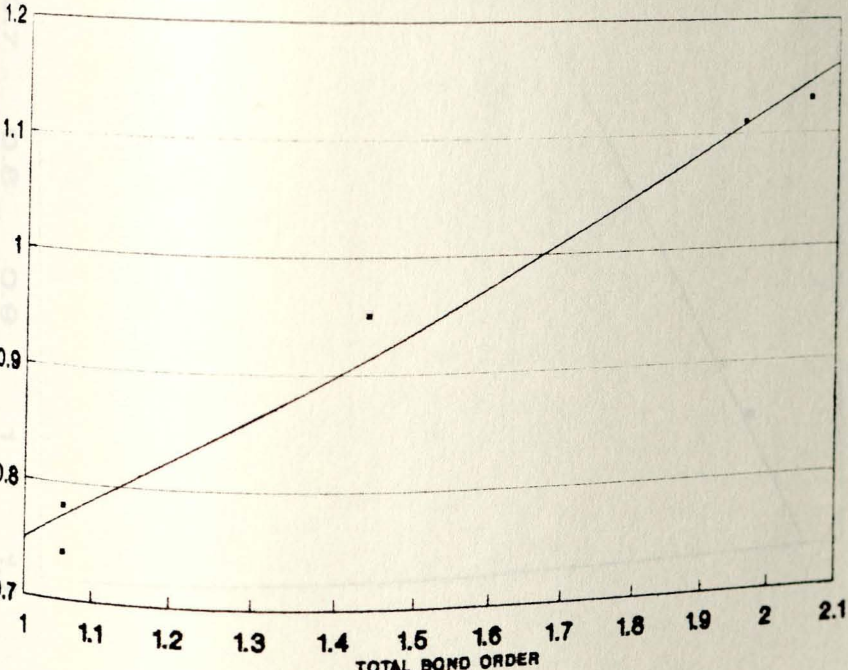
Fig. 10. Calculated total bond order versus bond reactivity.



ig. 11. Calculated pi-bond order versus bond reactivity.



ig. 12. Calculated pi-bond order versus HMO pi-bond order.



3. Calculated total bond order versus bond overlap population.

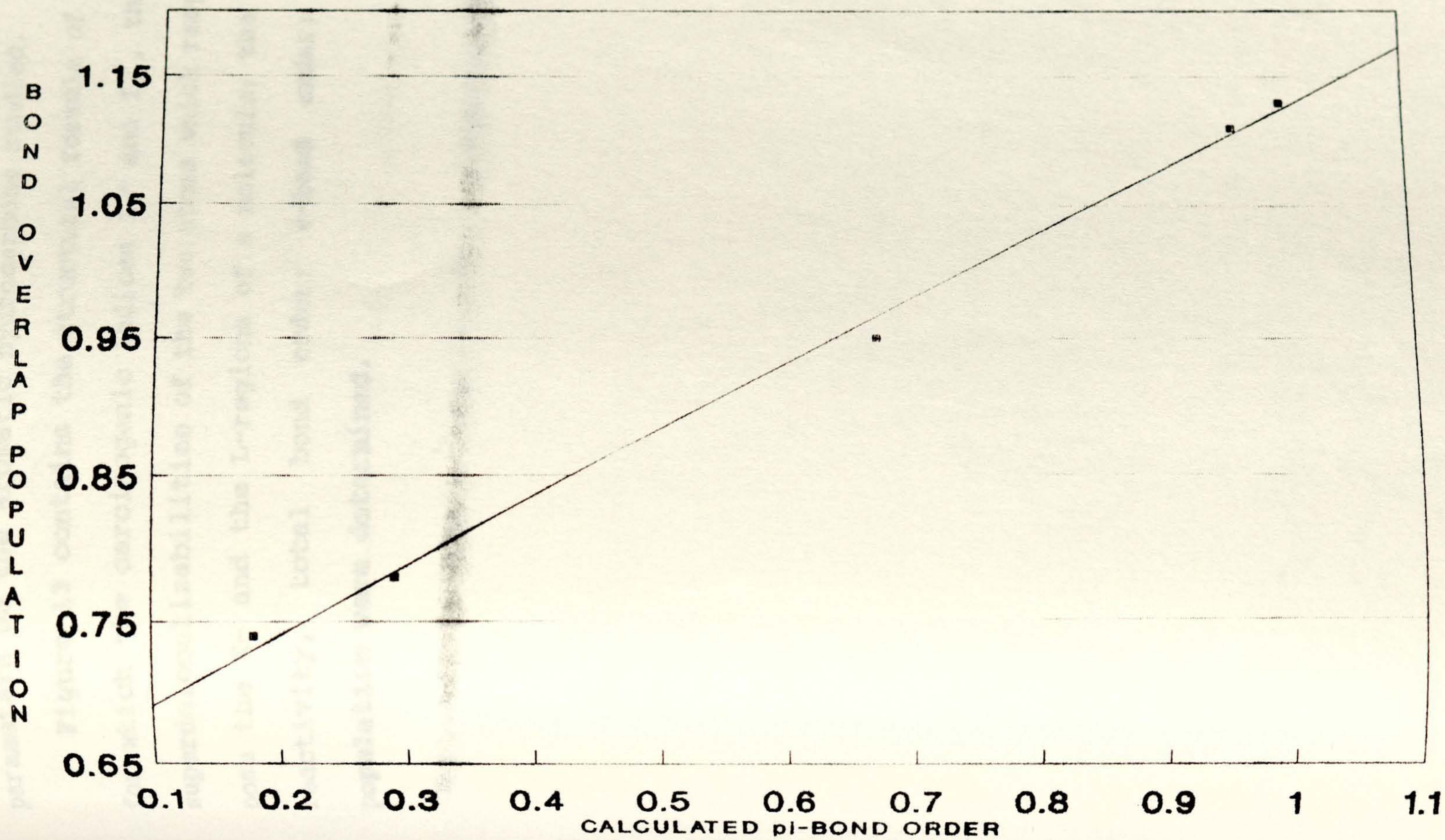


Fig. 14. Calculated pi-bond order versus bond overlap population.

Next we see the result of the calculation of the reactivity parameters on the aromatic hydrocarbons studied.

Figure 13 contains the structural formula of the molecules for which the carcinogenic indices I_K' and I_L', the sum of the superdelocalizabilities of the two atoms which respectively compose the K- and the L-regions of a molecule; the K-region bond reactivity, total bond order; π-bond order; and overlap population were determined.

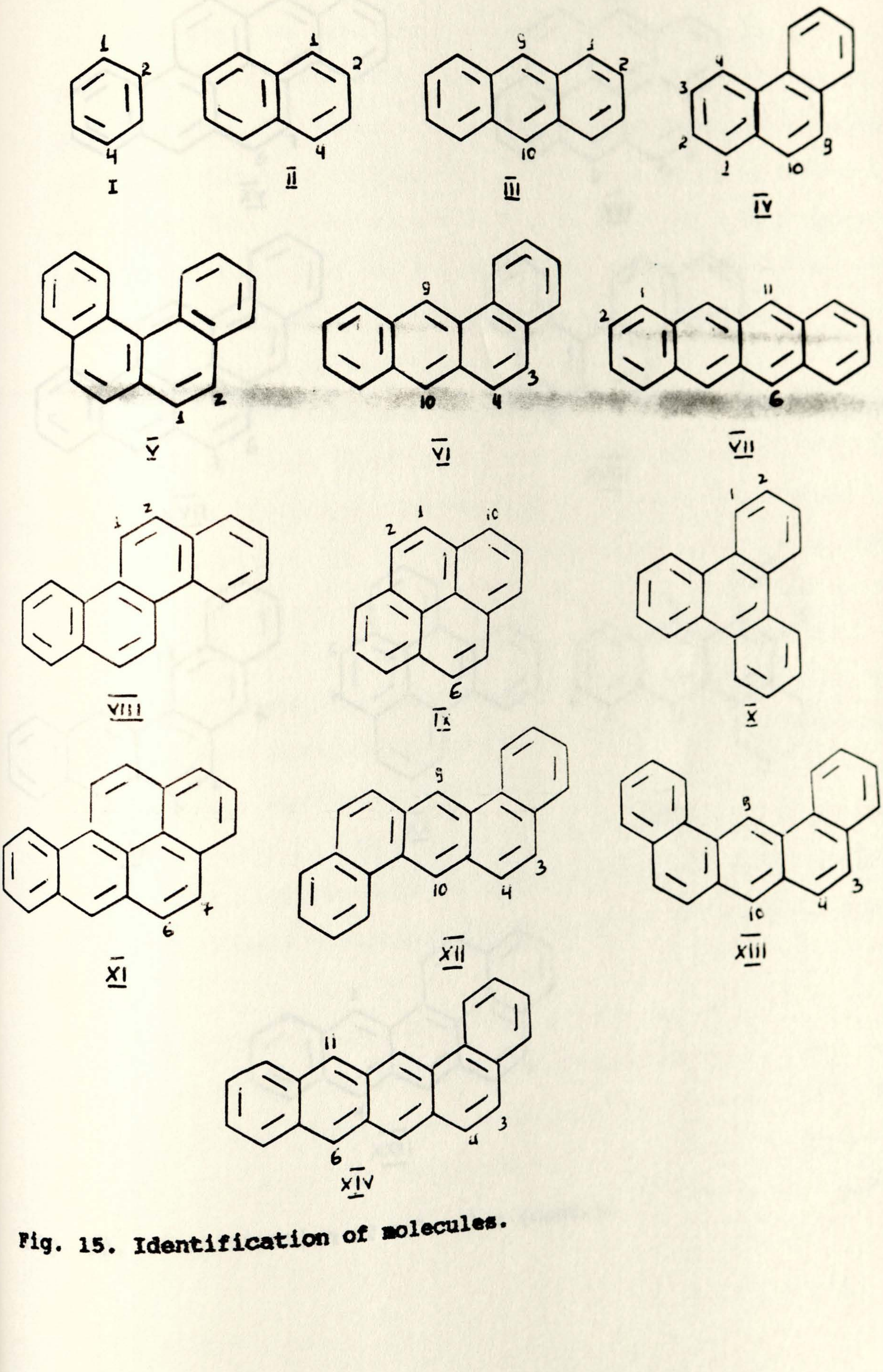
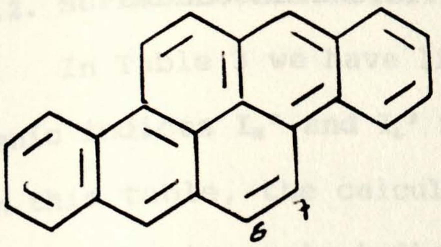
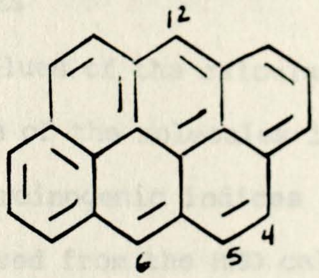


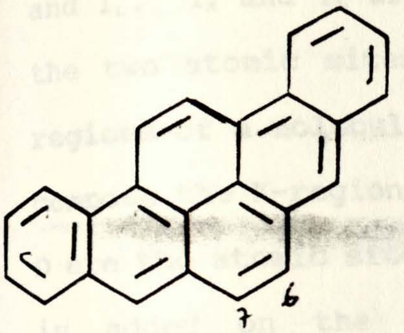
Fig. 15. Identification of molecules.



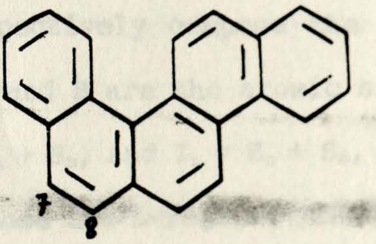
XV



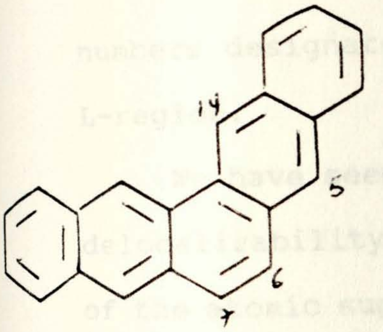
XVI



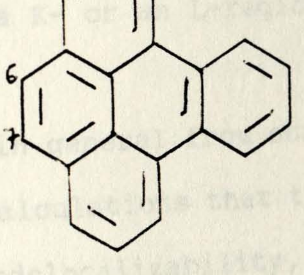
XVII



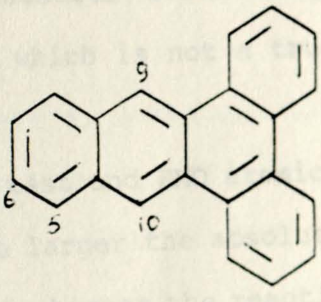
XVIII



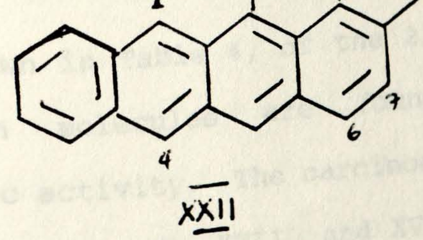
XIX



XX



XXI



XXII

Fig. 15. Identification of molecules (cont.).

5.2. SUPERDELOCALIZABILITY INDICES

In Table 3 we have listed values of the calculated carcinogenic indices I_K' and I_L' for each of the molecules in Figure 13. In this table, the calculated carcinogenic indices are compared to the carcinogenic indices derived from the HMO calculations I_K and I_L. I_K and I_L are the sum of the superdelocalizabilities of the two atomic sites which respectively compose the K- and L-regions of a molecule. When A and B are the atomic sites which compose the K-region then $I_K = S_A + S_B$, and $I_L = S_C + S_D$, when C and D are the atomic sites that compose the L-region. The prime sign is added on the calculated superdelocalizability index to distinguish it from the superdelocalizability index obtained by the HMO method. In Table 3 a parentheses around a pair of numbers designate a K- or an L-region which is not a true K- or L-region.

We have seen in general from our test and HMO atomic superdelocalizability calculations that the larger the absolute value of the atomic superdelocalizability, the higher the reactivity of the given site. From this it is clear that the relation, the larger the absolute value of the superdelocalizability indices the more likely a reaction will take place at the given bond, holds.

As shown in Table 4, of the 22 molecules we have studied, only seven molecules are found experimentally to show carcinogenic activity. The carcinogenic molecules are molecule, V, XI, XII, XIII, XV, XVII, and XVIII (Fig. 13). Molecules V, XI, XV, XVII, and XVIII possess only the K-region whereas molecules XII and XIII have both K- and L-regions.

It is recognized that the K-region reactivity is like an "on" switch; if the K-region is more reactive to a certain limit, carcinogenicity is favoured[16]. The L-region behaves like an "off" switch; it prevents carcinogenicity if it becomes too reactive, regardless of the K-region index. Having this fact i.e., the critical cancer-inducing reaction takes place at a bond of the K-region type, and reactions at L-region bypass the development of cancer, a criterion for carcinogenicity was stated using the HMO superdelocalizability indices. The criterion for an aromatic hydrocarbon to be carcinogenic using the HMO superdelocalizability indices is stated as follows:

(i) it must possess a K-region which is more reactive than 2.05(1/ß);

(ii) if the molecule also possesses an L-region, this region must be less reactive than 2.30(1/ß),

If we apply the criterion stated using the HMO superdelocalizability indices to the molecules studied, we will find compounds IX, XI, XIII, XVI, and XVII to be carcinogenic. The exception to the criterion specified using the HMO superdelocalizability are molecules V, IX, XVI and XVIII. They are also exceptions to the criteria based upon bond and carbon localization energies. Molecule V and XVIII which show carcinogenic activity are predicted, according to this criterion, to be non-carcinogenic. Molecules IX and XVI which are non-carcinogenic are predicted to show carcinogenic activities. The exceptional behaviour of molecule XVI has been explained by its possession of a highly reactive, extended, (pseudo) L-region, atoms 6 and 12[17] (Fig. 13). The non-carcinogenicity of molecule IX is explained by its close relation to the non-carcinogenic molecule XVI. In molecule IX, atoms 6 and 10 constitute a highly reactive, extended, (pseudo) L-region in the same fashion as atom 6 and 12 of molecule XVI[17] (Fig. 13). Molecules V and XVIII are exceptions to the criterion stated by the HMO superdelocalizability indices and this is assumed to be attributed to the non-planarity of the molecules[17]. The HMO method permits one to perform calculations on planar conjugated molecules only. By means of the superdelocalizabilities we calculated, a criterion for chemical carcinogenicity may be stated in the following way:

(i) The molecule must possess a K-region superdelocalizability index $\leq -10.70(1/a.u.)$;

(ii) if the molecule also possesses an L-region, this region must have superdelocalizability index $\geq -10.98(1/a.u.)$.

Again we formulated this criterion based on the fact that the critical cancer-inducing reaction takes place at a bond of the K-region type, and reactions at L-region bypass the development of cancer.

Molecules V and XVI are exceptions to the criteria specified using the superdelocalizability indices we calculated. They are also exceptions to the criteria based upon bond and carbon localization energies, and the superdelocalizability indices obtained by HMO method.

The exception of molecule V to the criteria stated using HMO superdelocalizability indices was assumed to be attributed to the non-planarity of the molecule[17]. In our calculations there is no restriction in geometry. That is, we can study planar or non-planar molecules. Molecule V which is non-planar was an exception again to the criteria stated using our super-delocalizability indices which is obtained using CNDO/2 method. Hence there should be a factor other than the non-planarity for the molecule to be exception to both criteria.

The exceptional behaviour of molecule XVI, using the HMO SD indices, has been explained by its possession of a highly reactive, extended, (pseudo) L-region, atoms 6 and 12[17]. In our calculation, molecule XVI is found again to possess a highly reactive, extended, (pseudo) L-region, atoms 6 (S_6 = -5.603) and 12 (S_{12} = -5.604).

Molecule XVIII, which is an exception to the criterion stated using the HMO superdelocalizability indices, is not an exception to the criterion stated by our superdelocalizability indices. In this study, the geometry of all molecules is optimized before the actual computation of the reactivity parameters (section 3.0). With CNDO method, there is no restriction as to geometry of the molecules. That is we can study, as mentioned above, planar or non-planar molecules. For this reason the non-planar molecule XVIII is not an exception to the criterion stated using the calculated superdelocalizability indices. This observation supports the reason given for the exceptional behaviour of molecule XVIII to the criterion stated by the HMO superdeclocalizability indices.

Molecule IX which is an exception to the criterion stated using the HMO superdelocalizability indices, is not an exception to our criteria.

The relationship between the K- and L-region superdelocalizability indices obtained by our procedure and the HMO method are shown in Figures 14 and 15, respectively. The Figures show that the higher the absolute value of our superdelocalizability indices the higher the HMO superdelocalizability indices. The correlation coefficient of the superdelocalizabilities we calculated and the ones obtained by HMO method in the K-region and L-region are 0.89 and 0.90, respectively.

Table 3. SD indices of the K- and L-regions of aromatic hydrocarbons calculated by CNDO/2 (I_K' and I_L') in units of 1/a.u.

COMPD	K-REGION	I_K'	I_K	L-REGION	I_L'	I_L	CAR CIN.[f]	CLASS
I	(1-2)	-10.27	1.67	(1-4)	-10.27	1.67	-	C
II	(1-2)	-10.49	1.87	(1-4)	-10.60	1.99	-	C
III	(1-2)	-10.55	1.99	9-10	-11.04	2.63	-	E
IV	9-10	-10.66	1.99				-	C
V	1-2	-10.69	1.97				+	F
VI	3-4	-10.70	2.08	9-10	-11.01	2.44	-	D
VII	(1-2)	-10.58	2.08	6-11	-11.17	3.01	-	D
VIII	1-2	-10.68	2.00				-	C
IX	1-2	-10.66	2.05				-	C
X	1-2	-10.32	1.80				-	C
XI	6-7	-10.70	2.14				++++	A
XII	3-4	-10.70	2.06	9-10	-10.98	2.29	++	A
XIII	3-4	-10.70	2.05	9-10	-10.98	2.29	+	A
XIV	3-4	-10.72	2.14	6-11	-11.15	2.86	-	D
XV	6-7	-10.73	2.17				++++	B
XVI	4-5	-10.70	2.16				-	X
XVII	6-7	-10.74	-				++++	A
XVIII	7-8	-10.70	1.98				+	F
XIX	6-7	-10.74	2.15	5-14	-11.07	2.41	-	D
XX	(5-6)	-10.69	2.03				-	C
XXI	6-7	-10.54	1.93	9-10	-11.00	2.31	-	E
XXII	6-7	-10.73	2.20	1-4	-11.12	2.72	-	D

[f] From Ref. 17

Table 4. Classification codes.

Class	Explanation
A	Molecule is active because both K- and L-regions obey the limiting conditions for carcinogenic potency.
B	Molecule is active because the K-region obeys the limiting condition for carcinogenic potency.
C	Molecule is inactive because a K-region suitable for carcinogenesis is absent.
D	Molecule is inactive because an L-region suitable for carcinogenesis is absent.
E	Molecule is inactive because both K- and L-regions are unsuitable for carcinogenic activity.
F	Exceptions to the SD criteria stated using our SD and that of HMO SD.
X	The only exception to the SD criteria not included in Class F.

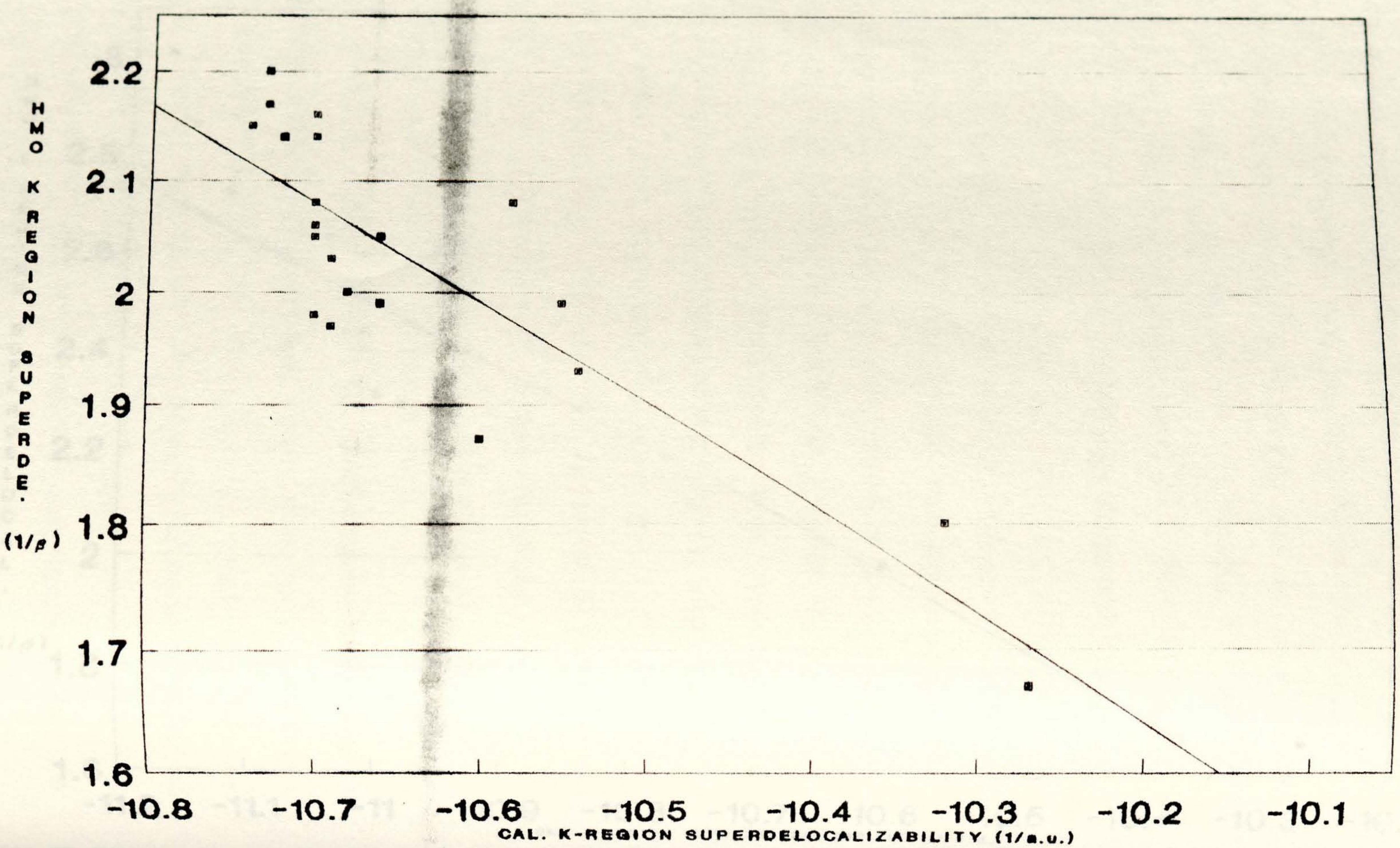


Fig. 16. CAL. K-region superdelocalizability index versus HMO K-region superdelocalizability index.

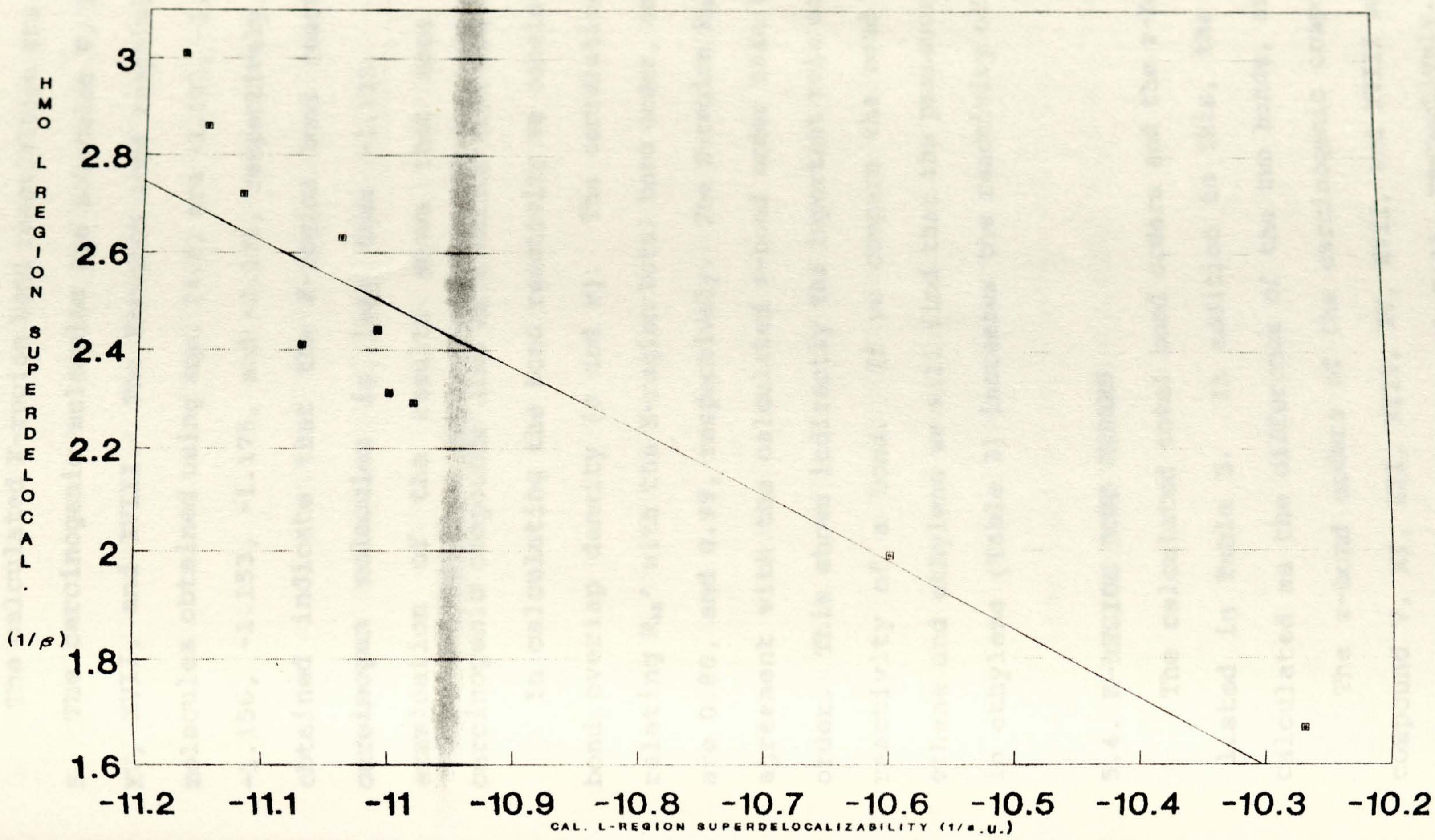


Fig. 17. CAL. L-region superdelocalizability index versus HMO L-region superdelocalizability index.

5.3. K-REGION BOND REACTIVITY

The calculated K-region bond reactivities are given in Table 5. **The carcinogenic molecules are molecules V, XI, XII, XIII, XV, XVII, and XVIII.** The K-region bond reactivities of these molecules obtained using eqn. (4.6) are -1.123, -1.166, -1.150, -1.156, -1.153, -1.175, and -1.131, respectively. The results obtained **indicate that the K-region bond reactivity of the carcinogen molecules is less than -1.123.** But, careful examination of the results shows that some of the non-carcinogenic compounds also have values less than -1.123.

In calculating the bond reactivity we considered the total bond overlap density (σ and π). The correlation coefficients relating R_{AB}' with the K-region total bond order, and π-bond order are 0.80, and 0.99, respectively. The K-region R_{AB}' is in better agreement with the calculated π-bond order than the total bond order. This shows indirectly the important role of π-bond in the reactivity of a bond. If we compare the bond reactivity of ethane and ethylene we will find that the presence of the π-bond in ethylene (Table 2) increases the reactivity of the molecule.

5.4. K-REGION BOND ORDERS

The calculated total bond orders and the π-bond orders are listed in Table 5. In addition to this, the σ-bond order, calculated as the difference of the two bonds, are shown.

The π-bond orders of the carcinogenic compounds, i.e. of compound V, XI, XII, XIII, XV, XVII, and XVIII are 0.80, 0.87, 0.85, 0.86, 0.85, 0.88, and 0.82, respectively. This result shows that all the carcinogenic compounds have π-bond orders greater than or equal to 0.80 , but it is found that some of the non-carcinogenic compounds have π-bond order greater than 0.80. For example, molecule XIX has no carcinogenic activity, but it has a calculated π-bond order in the K-region of 0.87, which exceeds that of XV, an active carcinogen.

No significant differences in the σ-bond orders of the carcinogenic and non-carcinogenic compounds are observed. This observation conforms with the result we obtained above (sec. 4.3), that is, a π-bond is important for the reactivity of a bond.

5.5. K-REGION BOND OVERLAP POPULATION

The result of the K-region bond overlap population calculation for the molecules studied is shown in Table 5. It is interesting to note that the K-region overlap population of all carcinogenic hydrocarbons are found to be greater than or equal to 1.026. Some of the non-carcinogenic molecules such as XVI (1.049) and XXII (1.050) have K-region bond overlap populations greater than 1.026. As a result not all molecules with a K-region overlap population exceeding 1.026 are carcinogenic.

Table 5. Calculated K-region SD indices (I_K'), SD indices obtained using HMO (I_K)[f], total bond order (W_{AB}), π-bond order (P_{AB}^{π}), overlap bond population (P_{AB}^{ov}), bond reactivity (R_{AB}), and σ-bond order (P_{AB}^{σ}).

COMPD	K-REG.	I_K'	I_K	R_{AB}	P_{AB}^{ov}	W_{AB}	P_{AB}^{π}	P_{AB}^{σ}
I	**(1-2)**	**-10.27**	**1.67**	**-0.997**	**0.951**	**1.44**	**0.67**	**0.77**
II	(1-2)	-10.60	1.87	-1.081	1.000	1.61	0.77	0.84
III	**(1-2)**	**-10.55**	**1.99**	**-1.113**	**1.019**	**1.67**	**0.81**	**0.86**
IV	**9-10**	**-10.66**	**1.99**	**-1.141**	1.038	1.73	**0.84**	**0.89**
V	**1-2**	**-10.69**	**1.97**	**-1.123**	**1.026**	**1.69**	**0.80**	**0.89**
VI	3-4	-10.70	2.08	-1.159	1.047	1.76	0.86	0.90
VII	(1-2)	-10.58	2.08	-1.130	1.029	1.70	0.83	0.87
VIII	1-2	-10.68	2.00	-1.126	1.028	1.69	0.82	0.87
IX	1-2	-10.66	2.05	-1.147	1.039	1.74	0.85	0.89
X	1-2	-10.32	1.80	-1.018	0.974	1.51	0.71	0.80
XI	6-7	-10.70	2.14	-1.166	1.051	1.78	0.87	0.91
XII	**3-4**	**-10.70**	**2.06**	**-1.150**	**1.041**	**1.74**	**0.85**	**0.89**
XIII	3-4	-10.70	2.05	-1.156	1.046	1.76	0.86	0.90
XIV	3-4	-10.72	2.14	-1.168	1.052	1.78	0.87	0.91
XV	6-7	-10.73	2.17	-1.153	1.042	1.74	0.85	0.89
XVI	4-5	-10.70	2.16	-1.165	1.049	1.78	0.87	0.91
XVII	**6-7**	**-10.74**	**-**	**-1.175**	**1.055**	**1.80**	**0.88**	**0.92**
XVIII	**7-8**	**-10.70**	**1.98**	**-1.149**	**1.043**	1.74	0.85	0.89
XIX	**6-7**	**-10.74**	**2.15**	**-1.165**	1.050	1.78	0.87	0.91
XX	6-7	-10.69	2.03	-1.149	1.043	1.74	0.85	0.89
XXI	**(5-6)**	**-10.54**	**1.93**	**-1.102**	**1.014**	1.65	0.79	0.86
XXII	6-7	-10.73	2.20	-1.172	1.054	1.79	0.88	0.91

[f] From Ref. 17

In Table 6 the K-region bond overlap population calculated using eqn. 4.13 for some of the molecules studied are listed. It is compared to the K-region bond overlap population obtained by the SAMO method (using eqn. 4.15).

As mentioned in the first part of this thesis, the aim of this study is to develop an all-valence electron semiempirical quantum chemical method for screening carcinogenic aromatic hydrocarbons. As seen in Table 3 we have seven carcinogenic compounds. These are compounds V, XI, XII, XIII, XV, XVII, and XVIII. Using our superdelocalizability indices we found only five carcinogenic compounds out of the 22 compounds in Fig. 13. The five carcinogenic compounds are compound XI, XII, XIII, XV, and XVIII. Only two compounds, V and XVII, out of the 22 compounds do not obey the criteria stated using our superdelocalizability indices. But, using HMO superdelocalizability indices we have four exceptions, molecules V, IX, XVI, and XVIII. Based on this result we can say that our superdelocalizability indices are better in predicting carcinogenic activity of aromatic hydrocarbons.

In case of the calculated K-region bond reactivity, total bond order, π-bond order and bond overlap population, it is found that the calculated parameters for some of the non-carcinogenic molecules are in the range of the carcinogenic molecules. As an example we can take compound XIV, which has no carcinogenic activity[58] but has calculated K-region π-bond 0.87, bond overlap population of 1.052, and bond reactivity of -1.168, which exceed that of XII, an active carcinogen.

In general, when we observe the results of all the calculated parameters, with the exception of the superdelocalizability indices, similar predictions are made as to the activities of the molecules. If one examines the values of the K-region bond reactivity, total bond order, π-bond order, and bond overlap population listed in Table 4 carefully, one will find that out of the 22 molecules, 17 are carcinogenic in all cases. This shows that in determining carcinogenic potency, a factor other than the K-region total bond order, π-bond order and/or bond overlap population should be considered. In other words, the K-region bond reactivity, total bond order, π-bond order, and bond overlap population can not be used as reactivity parameters in pre-screening aromatic hydrocarbons together or alone.

5.6. IONIZATION POTENTIAL

In addition to the reactivity parameters, the ionization potential of some of the molecules studied was also calculated. The ionization potentials obtained from the energy of the highest occupied molecular orbital using Koopmans' theorem[37], i.e., the negative of the orbital energies for occupied HF MOs are to be interpreted as IPs.

Table 7 shows the calculated ionization potentials (IP) along with the experimental ionization potential and the ionization potential, obtained by SAMO method. The calculated ionization potentials differ from the experimental, but this to be expected using the CNDO/2 method. The CNDO/2 is an all-valence electron method which involves a number of approximations (section 3.3.1). However, as the correlation coefficient shows the ordering within this series is in good agreement (correlation coefficient = 0.986) with the experimental value. The correlation coefficient of the SAMO method with the experimental value is 0.987.

The carcinogenic hydrocarbons for which the ionization potentials calculated (Table 7) have low ionization potentials and therefore can be expected to behave as strong electron donors. This is consistent with what is observed experimentally[59,60].

Table 6. K-region bond overlap population of some of the studied molecules calculated by CNDO/2 and SAMO method.

COMPOUND	**P_{AB}^{ov}(SAMO)[g]**	**P_{AB}^{ov}(CNDO)**	**CARCINOGENICITY[f]**
II	1.262	0.951	-
III	1.274	1.000	-
IV	1.286	1.038	-
IX	**1.286**	**1.039**	-
VI	1.293	1.047	-
XI	**1.292**	**1.051**	**++++**
XII	**1.288**	**1.041**	**++**
XIV	**1.296**	1.052	-
XVI	1.292	1.049	-
XVII	1.295	1.055	++++

[f] From Ref. 17
[g] From Ref. 18

Table 7. Ionization potentials (eV) of some of the studied aromatic hydrocarbons calculated by the CNDO/2 and SAMO methods.

COMPOUND	I_A(CNDO)(ev)	I_A(SAMO)(ev)	I_A(EXP.)(ev)[g]
I	13.33	9.80	9.25
II	10.91	7.70	8.15
III	9.48	6.30	7.37
IV	10.62	7.40	7.86
VI	9.60	6.40	7.47
IX	9.49	6.50	7.55
XI	8.93	5.90	7.19
XII	9.66	6.50	7.80
XIV	8.71	5.60	-
XVI	8.31	5.50	7.01
XVII	8.72	5.60	-

[g] From Ref. 18

6.0. CONCLUSION

The superdelocalizability indices we calculated have good correspondence with the experimentally determined carcinogenic activity of the molecules. The relationship of the superdelocalizability indices we calculated to the experimental carcinogenic reactivity is better than the superdelocalizability indices previously obtained using the HMO method. The method we used to obtain the superdelocalizability indices has an advantage over the HMO method. In the method we used, there is no restriction on the geometry of the molecule to be studied. One can treat any molecule with an optimized geometry irrespective of its geometry, whether planar or non-planar. Conversely, in HMO method one can study only molecules which are planar and conjugated.

The HMO atomic superdelocalizability is calculated by taking only the π-electron density into consideration. Our atomic superdelocalizability on the other hand is calculated by considering the total electron density (σ- and π-electrons). The consideration of the σ-electron density will enable one to study reactivity of non-conjugated molecules and even molecules without π-electrons. This is another advantage over the HMO method which is used to study only conjugated molecules.

The objective of this study was to develop an all-valence electron semiempirical quantum chemical method for screening carcinogenic aromatic hydrocarbons. Our results have closely demonstrated that the superdelocalizability indices can be used for studying the carcinogenic activity of aromatic hydrocarbons.

But, it is difficult to use the calculated K-region bond orders, bond overlap population, and bond reactivity as reactivity index to differentiate between non-carcinogenic and carcinogenic molecules. The reason is that the calculated values obtained for some of the non-carcinogenic compounds are in the range of the carcinogenic compounds. But, one can use the K-region bond reactivity, total bond order, π-bond order and bond overlap population together with other reactivity indices such as the superdelocalizability indices, because they can give valuable information as to the reactivity of the given K-region bond.

REFERENCES

1. R. Doll and R. Peto, J. Natl. Cancer Inst., 66, 1191 (1981).
2. L. N. Ferguson, J. Chem. Educ., 52, 688 (1975).
3. B. Pullman, "Molecular Association in Biology", Academic Press, N.Y., (1968).
4. C. L. Hewett, J. Chem. Soc., 293 (1940); F. Bergmann, Cancer Res., 2, 660 (1942).
5. R. Robinson, Brit. Med. J., I, 943 (1946).
6. C. F. Fieser, Amer. J. Cancer, 34, 37 (1938).
7. F. Bergmann, Cancer Res., 2, 660 (1942); G. M. Badger, Quart. Rev. (London), 5, 147 (1951).
8. G. M. Badger, Adv. Cancer Res., 2, 73 (1954).
9. M. S. Newman and S. Otsuka, J. Nat. Cancer Inst., 21, 721 (1968).
10. E. J. Moriconi and L. Salce, "Oxidation of Organic Compounds-III", Adv. in Chemistry Ser., Vol. 77, 65 (1968).
11. O. Schmidt, Z. Physik. Chem., 39, 59 (1938).
12 O. Schmidt, Z. Physik. Chem., 42, 83 (1938); 43, 185 (1939).
13. O. Schmidt, Naturwiss., 29, 146 (1941).
14. N. V. Svartholm, Arkiv Kermi Minerol. Geol., A15, 1 (1942).
15. A. Pullman, Bull. Soc. Chim. Fr., 21, 596 (1954).
16. A. Pullman and B. Pullman, Adv. Cancer Res., 3, 117 (1955).
17. M. A. Mainster and J. D. Memory, Bioch. Bioph. Acta, 148, 606 (1967).
18. B. J. Duke, D. R. Eilers, J. E. Eilers, S. Kang, A.

Liberles, and B. O'Leary, *Int. J. Quant. Chemistry.*, 155 (1975).

19. J. E. Eilers and D. R. Whitman, *J. Amer. Chem. Soc.*, 95, 2067 (1973).

20. R. G. Harvey, *Acc. Chem. Res.*, 14, 218 (1981); D. H. Phillips, *Nature (London)*, 303, 468 (1983).

21. D. M. Jerina, R. E. Lehr, H. Yagi, O. Hernandez, P. M. Dansette, P. G. Wislocki, A. W. Wood, R. L. Chang, W. Levin, and A. H. Conney, In " In Vitro Metabolic Activation in Mutagenesis Testing", F. de Serres, J. R. Bend, R. M. Philpot, Eds, Elsevier, Amsterdam, pp 159-177 (1976).

22. D. M. Jerina, J. M. Sayer, D. M. Thakker, H. Yagi, W. Levin, A. W. Wood, and A. H. Conney, In " Carcinogenesis: Fundamental Mechanisms and Environmental Effects", B. Pullman, P. O. P. Ts'o, and H. Gelboin, Eds., D. Reidel Publishing Co., Dordrecht, Holland, pp 1-12 (1981).

23. G. D. Berger, I. A. Smith, P. G. Seybold, and M. P. Serve, *Tetrahedron Letters*, 231 (1978).

24. S. Glassone, "Textbook of Physical Chemistry," 2nd. Ed. Macmillan, N. Y., pp. 1087 (1960).

25. G. W. Wheland, *J. Amer. Chem. Soc.*, 64, 900 (1942).

26. M. J. S. Dewar, *J. Amer. Chem. Soc.*, 74, 3357 (1952); M. J. S. Dewar, *Record Chem. Prog.*, 19, 1 (1958).

27. R. D. Brown, *J. Chem. Soc.*, 1951 (1950).

28. R. D. Brown, *Qurat. Revs.*, 6, 63 (1952).

29. K. Fukui, Y. Yonezawa and C. Nataga, *J. Chem. Phys.*, 27, 1247 (1957).

30. L. Salem, "The Molecular Orbital Theory of Conjugated Systems", W. A. Benjamin, Inc., N.Y., (1966).

31. G. H. Loew and T. Poulsen, Int. J. Quant. Chemisty, Quant. Biol. Symp., 8, 95 (1981).

32. M. Born and J. R. Oppenheimer, Ann. Physik, 84, 457 (1927).

33. D. R. Hartree, Proc. Cambridge Phil. Soc., 24, 89,111, 426 (1928).

34. J. C. Slater, Phys. Rev., 35, 509 (1938); 34, 1293 (1959).

35. C. C. J. Roothaan, Rev. Mod. Phys., 23, 69 (1951).

36. L. Pauling and E. B. Wilson, "Introduction to Quantum Mechanics", McGraw-Hill Book Company, N.Y., (1935).

37. T. Koopmans, Physica, 1, 104 (1933).

38. V. Fock, Z. Physik, 61, 126 (1930).

39. G. G. Hall, Proc. Roy. Soc. (London), A205, 541 (1955).

40. G. Naray-Szabo, P. R. Surjan, and J. G. Angyan, "Applied Quantum Chemistry", D. Reider Publishing Comp., Holland, (1987).

41. R. G. Parr, J. Chem. Phys., 20, 239 (1952).

42. J. A. Pople J. A. Santry, and G. A. Segal, J. Chem. Phys., 43, S129 (1965).

43. J. A. Pople, and G. A. Segal, J. Chem. Phys., 43, S136 (1965).

44. J. A. Pople and D. L. Beveridge, "Approximate Molecular Orbital Theory", McGraw-Hill, Inc., N.Y. (1970).

45. M. Goeppert-Mayer and A. L. Sklar, J. Chem. Phys., 6, 645 (1938).

46. J. A. Pople and S.A. Segal, J. Chem. Phys., 44, 3289, (1966).

47. R. S. Mulliken, J. Chem. Phys., 2, 782 (1934).

48. J. M. Sichiel and M. A. Whitehead, Theor. Chim. Acta, 7, 33 (1966).

49. R. Y. Boyd and M. A. Whitehead, J. Chem. Soc., Dalton Trans., 81 (1972).

50. J. Del Bene and H. H. Jaffe, J. Chem. Phys., 48, 1807, 4050 (1968).

51. E. Huckel, Z. Physik, 70, 204 (1931).

52. C. A. Coulson, Proc. Roy. Soc., A169, 413 (1939).

53. K. B. Wiberg, Tetrahedron, 24, 1083 (1968).

54. I. Mayer, Chem. Phys. Letter, 89, 390 (1982).

55. N. L. Allinger and Y. H. Yuh, Molecular Mechanics: MMP2 (Quantum Chemistry Program Exchange, Bloomington Indiana, 1981), Vol. 13, p. 395.

56. T. Clark, "A HandBook of Computational Chemistry", John Wiley and Sons, N.Y., (1985)

57. A. Streitwieser, "Molecular Orbital Theory for Organic Chemists", Wiley, N. Y., (1961).

58. S. S. Sung, C. R. , Acad. Sci. Paris, 274D, 1597 (1972).

59. D. E. Laskowski, in "Jerusalem Symp. on Quantum Chemistry and Biochemistry I, Physico-chemical Mechanisms of Carcinogensis", E. D. Bergemann and B. Pullman, Eds. (Israel Acad. Sci. Humanities, Jerusalem, 1969), p85.

60. D. E. Laskowski, Cancer Res., 27, 903 (1967).

DECLARATION

I, the undersigned, declare that this thesis is my work and that all sources of material used for the thesis have been duly acknowledged.

Name: Ahmed Mustefa Mohammed

Signature:

Chemistry Department, Addis Ababa University.

September, 1993.