Density Functional Theory Based Studies of Reactivity Descriptors to Predict the Regioselectivity of Chemical and Biological Systems

THESIS

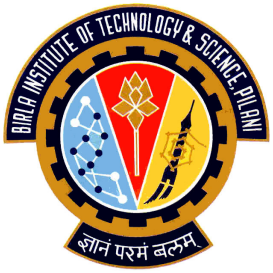
Submitted in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

By

Soumen Saha

Under the Supervision of
Dr. Ram Kinkar Roy



BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE
PILANI (RAJASTHAN) INDIA
2010

Dedicated to My Mother and Father

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE
PILANI (RAJASTHAN) INDIA

CERTIFICATE

This is to certify that the thesis entitled **"Density Functional Theory Based Studies of Reactivity Descriptors to Predict the Regio-selectivity of Chemical and Biological Systems"** submitted by **Soumen Saha, ID No. 2005PHXF411P** for award of Ph.D. Degree of the Institute, embodies original work done by him under my supervision.

Signature in full
of the Supervisor

Name in capital block
letters: **RAM KINKAR ROY**

Designation: Assistant Professor

Date:

ACKNOWLEDGEMENTS

It has been my great privilege to work with the inspiring and encouraging guidance of Dr. Ram Kinkar Roy, Assistant Professor, Chemistry Group, BITS, Pilani. I owe to him my sincere thanks for his invaluable suggestions and constructive criticism in planning, execution and completion of this study. I shall forever be indebted to him for inculcating in me the habit of working hard, being patient, team spirit and thinking independently.

It gives me great pleasure to thank Prof. B. M. Deb, Prof. D. Mukherjee, Prof. S. K. Ghosh, Prof. S. Pal and Prof. P. W. Ayers for their motivation, subtle guidance and valuable suggestions during various stages of this research work.

I am grateful to Prof. G. Sundar, Deputy Directors (Off-Campus); Dr. S. K. Saha, Assistant Professor (Chemistry Group); Dr. B. Layek, Assistant Professor (Physics Group) and Dr. A. Mukherjee, Assistant Professor (Physics Group), BITS, Pilani for their fruitful suggestions and constant help through the entire course of this work.

I extend my sincere thanks to my very affectionate senior Dr. P. Purkayastha, Assistant Professor (Department of Chemical Sciences, IISER, Kolkata), for ever- lasting support and fruitful company.

I thank Ritobrata Goswami, Saptarshi Basak, Tapta Kanchan Roy and Md. Sayem Alam for their unwavering supports throughout this work.

With profound reverence, I express my gratitude to the Vice Chancellor Prof. L. K. Maheswari, Director Prof. G. Raghurama, Deputy Director (Research & Education Development) Prof. R. N. Saha and Dean (Faculty Division I and Instruction Division) Prof. A. K. Sarkar for their ever willing guidance extended throughout the tenure of the present investigation.

My gratitude is due to the learned members of my Doctoral Advisory Committee, Dr. N. Rajesh, Dr. N. Nandi and Dr. P. U. Manohar for their cooperation and timely suggestions.

I am thankful to Prof. S. C. Sivasubramanium (Unit Chief, CAHU) and Dr. Dalip Kumar (Group Leader, Chemistry Group) for their kind words, constant cooperation and encouragement during various stages of this research work.

I feel privileged in expressing high regards to the faculty, Chemistry Group, BITS, Pilani for their ever-willing help and support.

I express my gratitude to Prof. A. K. Das, Dean, Research and Consultancy Division (RCD), BITS, Pilani for his constant official support, encouragement and making the organization of my research work through the past few years easy.

I am thankful to Dr. Hemant Jadav, Dr. Monica Sharma, Mr. Dinesh Kumar, Mr. Sharad Shrivastava, Mr. Gunjan Soni, Mr. Amit Kumar and Ms. Sunita Bansal and other office staff of Research and consultancy Division, BITS, Pilani for helping at different point of time during the course of my research work.

My sincere thanks to Dr. M. Ishwara Bhat, Librarian, BITS, Pilani and other staff of BITS, Library for their cooperation and help rendered while utilizing the library services.

The assistance provided by the office staffs of Chemistry Group, Sita-Ramji, Goukulji, Shankarji, Omprakashji and Pushplataji is heartly acknowledged.

I owe my special thanks to my friends Amit Tiwari, Bhupendra Mishra, Buchi Reddy, Deepak Pakalapati, Divya Sachdev, Garudachari B., Gautam Patel, K. P. Chandrashekhar, Kameshwar Rao, Kiran Soni, Pradeep Kumar L., Priyanka Bagaria, Maheswara Rao, Mallari Naik, Manoj Kumar, Maruthi Kumar, Sudershan Rao, Sonu, Sowmya, Swapna Sundaree, Swarna Kanchan, Thirumurthy and Vasanthakumar S. for their help and charming company. Their timely help and friendship shall always be remembered.

The Financial Assistance in the form of DST Junior as well as Senior Research Fellowship and CSIR Senior Research Fellowship is thankfully acknowledged.

Words are inadequate to express my feelings for my sister (Anamika) and my brother-in-law (Sandip) for their constant encouragement.

Date

(SOUMEN SAHA)

ABSTRACT

This thesis describes a novel study on Conceptual Density Functional Theory (DFT) based reactivity descriptors for use in predicting the regioselectivity of large systems, bimolecular systems. The challenges of bio-systems in particular the large number of atoms and high structural flexibility, made the way to a routine application of DFT more laborious. To cope with extended systems, a fragmentation based approach has been developed (given the name 'One-into-Many' model) which is seen to lead to a reliable determination of the regioselectivity of model biomolecular systems.

In the first part of the thesis, a brief review of literature, objectives and motivation for the present thesis are given. An overview of the Conceptual DFT is given, which provides the theoretical foundation of different reactivity descriptors. These descriptors enable us to understand experimental observations in an elegant way. An important aim of this thesis is to verify and interpret the correlation of these descriptors with the experimental studies at macromolecular level. Some advances achieved in last few years to predict the regioselectivity of the large bio-molecular systems using Conceptual DFT based reactivity descriptors are discussed in detail.

So far there has been no schematic study for modeling or analysing the regio-chemistry of large biomolecular systems using DFT-based reactivity descriptors. The second part of the thesis discusses the development of a new model (named as 'One-into-Many') to predict the regioselectivity of large chemical and biological systems. Large chemical and biological systems with multiple reactive sites are proposed to be broken into small fragments having at least one reactive site in each fragment. The environment around each reactive site is mimicked by incorporating a buffer zone. Local reactivity is evaluated for each reactive site adopting a new modified approach of local hardness, $\eta(\bar{r})$. Theoretical justification in favour of using the local hardness based descriptor to compare the intermolecular reactivity sequences (the fragments can be treated as individual molecules) has been discussed. When the 'One-into-Many' model is applied to predict the regioselectivity (towards an electrophilic attack on it) of the base-pairs in DNA (PDB ID: 1BNA) (*Proc. Natl. Acad. Sci. USA* **1981,** *78*, 2179) the generated results are found to be satisfactory in most cases.

The reactivity descriptor, used in 'One-into-Many' model as a key tool, is local hardness ($\eta(\bar{r})$) because its predominant component is electronic contribution to the molecular electrostatic potential (MEP). MEP has a long distantce effect, thus making it suitable for predicting intermolecular reactivity and so fitting the proposed model. However,

$\eta(\bar{r})$ (or better it's condensed form, $\eta(k)$) suffers from one severe limitation and that is the N-dependence problem (here, N is the total number of electrons of the system). In case of studying the regioselectivity (using 'One-into-Many' model) of DNA systems (i.e., in chapter II) N-dependence problem was solved automatically as the number of electrons in all base-pairs (whether it is 'single' or 'triple' or higher base-pair systems) are same. The broader applicability of local hardness, ($\eta(\bar{r})$) as a reliable intermolecular reactivity descriptor primarily depends on the removal of the $\frac{1}{N}$ dependence. In the third part of the thesis an attempt is made to illustrate the problem of evaluating $\eta(\bar{r})$ by using $\rho(\bar{r})$ (the electron density) or $Nf(\bar{r})$ (Fukui function multipled by N) as composite functions. A plausible scheme to bypass those problems of evaluating $\eta(\bar{r})$ is also demonstrated.

However, this is not a general solution applicable to all kinds of systems. In order to make 'One-into-Many' model widely applicable, it should be based on a descriptor, which has the essential quality of taking care of intermolecular reactivity aspects and at the same time N-dependence problem removed analytically. In the fourth part of the thesis a simple formalism to obtain hardness potential (as defined by Parr and Gazquez, *J. Phys. Chem.*, 1993, **97**, 3939) is presented. Use of hardness potential formally resolves the *N*-dependence problem of local hardness. It is shown that the hardness potential has the ability to describe the intermolecular reactivity sequence of a series of chemical systems. The corresponding electrophilic ($\Delta^{+}h(k)$) and nucleophilic ($\Delta^{-}h(k)$) variants of the hardness potential are also developed, which measure the reactivity toward a nucleophilic (i.e., Nu^{-}) and an electrophilic (i.e., El^{+}) reagent, respectively. Several examples are chosen and it is shown that the proposed reactivity descriptors correctly predict the expected trend in each case.

Finally, in chapter V the results and conclusions of the research presented in this thesis are summarized. Areas that require further exploration are identified and accordingly, future scope of work is presented.

LIST OF ABBREVIATIONS AND SYMBOLS

α	Polarizability
A	Adenine
ArsC	Arsenate reductase
BLYP	Becke exchange and Lee-Yang-Parr correlation functional
B3LYP	Becke three-parameter Lee-Yang-Parr functional
χ	Electronegativity
C	Cytosine
cc-pVDZ	Correlation consistent polarized valence double zeta (basis set)
cc-pVTZ	Correlation consistent polarized valence triple zeta (basis set)
CDASE	Comprehensive decomposition analysis of the stabilization energy
CI	Configuration interaction
ΔE_{SE}	Stabilization energy
$\Delta G^{\#}$	Free energy of activation
ΔN	Electron transfer
$\Delta S^{\#}$	Activation entropy
DC	Divide-and-conquer
DFT	Density functional theory
DMM	Density matrix minimization
DNA	Deoxyribonucleic acid
DNP	Double numeric with polarization
E	Energy
E_a	Activation energy
$E[\rho]$	Total electronic energy
EA	Electron affinity
EBF	Energy based fragmentation
EEM	Electronegativity equalization method
EFP	Effective fragment potential
El^+	Electrophile
$E_{XC}[\rho]$	Exchange-correlation energy
$f(\bar{r})$	Fukui function

$f(k)$	Condensed-to-atom Fukui function
$F[\rho]$	Hohenberg-Kohn universal Functional
FC	Frozen core
FF	Fukui function
FMM	fast multipole methods
G	Guanine
GGA	Generalized gradient approximation
GTO	Gaussian type orbital
η	Chemical hardness
$\eta(\bar{r})$	Local hardness
$\eta(k)$	Atomic local hardness
$\eta(\bar{r},\bar{r}')$	Hardness kernel
$H[\rho]$	Hardness functional
h	Planck's constant
$h(\bar{r})$	Hardness potential
$h(k)$	Atomic hardness potential
HIV	Human immunodeficiency virus
HOMO	Highest occupied molecular orbital
HPA	Hirshfeld population analysis
HSAB	Hard and soft acid and base
IP	Ionization potential
$J[\rho]$	Coulombic interaction energy
K	Equilibrium constant
k	Rate of reaction
k_B	Boltzmann constant
LDA	Local density approximation
LRD	Local reactivity descriptors
LSDA	Local spin density approximation
LUMO	Lowest occupied molecular orbital
μ	Chemical potential
MEDLA	Molecular electron density lego assembler
MEF	Molecular electrostatic field

MESP	Molecular electrostatic potential
MPA	Mulliken population analysis
MP*n*	Møller-Plesset perturbation theory of order *n*
MPP	Minimum polarizability principle
N	Total number of electrons
NOA	Natural orbital analysis
Nu^-	Nucleophile
P_k	Gross electronic population of atom k in the molecule
PEP	Phosphoenolpyruvate
PMH	Principle of maximum hardness
PWC	Perdew-Wang correlation functional
PW91	Perdew-Wang (1991) correlation functional
q_k	Effective atomic charge
QSAR	Quantitative structure-activity relationship
$\bar{r}$	Electronic position
RHF	Restricted Hartree-Fock
$\rho(\bar{r})$	Electron density
ROHF	Restricted open-shell Hartree-Fock
S	Softness
S_N2	Bimolecular nucleophilic substitution
$s(\bar{r})$	Local hardness
$s(k)$	Condensed-to-atom local softness
$s(\bar{r},\bar{r}')$	Softness kernel
SP-DFT	Spin-polarized density functional theory
STO	Slater type orbital
T	Thymine
T	Temperature
$T_k[\rho]$	Electronic kinetic energy
TFD	Thomas-Fermi-Dirac theory
TS	Transition state
UHF	Unrestricted Hartree Fock
w	Electrophilicity index

$w(\bar{r})$	Local electrophilicity index
$v(\bar{r})$	External potential
$V_{ee}[\rho]$	Electron-electron repulsion energy
$v_{XC}(\bar{r})$	Exchange correlation potential
ψ	Electronic wave function
ZPE	Zero point energy

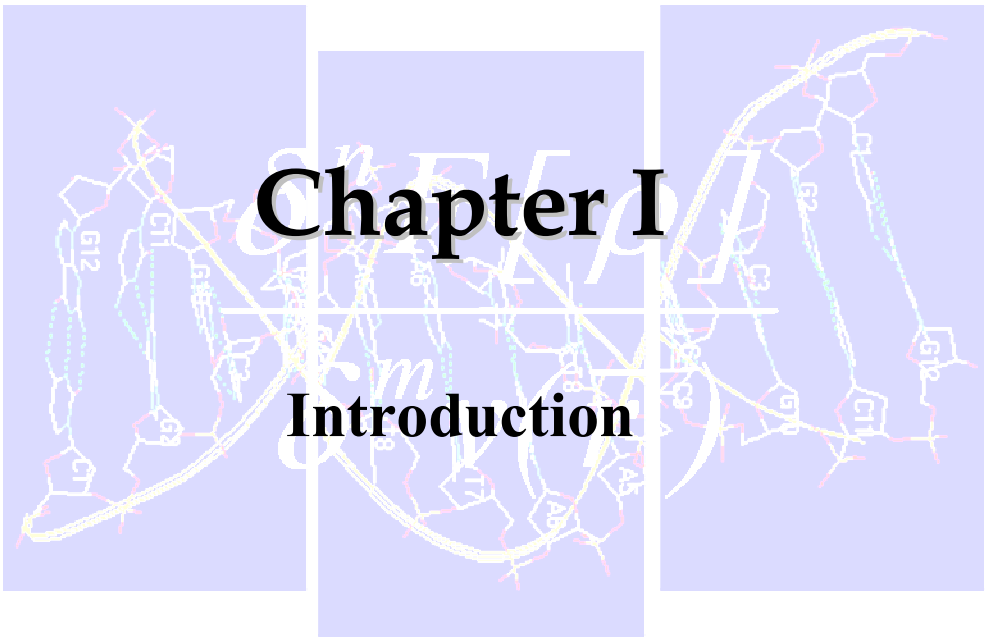
CONTENTS

LIST OF TABLES

Page No.

LIST OF FIGURES

Page No.



Chapter I

Introduction

1.1. Introduction

An important chemical concept prevalent in chemistry (organic chemistry, in particular) is regioselectivity. Regioselectivity[1,2] is defined as the property of a chemical reaction of producing one structural isomer in prefence to others that are theoretically possible. Understanding the regioselectivity of a reaction between two chemical species is useful not only for predicting the corresponding reaction mechanism but also for designing desired products. In the last few decades several electronic parameters, *viz*, Frontier Molecular Orbital (FMO),[3-6] Electron Localized Function (ELF),[7,8] Molecular electrostatic Potential (MEP)[9-17] etc., have been proposed and used extensively to explain the regioselectivity of a wide variety of reactions. Similarly, empirical principles such as the hard and soft acids and bases (HSAB),[18-22] Electronegativity equalization method (EEM),[23-28] etc. have been developed to rationalize chemical behaviour. However, most of these approaches remained empirical until a branch of density functional theory (DFT),[29-41] called "Conceptual DFT" or "Chemical Reactivity Theory", was initiated by R. G. Parr. Based on the idea that the electron density is the fundamental quantity for describing atomic and molecular ground states, Parr and co-workers, and later on a large community of theoretical chemists provided the theoretical basis for formal definitions of empirical concepts.[42-52] Conceptual DFT have been successful even to propose a new quantitative principle, the 'principle of maximum hardness' (PMH),[53-64] which can predict the most stable state of a chemical species.

Although, DFT provided a sound basis for the development of computational strategies for obtaining information about molecules at much lower cost than the conventional *ab initio*[65] wave function techniques, these methods are still not routinely feasible for large systems such as biological molecules and molecular systems with hundreds or thousands of atoms, due to the steep increase of their computational cost with increasing molecular size. To extend quantum chemical calculations or DFT calculations to macromolecules, theoretical chemists have come up with a variety of approaches, which allow Hartee-Fock (HF), DFT or post-HF calculations to achieve linear scaling. Linear-scaling methods are primarily based on the principle of quantum locality[66] or "near-sighted-ness",[67] that the properties of a certain observation region of

only one or a few atoms are only weakly influenced by factors that are spatially far away from this observation region. This can be achieved by limiting to a local region of space the physical span of the electronic degrees of freedom.[46] Careful consideration of such underlying physics and improved mathematical methods have led to linear scaling in, inter alia, the calculations of the Coulomb[68-74] and exchange[75,76] integrals, and in alternative approaches to the direct diagonalization of the Fock matrix.[77-85] These rigorous algorithms treat the molecule as a whole, being "black-box". Nevertheless, these algorithms begin to exhibit linear scaling only for quite large molecules, say, with several hundreds of atoms.

In addition to this category of linear-scaling algorithms that are aimed to calculate the properties of the whole system at once, there also exists a category of fragment-based approaches[86-113] which are capable of reproducing *ab initio* HF or post-HF results of large molecules quite accurately but with much lower computational costs. The basic idea in these is to divide a large molecule into a set of fragments, and then obtain the energy or molecular properties of this molecule from conventional quantum chemical calculations on a series of subsystems, each of which is constructed by connecting a fragment with its local surroundings. These methods not only offer a very considerable reduction of the computational costs but also allude to the chemical building blocks in larger systems, such as residues taken as fragments, and provide details of the interaction and other properties of these fragments-in-molecules. Molecular fragmentation approaches are of two main types. One is the density matrix-based fragmentation approach,[86,87,90,92,95,97,100,105,108] in which the density matrix of the target molecule is obtained by assembling the density matrices or localized molecular orbitals from various subsystems, which is then employed to calculate the total energy or some properties of the target molecule. Another type can be named as the energy-based fragmentation (EBF) approach.[91,93,96,99,101,102,104] In this approach, the total energy of a molecule is approximately estimated as linear combinations of the energies of its various subsystems, like, energy or heat of formation of a molecule being approximated as a sum of bond energies or enthalpies. In comparison to these density matrix-based approaches, energy-based approaches have one main advantage. Within energy-based approaches, the energy derivatives or other molecular properties of the target system can be computed as

combinations of the corresponding quantities from various subsystems while no such simple algorithms exist within density matrix-based approaches.[93,96,98,109]

The objective of the research work presented herein is to predict the regioselectivity of the large bio-molecular systems using Conceptual DFT (or chemical reactivity theory) based reactivity descriptors. Before going to actual descriptions of the recent works on DFT based reactivity descriptors used to predict the regioselectivity of large systems, a brief discussion on the background of the reactivity descriptors, which are just relevant for the above purpose, justified. In the second part of this chapter (i.e., Section 1.2), the Conceptual DFT is described briefly, which provides the theoretical foundations of different reactivity descriptors. The third part (i.e., Section 1.3) takes care of more recent developments enabling evaluation of the regioselectivity for a number of large biological systems using these reactivity descriptors. A systematic excursion of the thesis is presented in the last section (i.e., Section 1.4) of this introductory chapter.

1.2. Theoretical Background

A. Foundations

In 1964, Pierre Hohenberg and Walter Kohn[29] proved that for molecules with a nondegenerate ground state, the ground state molecular energy, wave function and all other molecular electronic properties are uniquely determined by the ground state electron density $\rho(x,y,z)$. One says that the ground state electronic energy E is a functional of ρ and writes

$$E = E[\rho] \tag{1.1}$$

where the square brackets denote a functional relation. Density Functional Theory (DFT) attempts to calculate E and other ground state molecular properties from the ground state electron density, ρ.

The total electronic energy is given by[114]

$$E[\rho] = F[\rho] + \int v(\bar{r})\rho(\bar{r})d\bar{r} \tag{1.2}$$

where the functional $F[\rho]$, so-called Hohenberg-Kohn functional,[29] is the sum of the kinetic energy functional $T[\rho]$ and the electron-electron repulsion energy functional $V_{ee}[\rho]$; $v(\bar{r})$ is the external potential (i.e., the potential acting on an electron at position $\bar{r}$ due to the presence of nuclei plus any other external field, if present). To turn from a formal relation to a practical tool, we need a second theorem also proved by Hohenberg and Kohn,[29] and a practical approach of the evaluation of which was developed by Kohn and Sham.[30] This second theorem, is a variational principle, is stating that the ground state density minimizes the energy of the system for a fixed number of electrons

$$\delta(E[\rho] - \mu\int \rho(\bar{r})d\bar{r}) = 0 \tag{1.3}$$

where μ is a Lagrange multiplier associated with the normalization constant $\int \rho(\bar{r})d\bar{r} = N$; here, N is the total number of electrons in the ground state of the system. Otherwise

$$\mu = v(\bar{r}) + \frac{\delta F[\rho]}{\delta\rho(\bar{r})} = constant \tag{1.4}$$

Kohn and Sham rewrote Eq (1.4) as an orbital equation having the form[30]

$$\left[-\frac{1}{2}\nabla^2 + v(\bar{r}) + v_{xc}(\bar{r}) + \int \frac{\rho(\bar{r}')}{|\bar{r} - \bar{r}'|}d\bar{r}'\right]\psi_i = \varepsilon_i\psi_i \tag{1.5}$$

where $v_{xc}(\bar{r})$ is the exchange-correlation potential, the functional derivative of the exchange-correlation energy functional $E_{xc}[\rho]$, i.e.,

$$v_{xc}(\bar{r}) = \frac{\delta E_{xc}[\rho]}{\delta\rho(\bar{r})} \tag{1.6}$$

In Eq (1.5), ψ_i's are the Kohn-Sham orbitals, the squares of which must sum up to the total electron density of the system

$$\rho(\bar{r}) = \sum_i |\psi_i|^2 \tag{1.7}$$

In recent years, many accurate forms of the exchange correlation functional have been derived.[115-118] A great strength of the density functional language is its appropriateness for defining and elucidating important universal concepts of molecular structure and molecular reactivity. It has become clear in recent years that there is also a

very important "noncomputational" or conceptual side to DFT.[43,45,46,49,51] In this aspect of the theory, the central quantities are the so-called response functions.[44,47,50,52] As such, they are reactivity descriptors or reactivity indicators for the molecule under consideration: these terms measure the response of the chemical system to perturbations in its number of electrons, N, and/or the external potential, $v(\bar{r})$. The reactivity descriptors allow one to predict what sorts of perturbations stabilize the molecule the most or alternatively, destabilize the molecule the least. This, in turn, allows one to predict toward what sorts of reagents the molecule will be most reactive. It also allows one to predict the regioselectivity of the reactions with those reagents.

B. Reactivity Descriptors

The response functions can be split into three groups: global, local, and nonlocal. The global quantities describe global responses against global perturbations. Such quantities do not depend on the spatial position $\bar{r}$ within the molecular framework, but characterize the entire system as an entity. Hence, they do not contain any information about regioselectivity. The local descriptors (i.e., $\bar{r}$ dependent) are associated to global/local responses of the system against local/global perturbations. These quantities are therefore suitable to describe the molecular selectivity because they vary locally from one position to another in a molecule. So the local reactivity descriptors are key in making predictions about regioselectivity. The nonlocal indices (i.e., quantities depending of two or more spatial positions, $\bar{r}$, $\bar{r}'$, etc.) are hence associated to local responses as a result of local perturbations. Nonlocal reactivity descriptors either measure a molecule's polarization with respect to its environment or the change in polarization associated with electron transfer. All these descriptors provide us a status to understand experimental observations in an elegant way. An important aspect of this thesis is to verify and interpret the correlation of these descriptors with the experimental studies at macromolecular level. Hence, it is very essential to know which parameters represent molecular structure and reactivity, and which represent the tendency of a given molecule to undergo a certain class of reactions.

(i) ***Global Reactivity Descriptors:*** Global reactivity descriptors measure the overall reactivity of a molecule. These reactivity descriptors can be considered as response functions describing the system's response to perturbations in the number of electrons N at constant $v(\bar{r})$.

After the introduction of DFT, advancement in the chemical reactivity was observed by concentrating on the interpretation of the Lagrangian multiplier μ in Eq (1.4). It has been an impressive contribution by Parr *et al*.[42] to identify this abstract multiplier as the partial derivative of the systems energy with respect to the number of electrons at constant external potential, $v(\bar{r})$ (i.e., identical nuclear charges and positions)

$$\mu = \left(\frac{\partial E}{\partial N}\right)_{v(\bar{r})} \tag{1.8}$$

The physical meaning of chemical potential in DFT is to measure the escaping tendency of an electron cloud. It is constant in three dimensional space for the ground state of an atom, molecule or solid and equals the slope of E versus N curve at constant external potential. Assuming continuity and differentiability of E, the quantity $-\left(\frac{\partial E}{\partial N}\right)_{v(\bar{r})}$ is easily seen to be a measure of the electronegativity, χ, of the atom. Thus, it is now pertinent to note that the chemical potential (μ) is exactly identical with the definition of one of the important concepts, electronegativity (χ), for which a number of definitions are available starting from Pauling's work.[119,120] Interestingly, Iczkowski and Margrave,[121] in an important contribution to the literature of electronegativity have defined the electronegativity (χ) of a system by the following,

$$\chi = -\left(\frac{\partial E}{\partial N}\right) \tag{1.9}$$

Mulliken's[122] definition of electronegativity is given as the arithemetic average of two experimentally measurable quantities, i.e., ionization potential (*IP*) and electron affinity (*EA*):

$$\chi = \frac{IP + EA}{2} \tag{1.10}$$

The expression is just the finite difference approximation to the term, $-\left(\frac{dE}{dN}\right)$. However, now within the framework of DFT, Parr and his collaborators[42] have provided the theoretically justified definition of the electronegativity, χ, to minus the chemical potential, μ in a natural way:

$$\chi = -\mu = -\left(\frac{\partial E}{\partial N}\right)_{v(\bar{r})} \tag{1.11}$$

The idea that electronegativity is a chemical potential originates with Gyftopoulos and Hatsopoulos.[123]

The operational definition of μ and χ are provided by the finite difference approximation[20] from $E(N)$ vs N curve, in which the first derivative $\left(\frac{\partial E}{\partial N}\right)$, μ is calculated as the average of the left- and right-hand side derivatives. The left derivative is obtained as the finite difference of energy of cation, $N-1$, and neutral, N, (usually neutral, but may be charged) electrons. This is simply equal to negative of *IP*. Similarly, the right derivative is obtained as difference of neutral (N) and anion ($N+1$) electrons. This is equal to the negative of *EA*.

$$\mu^{-} = \frac{E(N-1) - E(N)}{-1} = -IP \tag{1.12}$$

$$\mu^{+} = E(N+1) - E(N) = -EA \tag{1.13}$$

$$\left(\frac{\partial E}{\partial N}\right)_{v(\bar{r})} = \mu = \frac{1}{2}\left(\mu^{+} + \mu^{-}\right) = -\frac{1}{2}\left(IP + EA\right) \tag{1.14}$$

Thus, from Eq (1.11) electronegativity (χ) can be written as

$$\chi = -\mu = \frac{1}{2}\left(IP + EA\right) \tag{1.15}$$

The expression of χ originated from here is similar to that of Mulliken (i.e., Eq (1.11))[122]. As an approximation to Eq (1.15), one can relate chemical potential (μ) to the frontier orbital energies. This can be obtained through the Koopman's approximation[124,125] within the molecular orbital theory wherein *IP* and *EA* can be replaced by frontier orbital energies (i.e., HOMO and LUMO energy, in conventional

notation LUMO represents the *lowest unoccupied molecular orbital* in the species in question, and HOMO the *highest occupied molecular orbital*) as,[126-129,53,130,131]

$$-E_{HOMO} = IP \tag{1.16}$$

$$-E_{LUMO} = EA \tag{1.17}$$

Therefore, using Koopman's theorem,[124] we can write

$$\mu = -\chi = \frac{E_{LUMO} + E_{HOMO}}{2} \tag{1.18}$$

The physical significance is that the negative of χ represents a horizontal line at the energy midpoint between HOMO and LUMO. This approximation might be of some use when large systems are considered as it requires a single calculation (i.e., only for neutral system), whereas the evaluation of Eq (1.15) necessitates three calculations (i.e., for cationic and anionic system along with the neutral one), which is computationally expensive and sometimes very complicated to compute. Also, in the case of systems leading to metastable $N+1$ electron systems (typically anion), the problem of negative electron affinities is sometime avoided via Eq (1.18).[132-134]

Moreover, theoretical justification was provided for Sanderson's principle of electronegativity equalization[23,135,26,136] which states that when two or more atoms come together to form a molecule, their electronegativities become adjusted to the same intermediate value. Electronegativity, being synonymous with chemical potential, the correctness of Sanderson's principle immediately follows from the fact that the chemical potential of DFT is a property of an equilibrium state. The chemical potential (electronegativity) is expected to be sensitive to the external potential and may not be necessarily easy to calculate, but it is a concept securely rooted in DFT. Semiempirical electronegativity equalization methods now are widely used.[28]

E versus N plots are not straight lines but generally convex upward. Their curvatures define another property of substantial importance, the chemical hardness (η)[20]

$$\eta = \left(\frac{\partial^2 E}{\partial N^2}\right)_{v(\vec{r})} = \left(\frac{\partial \mu}{\partial N}\right)_{v(\vec{r})} \tag{1.19}$$

The chemical hardness was introduced by Pearson in the framework of his classification of Lewis acids and bases, leading to the introduction of the hard and soft acids and bases

principle (HSAB).[18,19,137,54,138-140] This principle states that hard acids prefer to bond to hard bases and soft acids to soft bases. A factor of two, included in the original definition of η, is omitted now as Parr himself recommended.[141,142] Again, using a finite difference approximation and a quadratic $E = E(N)$ curve, this equation reduces to

$$\eta = IP - EA \tag{1.20}$$

which, after using Koopmans approximation,[124] becomes

$$\eta = \varepsilon_{LUMO} - \varepsilon_{HOMO} \tag{1.21}$$

For an insulator or semiconductor, hardness is the band gap. When the gap is large (other things being equal), one expects high stability and low reactivity. When it is small, one expects low stability and high reactivity. These predictions are well borne out in the good correlation that exists between HOMO-LUMO gap and the organic chemists' concept of aromaticity.[143] This finding is nicely captured in the maximum hardness principle also, proposed by Pearson,[53] which states that "molecules will arrange themselves to be as hard as possible". Parr and Chattaraj provided a rigorous proof for this principle based on the fluctuation-dissipation theorem.[55,144-147]

The inverse of the global hardness is called the global softness[141,138]

$$S = \frac{1}{\eta} = \left(\frac{\partial N}{\partial \mu}\right)_{v(\bar{r})} \tag{1.22}$$

which was empirically shown to be proportional to the polarizability of the system.[148-152,147,153] The hardness can be thought of as a resistance to charge transfer, while the softness measures the ease of transfer.

Drawing analogy from classical thermodynamics, Parr and Pearson[20] developed the formalism for energy lowering i.e., the stabilization energy (SE), due to electron transfer between two chemical species A and B. If chemical potentials of the two species are μ_A^o and μ_B^o respectively, and $\mu_B^o > \mu_A^o$ (i.e., A is more electronegative than B) then electrons flow from B to A in the formation of AB. Assuming there are no other complicating factors it can be shown from the definition of μ and η that when electron transfer (ΔN) is small,

$$E_A = E_A^o + \mu_A^o(N_A - N_A^o) + \frac{1}{2}\eta_A(N_A - N_A^o)^2 + \ldots \tag{1.23a}$$

$$E_B = E_B^o + \mu_B^o(N_B - N_B^o) + \frac{1}{2}\eta_B(N_B - N_B^o)^2 + ... \quad (1.23b)$$

[here, terms from third order onwards are neglected and it is assumed that $\left(\frac{\partial^2 E}{\partial N^2}\right)_{v(\bar{r})} = \eta$]. Now ignoring all other effects, the total energy can be written as,

$$E_A + E_B = E_A^o + E_B^o + (\mu_A^o - \mu_B^o)\Delta N + \frac{1}{2}(\eta_A + \eta_B)(\Delta N)^2 + ...$$

or, $$(E_A + E_B) - (E_A^o + E_B^o) = \Delta E_A + \Delta E_B$$

$$= \Delta(E_A + E_B) = (\mu_A^o - \mu_B^o)\Delta N + \frac{1}{2}(\eta_A + \eta_B)(\Delta N)^2 + ... \quad (1.24)$$

where, $$\Delta N = N_B^o - N_B = N_A - N_A^o \quad (1.25)$$

Thus, when $\mu_B^o > \mu_A^o$; a positive ΔN i.e., a flow of electron from B to A, will stabilize the system (particularly for small ΔN). Now electron transfer will be stopped when, $\frac{\Delta(E_A + E_B)}{\Delta N} = 0$. Hence, from Eq (1.24) one can write,

$$\frac{\Delta(E_A + E_B)}{\Delta N} = 0 = (\mu_A^o - \mu_B^o) + (\eta_A + \eta_B)\Delta N$$

or $$(\mu_A^o + \eta_A \Delta N) - (\mu_B^o - \eta_B \Delta N) = 0$$

or $$\mu_A = \mu_B \quad (1.26)$$

where, $$\mu_A = \left(\frac{\partial E_A}{\partial N_A}\right)_{v(\bar{r})} = \mu_A^o + \eta_A \Delta N \quad (1.27a)$$

$$\mu_B = \left(\frac{\partial E_B}{\partial N_B}\right)_{v(\bar{r})} = \mu_B^o - \eta_B \Delta N \quad (1.27b)$$

Hence, from Eqs (1.26), (1.27a) and (1.27b), we can write,

$$\Delta N = \frac{\mu_B^o - \mu_A^o}{(\eta_A + \eta_B)} \quad (1.28)$$

Substituting the values of ΔN from Eq (1.28) in Eq (1.24), we can write,

$$(E_A - E_A^o) + (E_B - E_B^o) = (\mu_A^o - \mu_B^o)\frac{\mu_B^o - \mu_A^o}{(\eta_A + \eta_B)} + \frac{1}{2}(\eta_A + \eta_B)\left(\frac{\mu_B^o - \mu_A^o}{(\eta_A + \eta_B)}\right)^2$$

or, $$\Delta E_{SE} = \Delta E_A + \Delta E_B = \Delta(E_A + E_B)$$

$$= -\frac{(\mu_B^o - \mu_A^o)^2}{(\eta_A + \eta_B)} + \frac{(\mu_B^o - \mu_A^o)^2}{2(\eta_A + \eta_B)} = -\frac{(\mu_B^o - \mu_A^o)^2}{2(\eta_A + \eta_B)}$$

or, $$\Delta E_{SE} = \Delta E_A + \Delta E_B = -\frac{(\mu_B^o - \mu_A^o)^2}{2(\eta_A + \eta_B)} \quad (1.29)$$

Here, Eq (1.29) represents the stabilization energy due to transfer of ΔN amount of electron from B to A (from Eq (1.29) it is obvious that ΔE_{SE} is negative i.e., energy is lowered due to charge transfer).[20,154]

Another global reactivity descriptor is global electrophilicity (w), also proposed by Parr *et al*.[155] while tried to validate the experimental findings of Maynard *et al*.[156] A model was used according to which, when electrophilic system (atom, molecule, or ion) immersed in an idealized zero-temperature free electron sea of zero chemical potential (eg., a protein or a DNA coil), there would be an electron flow of amount ΔN from the sea to the system until the chemical potential of the system becomes zero. The change in the electronic energy as a function of the change in the number of electrons, ΔN up to second order, at constant external potential $v(\bar{r})$ is

$$\Delta E = \mu \Delta N + \eta \frac{\Delta N^2}{2} \quad (1.30)$$

The saturation situation by soaking up the maximum amount of electrons, $\Delta N_{\max}$, of the system can be characterized by putting

$$\frac{\Delta E}{\Delta N} = 0 \quad (1.31)$$

implying

$$\Delta N_{\max} = -\frac{\mu}{\eta} \quad (1.32)$$

which yields stabilization energy,

$$\Delta E = -\frac{\mu^2}{2\eta} \quad (1.33)$$

In Eq (1.33), the numerator (μ^2) is quadratic and, hence, positive and the denominator (2η) is positive due to the convexity of the energy vs. N curve and hence, ΔE is

negative: charge transfer is an energetically favorable process. In view of the analogy between classical electricity ($power \equiv W = -\frac{V^2}{R}$), Parr *et al.*[155] defined $w = -\Delta E$ as a measure of electrophilicity of the system (atom, molecule, or ion). The resulting equation is

$$w = \frac{\mu^2}{2\eta} \tag{1.34}$$

This quantity w is called the "electrophilicity index". Kinetic and thermodynamic aspects of w were investigated by Chattaraj and collaborators[157] by correlating it with the relative experimental rates of different types of reactions. However, a thorough discussion, aided by analytical reasoning, on the thermodynamic and kinetic aspects of w were reported by Bagaria and Roy.[158] The 'thermodynamic' aspect helps to explain, qualitatively, favourable product formation. This aspect of w is established from the condition of maximal flow of electrons, i.e., when $\left(\frac{\Delta E}{\Delta N}\right)_{v(\vec{r})} = 0$, $\Delta E \approx -w = -\frac{\mu^2}{2\eta}$. As $\eta > 0$, $\Delta E < 0$, i.e., charge transfer is an energetically favorable process. The 'kinetic' aspect is used to describe the rate of the reaction. This can be realized from the expression of w (i.e., of Eq (1.34)) in terms of first vertical IP and first vertical EA as (by using Eqs (1.14) and (1.20)),

$$w = \frac{\mu^2}{2\eta} = \frac{\left[-(IP + EA)/2\right]^2}{2(IP - EA)} = \frac{(IP + EA)^2}{8(IP - EA)} \tag{1.35}$$

In a chemical reaction, where the substrate acts as an electron acceptor, it is expected that a substrate with higher *EA* value will enhance the rate of the reaction than that with a lower *EA*. Therefore, the rate of the reaction can be correlated with *EA* and hence with global electrophilicity (w) value. If the substrate is an electron acceptor then higher w value will favor the reaction and for electron donor substrate naturally the lower w value will favor the reaction leading to the lower activation energy (E_a), or free energy of activation ($\Delta G^{\#}$).

It also was reasoned[158] that the above correlation of global electrophilicity (w) with the activation energy is not justified for all types of reactions. Only for single-step

reactions is it safe to carry out such correlation. For multi-step reactions the overall rate depends on the rate-determining step in which the substrate may not be directly involved.

More recently, Bagaria *et al.*[159] extended the use of global electrophilicity descriptor, as proposed by Parr *et al.*,[155] to the system where the donor is not a perfect one and the acceptor is of comparable size to that of the donor (*viz*, when both are ordinary organic molecules). It was then proposed that the energy fragments (generated after decomposing the stabilization energy, i.e., $\left|\Delta E_{A(B)}\right|$ and $\left|\Delta E_{B(A)}\right|$) together with the global electrophilicity descriptor of the acceptor (w_A), could explain the rate determining step of a multistep chemical reaction.[159] They also showed that Eq (1.33) is a special case of Eq (1.29), when both μ_B^0 and η_B are assumed to be zero in case of a idealized donor (normally very large biological systems, e.g., DNA-coil, protein).

Several other global reactivity descriptors e.g., nucleophilicity,[160-164] electrofugality and nucleofugality,[165-167] potentialphilicity and potentialphobicity,[168] chargephilicity and chargephobicity[169] have been also proposed recently, which are all conceptually related to w.

(ii) *Local Reactivity Descriptors*: Parallel to the development of global reactivity descriptors, some local reactivity descriptors have also been proposed which have potential use in predicting local (site) reactivity (selectivity) of a chemical species. Local properties may vary from point to point in space and are one-point ($\bar{r}$) functions, and so are the key in determining which sites in a molecule are most reactive.

If a change from one ground state to another is considered then one finds the fundamental equation for the change in $E[N, v(\bar{r})]$ as,

$$dE = \left(\frac{\partial E}{\partial N}\right)_{v(\bar{r})} dN + \int \left(\frac{\delta E}{\delta v(\bar{r})}\right)_N \delta v(\bar{r}) d\bar{r}$$

or,
$$dE = \mu dN + \int \left(\frac{\delta E}{\delta v(\bar{r})}\right)_N \delta v(\bar{r}) d\bar{r} \qquad (1.36)$$

thus, one has the most fundamental local reactivity descriptor, the ground state electron density $\rho(\bar{r})$ [170-173,31,174-176]

$$\rho(\bar{r}) = \left(\frac{\delta E}{\delta v(\bar{r})} \right)_N \tag{1.37}$$

Similarly, the change in chemical potential associated with a change in N and/or $v(\bar{r})$ is given by the formula

$$d\mu = \left(\frac{\partial \mu}{\partial N} \right)_{v(\bar{r})} dN + \int \left(\frac{\delta \mu}{\delta v(\bar{r})} \right)_N \delta v(\bar{r}) d\bar{r} \tag{1.38a}$$

Or, introducing the symbol $\eta = \left(\frac{\partial \mu}{\partial N} \right)_{v(\bar{r})}$

$$d\mu = \eta dN + \int \left(\frac{\delta \mu}{\delta v(\bar{r})} \right)_N \delta v(\bar{r}) d\bar{r} \tag{1.38b}$$

and we arrive another important space-dependent (local) derivative of chemical potential,

$$f(\bar{r}) = \left(\frac{\delta \mu}{\delta v(\bar{r})} \right)_N = \left(\frac{\partial \rho(\bar{r})}{\partial N} \right)_{v(\bar{r})} \tag{1.39}$$

which is known as Fukui Function (FF).[177,178,141,179] This quantity integrates to unity, $\int f(\bar{r}) d\bar{r} = 1$. The second formula for $f(\bar{r})$ in Eq (1.39) is a "Maxwell relation"[180] following from the fact that dE is an exact differential. There is a discontinuity[177,181,182] in the derivative of the Fukui function just as there is for chemical potential.[183] When an electron is being added, one has $f^+(\bar{r})$; when it is being subtracted one has $f^-(\bar{r})$; one also has the average $f^0(\bar{r})$. Parr and Yang[177] have defined the left ($f^-(\bar{r})$), right ($f^+(\bar{r})$) and central ($f^0(\bar{r})$) derivatives of Eq (1.39). These three Fukui functions can be written by applying a finite difference approximation and the frontier-orbital theory[3-6] of reactivity as,

$f^-(\bar{r}) \cong \rho_N(\bar{r}) - \rho_{N-1}(\bar{r}) \approx \rho_{HOMO}(\bar{r})$ measures reactivity toward an electrophilic (El^+) reagent (derivative as N increases from $N - \delta \to N$), (1.40)

$f^{+}(\bar{r}) \cong \rho_{N+1}(\bar{r}) - \rho_{N}(\bar{r}) \approx \rho_{LUMO}(\bar{r})$ measures reactivity toward a nucleophilic (Nu^{-}) reagent (derivative as N increases from $N \to N+\delta$), (1.41)

and

$$f^{0}(\bar{r}) = \frac{1}{2}\left[f^{+}(\bar{r}) + f^{-}(\bar{r})\right] \cong \frac{1}{2}[\rho_{N+1}(\bar{r}) - \rho_{N-1}(\bar{r})] \approx \frac{1}{2}\left[\rho_{HOMO}(\bar{r}) + \rho_{LUMO}(\bar{r})\right]$$

measures reactivity toward an innocuous (radical) reagent (mean of left and right derivatives) (1.42)

where, $\rho_{N}(\bar{r})$, $\rho_{N-1}(\bar{r})$ and $\rho_{N+1}(\bar{r})$ represent the electron density at a point $\bar{r}$ for the N, $N-1$ and $N+1$ electron system, respectively.

As chemists are interested with reactivities of atomic sites in reactions involving neutral systems and their monopositive and mononegative ions (i.e., when the electron number is changing by 1, instead of an infinitesimally small amount, δ), it would be more useful, albeit approximate, if $f(\bar{r})$ indices of an atom in a molecule could be evaluated. Yang and Mortier[184] proposed such approximate atomic $f(\bar{r})$ indices (or condensed-to-atom Fukui functions) applying finite difference approximation to the condensed electronic population on any atom (say for atom k) as

$$f^{-}(k) \cong P_{N}(k) - P_{N-1}(k) \quad (1.43)$$

$$f^{+}(k) \cong P_{N+1}(k) - P_{N}(k) \quad (1.44)$$

and

$$f^{0}(k) \cong \frac{1}{2}[P_{N+1}(k) - P_{N-1}(k)] \quad (1.45)$$

where $P(k)$ denotes the electronic population on atom k. Parr and Yang[177] proposed that larger value of Fukui function indicates more reactivity. Hence, greater the value of the condensed Fukui function, the more reactive is the particular atomic center in the molecule.

Moreover, one of the often-cited problems with Fukui function is that of its negative values.[185-200] A negative Fukui function value means that when adding an electron to the molecule, in some spots the electron density is reduced (i.e., for nucleophilic attack). Alternatively when removing an electron from the molecule, in some spots the electron density grows larger (i.e., for electrophilic attack). If Fukui function indices are expected to be positive values, then the above equalities should not occur, which is unreasonable and also has yet not been formally shown whether such behavior is physically correct or not. But it has been emphasized that Fukui function should be normalized,[45] i.e., they should sum to one.

To treat the problem regarding the negative Fukui function, Hirshfeld population analysis (HPA)[201] (also known as stockholders charge-partitioning technique), as proposed by Hirshfeld, was used and shown that HPA yields only positive Fukui functions.[185,59,189,160,190,202,192] Also, it was shown that electronic population derived on the basis of HPA produces more reliable intramolecular reactivity trends when compared to those obtained from Mulliken population analysis (MPA),[203] natural bond orbital (NBO) analysis,[204-207] and molecular electrostatic potential (MESP) based methods.[9] Even though it is difficult to evaluate the superiority of one method to the others, studies by Roy *et al.*[185,189,190] clearly demonstrated that HPA is superior to other charge-partitioning schemes. Subsequently, there have been quite a number of studies in this area,[187,208,192] which have also analytically shown that HPA is a superior charge-partitioning scheme because it suffers from minimum missing information when atoms form a molecule.[186,209,187,188,202,191,210,192-194,196-200] But in this HPA technique also, there is no formal prescription for evaluating atomic charges (i.e., q_k) in the corresponding ionic species. Also, what would be the weight factors '$w_k(\bar{r})$' for the atoms in the corresponding ionic species is not clearly outlined. In the first study in this series Roy *et al.*,[185] have shown that condensed Fukui function can be positive only when same weight factor for the neutral, cationic and anionic species is considered. It is true that such an approximation is crude and not a generalized method.

In order to mitigate the problems associated with the above Hirshfeld scheme, in 2007, Bultinck *et al.*[211] have proposed an alternative, iterative version of the Hirshfeld

partitioning procedure, known as "Hirshfeld-I" method. They have verified this method on a test set of 168 molecules containing C, H, N, O, F and Cl atoms. On the basis of this study, they ensure that this iterative scheme (i) eliminates arbitrariness in the choice of the promolecule, so the atomic populations are determined solely by the molecular electron density, (ii) increases the magnitudes of the charges, and (iii) also treats open shell species without problem. But right now, it is difficult to comment on its universal validity, as this method has yet not been used much by other researchers working in this area. However, it has been recognized that HPA is trustworthy[212] as long as small atoms (especially hydrogen atoms) are not embedded in regions with substantial negative or positive deformation densities. It also seems that HPA is rather trustworthy when "large" changes in atomic charge (on the order of a tenth of the charge on the electron) are of interest and less trustworthy when small nuances are being studied. For systems that fail to meet these criteria, alternative population analysis schemes should be considered.

If negative Fukui function indices even occur at equilibrium geometries, then the molecule would be expected to have very interesting magnetic and redox properties.[213-215] This is important in view of the fact that although the problem of negative Fukui function indices has been looked into in detail, no definitive answer has been given yet to the question whether negative values are physically acceptable or are artifacts. According to some computational studies, it is truly impossible to exclude negative Fukui function.[216,187,217-219] Further, it has been pointed out that the possibility of negative atom condensed Fukui function values depend critically on the properties of the hardness matrix.[202,220,193,213]

In any case $f(\bar{r})$ is established as an index of considerable importance for understanding molecular behaviour – the natural reactivity index of density functional theory. Note that $f(\bar{r})$ is defined independently of any model, while the concepts of classical frontier theories are framed in the language of the independent-particle model.

The Fukui function is a powerful local reactivity indicator for regioselectivity but it is not expected to provide an accurate indication of the overall reactivity of a molecule. When a reactivity indicator that reflects overall reactivity is needed, workers in Conceptual DFT usually work in the grand canonical ensemble.[221] Reactivity descriptors

in the grand canonical ensemble are obtained by replacing derivatives with respect to the number of electrons, N, with derivatives with respect to the electronic chemical potential, μ (the electronic chemical potential measures the intrinsic strength of Lewis acids and bases, so reactivity descriptors in the grand canonical ensemble represent how a molecule's reactivity changes as its electron-withdrawing power or electronegativity decreases). In the grand canonical ensemble, the Fukui function, $f(\bar{r})$, is replaced by the local softness, $s(\bar{r})$ [141]

$$s(\bar{r}) = \left(\frac{\partial \rho(\bar{r})}{\partial \mu}\right)_{v(\bar{r})} = \left(\frac{\partial \rho(\bar{r})}{\partial N}\right)_{v(\bar{r})} \left(\frac{\partial N}{\partial \mu}\right)_{v(\bar{r})} = f(\bar{r})S \qquad (1.46)$$

where S is global softness (vide Eq (1.22)).

Thus local softness in such a reactivity parameter which describes the response of any particular site of a chemical species (in terms of change in electron density $\rho(\bar{r})$) to any global change in its chemical potential values. The parameter $s(\bar{r})$ obeys the condition,

$$\int s(\bar{r})d\bar{r} = S \qquad (1.47)$$

The Fukui function in Eq (1.46) can be identified with the Fukui function from above (Eq (1.40)), the Fukui function from below (Eq (1.41)), or from the average of the two (Eq (1.42)). Similarly, the three approximate atomic $f(\bar{r})$ indices (from Eqs (1.43)-(1.45)), when multiplied by S, provide three different local softnesses for any particular atom (k). These can be written as

$$s^{-}(k) \cong \left[P_N(k) - P_{N-1}(k)\right]S \qquad \text{(suited for studies of electrophilic attack)} \qquad (1.48)$$

$$s^{+}(k) \cong \left[P_{N+1}(k) - P_N(k)\right]S \qquad \text{(suited for studies of nucleophilic attack)} \qquad (1.49)$$

and

$$s^{0}(k) \cong \frac{1}{2}\left[P_{N+1}(k) - P_{N-1}(k)\right]S \qquad \text{(suited for studies of radical attack)} \qquad (1.50)$$

From Eq (1.46) it is obvious that local softness contains the same information as Fukui function plus additional information about the total molecular softness. The Fukui function may be thought of as a normalized local softness.[141] Therefore, either the Fukui function or local softness can be used in the studies of intramolecular reactivity sequences (i.e., relative site reactivity in a molecule).[222] But only $s(\bar{r})$ (and not $f(\bar{r})$) should be a better descriptor of the global reactivity with respect to a reaction partner having a given hardness (or softness), as stated in the HSAB principle.[18]

There is an interesting fluctuation formula for this quantity in finite-temperature DFT, where the averages are over all members of a grand ensemble at temperature T.[141] This formula and other similar DFT fluctuation formulae[223,224] may provide a basis for fluctuation theories of catalysis. $s(\bar{r})$ is measurable using scanning tunnel microscopy. For an infinite system, $s(\bar{r})$ is approximately the local density of states at the Fermi level and S the total density of states at the Fermi level.[141,225]

It has been argued that the individual values of s_k^+ and s_k^- are strongly influenced by the basis set or correlation effects. But the ratio of s_k^+ and s_k^-, involving two differences of electron densities of the same system differing by one in their number of electrons, at constant nuclear framework, are expected to be less sensitive to the basis set and correlation effects. Based on this argument, Roy *et al*.[226] introduced two new reactivity descriptors to find out the preferable reactive sites. These are defined as relative electrophilicity (s_k^+ / s_k^-) and relative nucleophilicity (s_k^- / s_k^+) of any particular atom k, and helps to locate the preferable site (or atom) in a molecule for nucleophilic and electrophilic attack on it, respectively. That is, relative nucleophilicity is the nucleophilicity of any site as compared to its own electrophilicity and relative electrophilicity is the electrophilicity of any site as compared to its own nucleophilicity.

There is no unique simple inverse of $s(\bar{r})$. Berkowitz and Parr[227] have given a derivation of local softness that reveals its relation to its reciprocal property, local hardness.[228-230]

Substitution of Eq (1.4) into Eq (1.2) follows, for a ground state

$$E[\rho] = N\mu - \left[\int \frac{\delta F[\rho]}{\delta\rho(\bar{r})}\rho(\bar{r})d\bar{r} - F[\rho]\right]$$

$$= N\mu - H[\rho] \tag{1.51}$$

where the *hardness functional* $H[\rho]$ is defined by the formula[231]

$$H[\rho] = \int \frac{\delta F[\rho]}{\delta\rho(\bar{r})}\rho(\bar{r})d\bar{r} - F[\rho] \tag{1.52}$$

$H[\rho]$ is what must be added to E to give $N\mu$. Note that a leading term in $H[\rho]$ is $J[\rho]$, the classical part of $V_{ee}[\rho]$.

The total differential of Eq (1.51), associated with a change of the system of interest, is simply

$$dE[\rho] = Nd\mu + \mu dN - dH[\rho] \tag{1.53}$$

Comparing this with Eq (1.36), one can obtain the equation

$$Nd\mu = dH[\rho] + \int \rho(\bar{r})\delta v(\bar{r})d\bar{r} \tag{1.54}$$

The differential of $H[\rho]$ has a surprisingly simple form. From Eq (1.52) one has[42]

$$dH[\rho] = -dF[\rho] + d\left[\int \frac{\delta F[\rho]}{\delta\rho(\bar{r})}\rho(\bar{r})d\bar{r}\right]$$

$$= -\int \frac{\delta F[\rho]}{\delta\rho(\bar{r})}d\rho(\bar{r})d\bar{r} + \int \frac{\delta F[\rho]}{\delta\rho(\bar{r})}d\rho(\bar{r})d\bar{r} + \int\left[d\left(\frac{\delta F[\rho]}{\delta\rho(\bar{r})}\right)\right]\rho(\bar{r})d\bar{r}$$

$$= \iint \frac{\delta^2 F[\rho]}{\delta\rho(\bar{r})\delta\rho(\bar{r}')}d\rho(\bar{r}')\rho(\bar{r})d\bar{r}d\bar{r}' \tag{1.55}$$

So, from Eq (1.53), $$d\left(E[\rho] - N\mu\right) = -\iint \frac{\delta^2 F[\rho]}{\delta\rho(\bar{r})\delta\rho(\bar{r}')}d\rho(\bar{r}')\rho(\bar{r})d\bar{r}d\bar{r}' \tag{1.56}$$

Following Ghosh and Berkowitz,[229] local hardness $\eta(\bar{r})$ is defined as:[228]

$$\eta(\bar{r}) = \frac{1}{N}\int \frac{\delta^2 F}{\delta\rho(\bar{r}')\delta\rho(\bar{r})}\rho(\bar{r}')d\bar{r}' \tag{1.57}$$

One can find accurate to all orders,

$$\frac{\delta H[\rho]}{\delta\rho(\bar{r})} = N\eta(\bar{r}) = h(\bar{r}) \tag{1.58}$$

Parr and Gázquez called $h(\bar{r})$ the hardness potential for the system.[231]

Introducing the symbol of $\eta(\bar{r})$, one can rewrite Eq (1.56) as

$$d\left(E[\rho]-N\mu\right)=-N\int\eta(\bar{r})d\rho(\bar{r})d\bar{r} \tag{1.59}$$

Eliminating dE from Eqs (1.36) and (1.59), was obtained

$$d\mu=\int\eta(\bar{r})d\rho(\bar{r})d\bar{r}+\frac{1}{N}\int\rho(\bar{r})\delta v(\bar{r})d\bar{r} \tag{1.60}$$

This is the local counterpart of Eq (1.38) in the sense of Nalewajski,[44] in which it now appears on the local hardness in place of the global hardness.

From Eq (1.60) one can find another formula for $\eta(\bar{r})$,[228]

$$\eta(\bar{r})=\left(\frac{\delta\mu}{\delta\rho(\bar{r})}\right)_{v(\bar{r})} \tag{1.61}$$

Eq (1.61) is an example of an ambiguous "constrained functional derivative".[228,232,223,233-241,242-249] The functional derivative is ambiguous because of the interdependence of $\rho(\bar{r})$ and $v(\bar{r})$.[223] It is interesting to note that the local hardness also appears in a natural way when the chain rule is applied to the global hardness:

$$\eta=\left(\frac{\partial^2E}{\partial N^2}\right)_{v(\bar{r})}=\left(\frac{\partial\mu}{\partial N}\right)_{v(\bar{r})}=\int\left(\frac{\delta\mu}{\delta\rho(\bar{r})}\right)_{v(\bar{r})}\left(\frac{\delta\rho(\bar{r})}{\delta N}\right)_{v(\bar{r})}d\bar{r}$$

$$=\int\eta(\bar{r})f(\bar{r})d\bar{r} \tag{1.62}$$

An explicit expression for $\eta(\bar{r})$ can be deduced from Eq (1.57) and the definition of Fukui function (Eq (1.39))[232]

$$\eta(\bar{r})=\int\frac{\delta^2F}{\delta\rho(\bar{r}')\delta\rho(\bar{r})}f(\bar{r}')d\bar{r}' \tag{1.63}$$

Local hardness and local softness are reciprocals in the sense that

$$\int\eta(\bar{r})s(\bar{r})d\bar{r}=1 \tag{1.64}$$

One can simplify the definitions of local hardness by writing[234]

$$\eta_\lambda(\bar{r})=\left(\frac{\delta\mu}{\delta\rho(\bar{r})}\right)_{v(\bar{r})}=\frac{1}{N}\int\frac{\delta^2F}{\delta\rho(\bar{r}')\delta\rho(\bar{r})}\lambda[\rho(\bar{r}')]d\bar{r}' \tag{1.65}$$

where, $\lambda[\rho(\bar{r}')]$[223] is a composite function that integrates to N (i.e., total number of electrons of the system),

$$\int \lambda[\rho(\bar{r}')]d\bar{r}' = N \tag{1.66}$$

Two important choices of the composite function $\lambda[\rho(\bar{r}')]$ are $\rho(\bar{r}')$ [229,234,243,244,246,247,250-252] and $Nf(\bar{r}')$, [232,235,237,245,239,248,241,249] when the following possibilities emerge:

$$\lambda[\rho(\bar{r}')] = \rho(\bar{r}') \qquad \text{yielding} \qquad \tilde{\eta}_D(\bar{r}) = \frac{1}{N}\int \frac{\delta^2 F}{\delta\rho(\bar{r}')\delta\rho(\bar{r})}\rho(\bar{r}')d\bar{r}' \tag{1.67}$$

and $$\lambda[\rho(\bar{r}')] = Nf(\bar{r}') \qquad \text{yielding} \qquad \tilde{\eta}_F(\bar{r}) = \int \frac{\delta^2 F}{\delta\rho(\bar{r}')\delta\rho(\bar{r})}f(\bar{r}')d\bar{r}' \tag{1.68}$$

But $\tilde{\eta}_F(\bar{r})$ is shown to be equal to the global hardness η at every point of space,[223] when the exact functional $F[\rho]$ is used[253] in Eq (1.68)

$$\eta = \left(\frac{\partial\mu}{\partial N}\right)_{v(\bar{r})} = \frac{\partial}{\partial N}\left(\frac{\delta F[\rho]}{\delta\rho(\bar{r})}\right) = \int \frac{\delta^2 F[\rho]}{\delta\rho(\bar{r})\delta\rho(\bar{r}')}\left(\frac{\partial\rho(\bar{r}')}{\partial N}\right)d\bar{r}'$$

$$= \int \frac{\delta^2 F[\rho]}{\delta\rho(r)\delta\rho(\bar{r}')}f(\bar{r}')d\bar{r}' = \tilde{\eta}_F(\bar{r}) \tag{1.69}$$

At first sight this form seems to be less appropriate, as "*unlike the chemical potential there is nothing in the concept of hardness which prevents it from having different values in different parts of the molecule*".[138]

Based on the global electrophilicity index w (Eq (1.34)) as defined by Parr *et al*.,[155] Pérez *et al*.[254] introduced an useful expression for the local electrophilicity index $w(k)$ in terms of the electrophilic Fukui function and local softness. From Eq (1.34) and using the inverse relationship between global hardness and global softness[141] (Eq (1.22)) one may obtain

$$w = \frac{\mu^2}{2\eta} = \frac{\mu^2}{2}S = \frac{\mu^2}{2}\sum_k s^+(k) = \sum_k w(k) \tag{1.70}$$

Afterward, Chattaraj *et al*. proposed a broader and general local reactivity descriptor by using the resolution of identity.[255] This is named as the "philicity" index $w(\bar{r})$, [255-258] which encompasses all types of reactions (i.e., electrophilic, nucleophilic, and radical reactions). This local philicity $w(\bar{r})$ is promised to be a more powerful quantity than global reactivity descriptors because the former contains the information of the latter in addition to the site selectivity of a molecule toward electrophilic, nucleophilic, and

radical attacks. Also, according to the argument of the authors, "because the global electrophiliciy of two different molecules are different, best sites of two different molecules for a given reaction can be explained only in terms of the 'philicity' and not Fukui function". So, they proposed the existence of a local electrophilicity index ($w(\bar{r})$) that varies from point to point in an atom, molecule, ion or solid and is defined as

$$w = \int w(\bar{r})d\bar{r} \tag{1.71}$$

By using the resolution of identity as represented by $\int f(\bar{r})d\bar{r} = 1$, the best choice of $w(\bar{r})$ was proposed to be

$$w = w\int f(\bar{r})d\bar{r} = \int wf(\bar{r})d\bar{r} = \int w(\bar{r})d\bar{r} \tag{1.72}$$

where $$w(\bar{r}) = wf(\bar{r}) \tag{1.73}$$

To take care of all types of reactions three different forms of $w(\bar{r})$ was defined as

$$w^{\alpha}(\bar{r}) = wf^{\alpha}(\bar{r}) \tag{1.74}$$

where $\alpha =$ +, -, and 0 for attacks by a nucleophile, electrophile, and radical, respectively. It is obvious that Eq (1.73), when integrated, generates w, i.e., the global electrophilicity. This is true for $\alpha =$ +, -, and 0. However, in the presence of a physicochemical perturbation, some particular atom (or atoms) is (are) better equipped toward electrophilic (or nucleophilic) attack on it. As $w^{\alpha}(\bar{r})$ takes care of all types of reactions, it is claimed to be more general and is called the local philicity index. The corresponding condensed-to-atom forms of the philicity index for atom *k* is written as

$$w^{\alpha}(k) = wf^{\alpha}(k) \tag{1.75}$$

In a study by Roy,[259] it has been shown that the philicity index $w(\bar{r})$ and the local softness $s(\bar{r})$ generate identical intramolecular reactivity (or site selectivity) trends. This is because $w(\bar{r})$ and $s(\bar{r})$ are analytically related as follows:

$$w(\bar{r}) = wf(\bar{r}) = \frac{\mu^2}{2\eta} f(\bar{r}) = \mu^2 Sf(\bar{r}) = \mu^2 s(\bar{r}) \tag{1.76}$$

That is, $w(\bar{r})$ can be obtained after multiplying the $s(\bar{r})$ by μ^2 which is constant for a particular system but varies from system to system. Therefore, it has been concluded that $w(\bar{r})$ will not provide any extra information than that of $s(\bar{r})$ or $f(\bar{r})$ on intramolecular reactivity trends. It may be noted that Chattaraj himself also later on mentioned that for

intramolecular reactivity, philicity, local softness and FF furnished the same trend.[260] Roy *et al*.,[261] in one interesting study made a significant revelation regarding the correlation between global and local reactivity descriptors. It was argued that the claim [i.e., global trend of electrophilicity (or nucleophilicity) originates from the local behavior of the molecules, or precisely of that atomic site which is most prone to electrophilic (or nucleophilic) attack] is logical for systems having only one distinctly strong site (electrophilic or nucleophilic) but does not hold true for systems having more than one site of comparable strength. For the justification of this argument, a thorough study was carried out by Roy *et al*.,[261,262] using numerical demonstrations and analytical reasoning. Finally, it was concluded that reliable intermolecular reactivity trend can be generated by global electrophilicity (or may be local hardness) and that is possible with local electrophilicity only for the systems having one distinctly strong site. In another interesting article Ayers *et al*.,[221] have discussed the 'extensive', 'intensive' and 'subintensive' nature of DFT based reactivity descriptors.

(iii) ***Nonlocal Reactivity Descriptors:*** These are reactivity descriptors which depends on two or more spatial positions, $\bar{r},\bar{r}'$, etc. Interest in these reactivity descriptors originates from the fact that local descriptors are defined as responses to a global perturbation, whereas the chemical reaction is typically local. In the detailed consideration of a change of any chemical system from one ground state to another, or in the determination of a ground state by any trial and error process in which ρ is guessed repeatedly, it has been recognized that nonlocal quantities play an important role.[227]

If we consider the ground state of a system of interest which changes only from one ground state to another, $\rho(\bar{r})$ determines everything by the original Hohenberg-Kohn theorems[29] including μ and $v(\bar{r})$. It therefore determines the modified potential $u(\bar{r})$ as (Eq (1.4)),

$$u(\bar{r}) = v(\bar{r}) - \mu = -\frac{\delta F[\rho]}{\delta \rho(\bar{r})} \tag{1.77}$$

where, $u(\bar{r})$ is also a functional of $\rho(\bar{r})$. The functional derivative of $u(\bar{r})$ with respect to $\rho(\bar{r}')$ therefore exists. This defines the hardness kernel, $\eta(\bar{r},\bar{r}')$ [263,44,141,227,225]

$$\eta(\bar{r},\bar{r}') \equiv -\frac{\delta u(\bar{r}')}{\delta\rho(\bar{r})} = -\frac{\delta u(\bar{r})}{\delta\rho(\bar{r}')} = \frac{\delta^2 F_E}{\delta\rho(\bar{r}')\delta\rho(\bar{r})} \tag{1.78}$$

the last equality coming from Eq (1.77). Recal the first definition of local hardness, $\eta(\bar{r})$ (Eq (1.57)), by introducing the symbol of $\eta(\bar{r},\bar{r}')$, one can find

$$\eta(\bar{r}) \equiv \frac{1}{N}\int \eta(\bar{r},\bar{r}')\rho(\bar{r}')d\bar{r}' \tag{1.79}$$

Similarly, another fact is that $u(\bar{r})$ determines all properties- not only $v(\bar{r})$ but also N, and hence $\rho(\bar{r})$. The functional derivative of $\rho(\bar{r})$ with respect to $u(\bar{r})$ therefore exists. This defines the softness kernel $s(\bar{r},\bar{r}')$[263,44,141,227,225]

$$s(\bar{r},\bar{r}') \equiv -\frac{\delta\rho(\bar{r}')}{\delta u(\bar{r})} = -\frac{\delta\rho(\bar{r})}{\delta u(\bar{r}')} \tag{1.80}$$

and the local softness $s(\bar{r})$

$$s(\bar{r}) \equiv \int s(\bar{r},\bar{r}')d\bar{r}' \tag{1.81}$$

Moreover, since both the functional derivatives exist,

$$\int \frac{\delta\rho(\bar{r})}{\delta u(\bar{r}')}\frac{\delta u(\bar{r}')}{\delta\rho(\bar{r}'')}d\bar{r}' = \delta(\bar{r}''-\bar{r}) \tag{1.82}$$

so that

$$\int s(\bar{r},\bar{r}')\eta(\bar{r}',\bar{r}'')d\bar{r}' = \delta(\bar{r}''-\bar{r}) \tag{1.83}$$

the hardness and softness kernels are true inverses.

Multiplying Eq (1.83) by $\rho(\bar{r}'')$ then integrating over $\bar{r}''$ and making use of Eq (1.79) one can write,

$$\int s(\bar{r},\bar{r}')\eta(\bar{r}')d\bar{r}' = \frac{\rho(\bar{r})}{N} \tag{1.84}$$

Integrating this over $\bar{r}$, and employing Eq (1.81), gives Eq (1.64)

$$\int \eta(\bar{r})s(\bar{r})d\bar{r} = 1 \tag{1.64}$$

To achieve Eq (1.46), writing

$$d\rho(\bar{r}) = \int \frac{\delta\rho(\bar{r})}{\delta u(\bar{r}')}du(\bar{r}')d\bar{r}'$$

$$= -\int s(\bar{r},\bar{r}')du(\bar{r}')d\bar{r}' \qquad \text{(by using Eq (1.80))}$$

$$= -\int s(\bar{r},\bar{r}')[\delta v(\bar{r}') - d\mu]d\bar{r}' \quad \text{(applying Eq (1.77))}$$

$$= [\int s(\bar{r},\bar{r}')d\bar{r}']d\mu - \int s(\bar{r},\bar{r}')\delta v(\bar{r}')d\bar{r}' \tag{1.85}$$

utilizing $s(\bar{r}) \equiv \int s(\bar{r},\bar{r}')d\bar{r}'$ from Eqs (1.38b) and (1.39) we get,

$$d\mu = \eta dN + \int f(\bar{r}')\delta v(\bar{r}')d\bar{r}' \tag{1.38c}$$

generating

$$d\rho(\bar{r}) = s(\bar{r})\eta dN + \int s(\bar{r})f(\bar{r}')\delta v(\bar{r}')d\bar{r}' - \int s(\bar{r},\bar{r}')\delta v(\bar{r}')d\bar{r}'$$

or,
$$d\rho(\bar{r}) = s(\bar{r})\eta dN + \int [-s(\bar{r},\bar{r}') + s(\bar{r})f(\bar{r}')]\delta v(\bar{r}')d\bar{r}' \tag{1.86}$$

Also, $\rho = \rho[N,v]$ implies

$$d\rho(\bar{r}) = f(\bar{r})dN + \int \left[\frac{\delta\rho(\bar{r})}{\delta v(\bar{r}')}\right]_N \delta v(\bar{r}')d\bar{r}' \tag{1.87}$$

N and $v(\bar{r})$ being independent, coefficients of dN and $\delta v(\bar{r})$ in Eqs (1.86) and (1.87) must be equal. Consequently, we have

$$s(\bar{r}) = \frac{f(\bar{r})}{\eta} = f(\bar{r})S = \left(\frac{\partial\rho(\bar{r})}{\partial N}\right)_{v(\bar{r})}\left(\frac{\partial N}{\partial \mu}\right)_{v(\bar{r})} = \left(\frac{\partial\rho(\bar{r})}{\partial \mu}\right)_{v(\bar{r})} \tag{1.46}$$

The derivative $\left[\frac{\delta\rho(\bar{r})}{\delta v(\bar{r}')}\right]_N$ is the conventional linear response function, denoted by $\chi(\bar{r},\bar{r}')$.[264,265] It is connected to the local softness, global softness and the softness kernel via an exact formula[227]

$$\chi(\bar{r},\bar{r}') = \left[\frac{\delta\rho(\bar{r})}{\delta v(\bar{r}')}\right]_N = -s(\bar{r},\bar{r}') + \frac{s(\bar{r})s(\bar{r}')}{S} \tag{1.88}$$

Eq (1.88) shows that the chemical reactivity, as measured by the softness kernel, is the sum of two contributions:[47] (i). the nonlocal response function of the system that contains contributions of all the MOs to the reactivity; and (ii) the electronic reactivity contained in the local softness, which is dominated by the frontier orbitals. This shows that the polarization changes in the electronic distribution (response to the external potential displacements) can be determined from the softness properties calculated for the fixed nuclear geometry (external potential).

C. Other Developments

Apart from the above developments of global, local and nonlocal reactivity descriptors in Conceptual DFT, some other parallel developments in the area are worth mentioning.

The defined reactivity and selectivity descriptors are inadequate to study the reactions which involve changes in spin multiplicity. For this purpose, the conceptual spin-polarized density functional theory (SP-DFT) was introduced by Galvan, Vela, and Gazquez.[266] This fact derives from the explicit consideration of the electron density and spin density (i.e., $\rho(\bar{r})$ and $\rho_S(\bar{r})$), respectively, written in terms of the spin-up $\rho_\alpha(\bar{r})$ and spin-down $\rho_\beta(\bar{r})$ components as

$$\rho(\bar{r}) = \rho_\alpha(\bar{r}) + \rho_\beta(\bar{r}) \tag{1.89}$$

and

$$\rho_S(\bar{r}) = \rho_\alpha(\bar{r}) - \rho_\beta(\bar{r}) \tag{1.90}$$

which integrates to the electron number, N, and spin number, N_S, respectively.

$$N = N_\alpha + N_\beta = \int \rho(\bar{r}) dr \tag{1.91}$$

$$N_S = N_\alpha - N_\beta = \int \rho_S(\bar{r}) dr \tag{1.92}$$

Spin-polarized DFT allows one to get some insight into the chemical properties related to the change in spin number. In recent years, many studies have appeared on the basis of which one can say that in some cases spin-polarization plays an important role.[266-295]

Here, so far we have placed emphasis on the effects of change of N and change of $v(\bar{r})$ on the *electron density*. The other elementary extension is shifts in the *nuclear* positions which must be incorporated in a complete theory.[296,224,225,297-305]

These reactivity indices have become very useful in predicting the regioselectivity of chemical reactions. The procedure and effort for determining the properties of large molecule are not same as those for calculation in small molecules. Thus could be an important area of further research. One can address the modeling techniques to take care of the extended systems. We now review some of the landmark works modeling the biological systems in the framework of Conceptual DFT. In the succeeding section we will discuss the regioselectivity for a number of large biological systems.

1.3. Literature Review

Realistic modeling of large systems is still out of reach. The bottleneck in exploring large chemical and biological systems is the calculation that is needed to be performed for the whole system. Thus, the larger the system becomes, the lower the level of calculation we have to opt for. Although the regioselectivity plays an important role in understanding a reaction –due to this bottleneck- very little conceptual DFT based works have been done involving large biological systems. However, one way to avoid this difficulty is to use some adequate modeling techniques. This section takes care of more recent developments enabling evaluations of the regio-chemistry for a number of large biological systems using several modeling schemes within the framework of Conceptual DFT.

In a very recent study, Barrientos-Salcedo *et al.*,[306] successfully reproduced experimental finding of reactivity segment as well as the reactive atomic sites of TP53 using DFT based reactivity descriptors. Therein, PNC-27 peptide derived amino acid sequence PPLSQETFSDLWKLL (aa 12-26)[307] was analyzed in three fragments: PPLSQ, ETFS, and DLWKLL (with carboxyl terminal ends in all cases). The chemical structure of the amino acids 12-26 (1Q2F, DOI: 10.2210/pdb1q2f/pdb)[307] was taken from the Protein Data Bank. The chemical structures of the three fragments studied in this work are shown in Figure 1.1, while the convention of atom-numbering for heavy atoms in this study is shown in Figure 1.2. They revealed[306] that PPLSQ, ETFS, and DLWKLL fragments studied, have important electrophilic sites such as Q16 (C71), D21 (C12), E17 (C17), P13 (C19), L26 (C103), S15 (C52), S20 (C53), L14 (C33), T18 (C18) and L25 (C82), suggesting that these amino acids are exposed to nucleophilic attacks on these atoms. Also, from the negative charge on nitrogen atoms such as Q16 (N76 and N59), K24 (N80), E17 (N1), D21 (N1), S20 (N42), and W23 (N23) and oxygen atoms S20 (O57), T18 (O24), S15 (O56), D21 (O15 and O16), and Q16 (O75), respectively, they observed[306] that these have larger negative charges as compared to the remainder of the atoms; therefore, electrophilic attacks might occur on these sites as well. These results are consistent with the experimental result of Kanovsky *et al*.,[308] and reinforce the proposal that the segment (17-20, ETFS) is essential for the biological effect. From the global reactivity descriptors values for the fragments, such as ionization potential (IP),

hardness (η) (i.e., $\eta = \frac{IP - EA}{2}$),[20] electrophilicity index (w) (Eq (1.34)),[155] and the spatial extent measured though the $\langle R^2 \rangle$, Barrientos-Salcedo *et al*. observed[306] that the ETFS fragment exhibits a larger value for ionization potential, which might be related to the greater global chemical stability of these fragments, i.e., larger ionization potential values may indicate smaller oxidative effects, and they observed that PPLSQ and DLWKLL show a decreasing potential order.

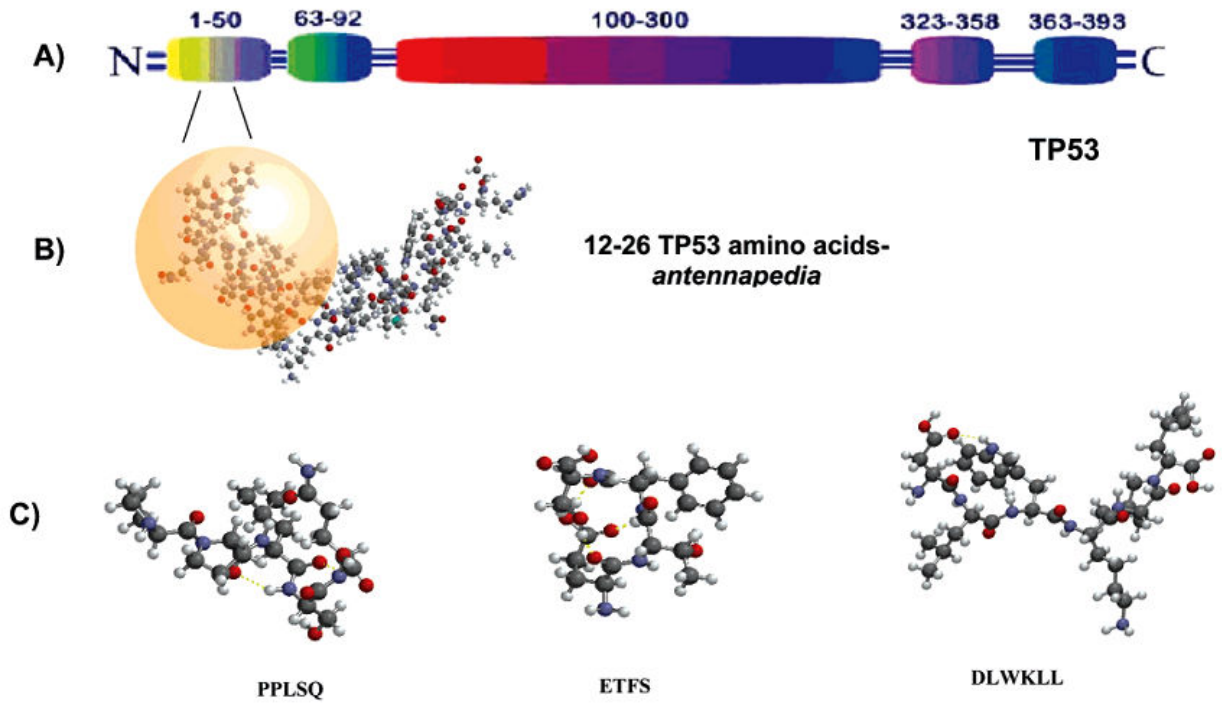


Figure 1.1: Images showing (A) TP53 protein, (B) amino acids 12-26 of TP53 protein (shaded circle) and penetratin and (C) fragments analyzed in ref 306 (ball and stick model).

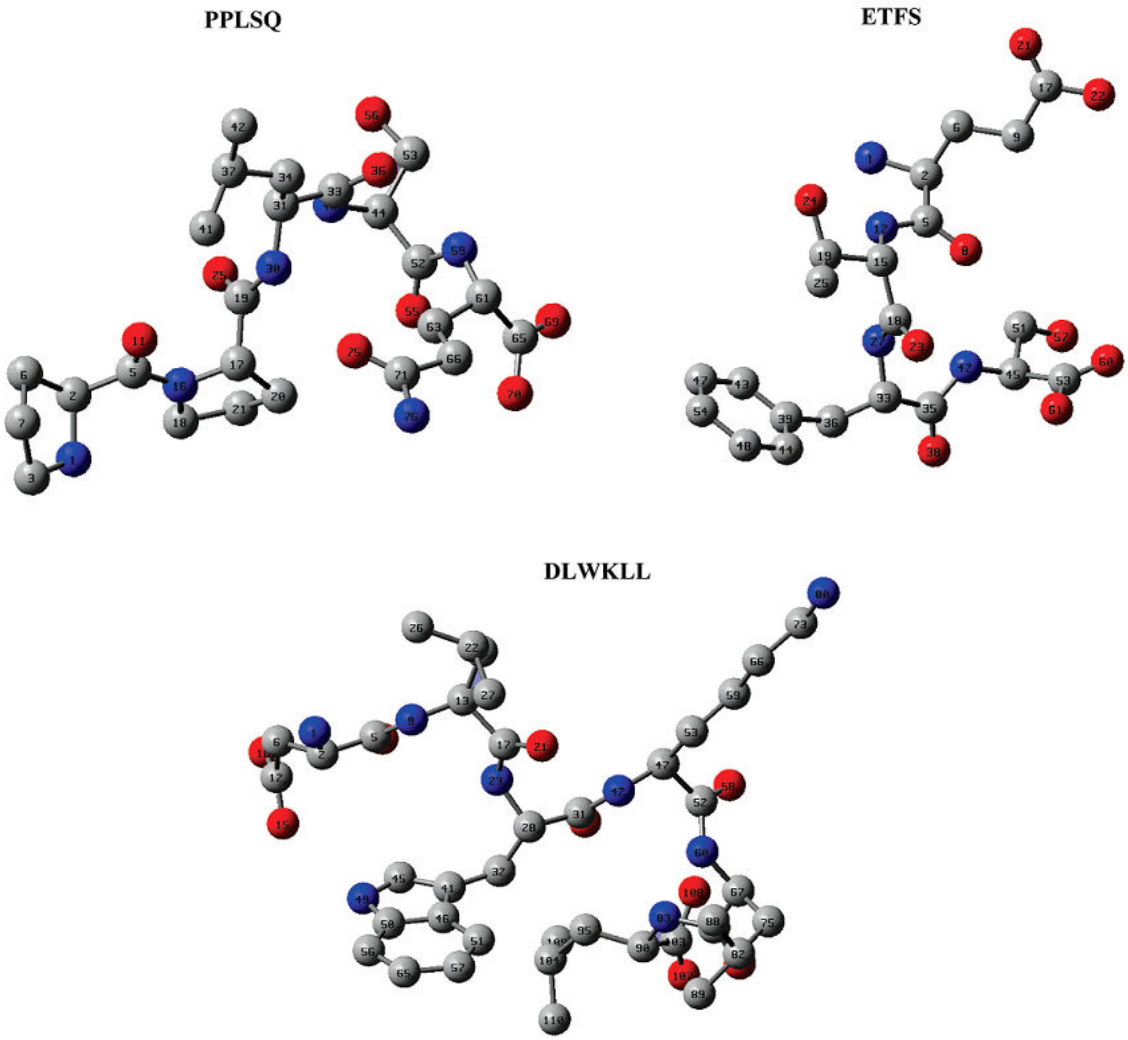


Figure 1.2: Convention used in numbering heavy atoms of atomic charges in ref 306 (ball and stick model).

On the basis of an energy perturbation method, Li and Evans[309,310] presented a slightly different formulation, indicating that, for a hard reaction, the site of minimal Fukui function[177,311,141,312,179,313] is preferred, whereas for a soft reaction, the site of maximal Fukui function is preferred. In a contribution, Li and Evans quantified the chemical reactivity of C3 of phosphoenolpyruvate (PEP) using charge and $f(\bar{r})$.[310] Conceptual DFT leads to hypothesis that this dual nature of C3 of PEP depends on the ionization state of PEP and on the value of the dihedral angle between the carboxylate and the C2-C3 double bond. The charge and $f(\bar{r})$ (the Fukui function is approximated by HOMO density divided by two[311]) of the C3 atom change when the conformation of the molecule varies.

The gas-phase proton affinity of the amino acids was investigated by Baeten *et al.*,[314] where electronegativities and hardnesses were determined for artificially constructed amino acid groups, in both the α-helix and β-sheet conformations. Group hardness[315-322] (using $\eta = \frac{IP - EA}{2}$) was found to play the dominant role, whereas group electronegativity (using Eq (1.10)) had only a minor influence on the sequence.[314]

As an example of how the Conceptual DFT has predicted reactivity of ligands for the nucleocapsid p7 (NCp7) Zn fingers, Rice and co-workers studied,[323,156] via $f^{-}(\bar{r})$ (Eq (1.43))[177,311,141,179] and $s^{-}(\bar{r})$ (Eq (1.48))[141], the regional reactivity of the two retroviral zinc fingers of the HIV-1 NCp7 protein, representing antiviral targets. On the basis of the sum of the thiolate Fukui indices, the reactivity of finger 2 was predicted to be greater than that of finger 1. The thiolate of Cys 49 in the carboxyl terminal finger 2 turns out to be the most susceptible to electrophilic attack, providing a rationale for experimental evidence for antiviral agents that selectively target retroviral nucleocapsid protein Zn fingers.

In the catalytic reaction of serine proteases the basicity of a histidine and the nucleophilicity of a serine together with an aspartate residue belonging to the "catalytic triad" (Asp32-His64-Ser221 for subtilisin), are of great importance. The influence of amino acid substitution on the basicity and the nucleophilicity of these important amino acids was investigated by Baeten *et al.*[324] As a possible reactivity index for the serine nucleophilicity both local softness[141] and local hardness[228-230] were examined.[324] They

found that local softness (Eq (1.48)) is not suitable for describing the nucleophilicity of Ser221. Local hardness, approximated by the minimum of the MESP, $V(\bar{R})_{min}$, and the atomic charge $q(k)$ for the serine hydroxyl oxygen (O_γ), however performs very well.

Mignon *et al.*[325] also used local hardness,[228-230] approximated by atomic charge (i.e., Mulliken charge) for explaining nucleophilicity of the $2'$-hydroxyl in the Active Sites of RNase A (bovine pancreatic ribonuclease A) (EC 3.1.27.5)[326] and RNase T1 (EC 3.1.27.3).[327,328] They have shown that negative charge builds up on the $2'$-oxygen atoms upon substrate binding. The increased nucleophilicity results from stronger hydrogen bonding to the catalytic base, which is mediated by a hydrogen bond from the charged donor.

On the basis of the group softness (sum of the local softness[141] of the involved atomic centers),[315,317,320,329] Conceptual DFT studies by Rivas *et al.* have led to the suggestion of a direct hydride transfer between the reactive regions in nicotinamide and lumiflavine.[330,331] Different atomic centers of lumiflavine and nicotinamide were tested, but the smallest difference in group softness was found between the C3, Ht (terminal hydrogen atom), and C5 atoms of nicotinamide and the C4a, N1, and N5 atoms of lumiflavine, supporting the hydride transfer[332-338] between these regions.[330] In the lumiflavine molecule, the local electrophilicity[255] (i.e., $w^+(k)$; Eq (1.75)) of the N5 atom is higher than the local electrophilicity of the N1 and C4a atoms. As such, the N5 atom will most likely receive the hydride ion. When N1 is protonated, the local electrophilicity of N5 increases almost 4-fold, while, when N5 is protonated, the electrophilicity of C4a increases 10 times. As such, protonation of N1 leads to hydride transfer to C4a via N5.

In a series of subsequent papers, Roos *et al.*[339-342,331] scrutinized the experimental findings of the enzymatic reaction mechanism of *Staphylococcus aureus* arsenate reductase (ArsC) within the Conceptual DFT framework. To implement the Conceptual DFT for describing pI258 ArsC enzyme, one needs an adequate model, combining accuracy with computational tractability. Roos *et al.* in their studies applied an interesting modeling technique.[339-342] For electrophile, the model system of choice was constructed starting from the X-ray structure of the Cys15Ala mutant of ArsC complexed with arsenite (product of the first reaction step, PDB 1LJU).[343] Their model included the

complete conserved catalytic sequence motif, Cys10-X-X-Asn13-X-X-Arg16-Ser17, since the backbone amides of this substrate binding loop form hydrogen bonds with the oxygen atoms of the substrate. Amino acids 10 and 17 were terminated respectively with -NH_2 and -$CONH_2$. The side chains of residues 11, 12, 14, and 15 were terminated on a C_α, since they are positioned at the periphery of the substrate binding loop where no interaction with the substrate occurs. The three well positioned water molecules present in the active site of the PDB structure 1LJU were incorporated. Dianionic arsenate was taken as substrate. The resulting model is called "wild type (WT)".[339-342] The Ser17Ala Arg16Ala and the Asn13Ala mutants were built "in silico",[156] starting from the coordinates of the WT model. The enzymatic environment of the Cys82 nucleophile was built by using the coordinates from free wild-type ArsC (PDB file1LJL).[343] Thr11 was modelled by $HOCH_3$ and the α-helix at residues 82–89 was taken as a whole and was terminated on both sides with $CONH_2$. For the Cys89 nucleophile, the coordinates of the partially unfolded residue 82–89 helix were taken from the Cys89Leu mutant (A chain in PDB file 1LK0).[343] In both structures, hydrogen atoms were placed and optimized.

The Conceptual DFT is used to assess the nucleophilic attack of Cys10 on arsenate. The difference in local softness (i.e., $\Delta s(k) = \left| s^-(k) - s^+(k) \right|$) between the attacking nucleophilic sulfur atom of Cys10 and the receiving electrophilic arsenic atom of arsenate is minimal when dianionic arsenate was considered.[339,340] Moreover, in an another study Roos *et al*. demonstrated that both the Conceptual DFT-based reactivity analysis and the calculated thermodynamics point to a monoanionic Cys10–arseno adduct in ArsC prior to the nucleophilic attack by Cys82.[344,331] Conceptual DFT analysis indicates Ser17 to be the major activator of the electrophilic Cys10–arseno adduct. Calculation of the nucleofugality (i.e., $\Delta E_{nucleofuge} = \frac{(IP - 3EA)^2}{8(IP - EA)}$)[165-167] indicates that the enzyme increases the leaving-group capacity of OH^- (first reaction step) and of $HAsO_3^{2-}$ (second reaction step).

A most recent study by Roos *et al.*[342,331] provided fresh insight into the mechanism behind the dissociation of the mixed disulfide complexes between thioredoxin (Trx)[345] and its substrates. As a key model, the complex between Trx and its

endogenous substrate, arsenate reductase (ArsC), was used.[342] With DFT-based reactivity analysis, molecular dynamics simulations, and biochemical complex formation experiments with Cys-mutants, Trx mixed disulfide dissociation was studied by Roos *et al.*[342,331]. Information regarding the selectivity of the nucleophilic attack was obtained from a DFT based reactivity analysis.[342] In the Trx-ArsC complex, four possible reactions between the attacking nucleophilic cysteines (Cys32Trx and Cys82Trx) and the accepting electrophilic disulfide (Cys29Trx-Cys89ArsC) can be considered. The minimal local softness[141] difference (i.e., $\Delta s(k) = \left| s^{-}(k) - s^{+}(k) \right|$) of the interacting sulfur atoms favors the nucleophilic attack of Cys32Trx on Cys29Trx.[342] By studying the s_k^+ / s_k^- [226] of the sulfur atoms of the nucleophilic Cysteines in the Bs_Trx-ArsC complex they found that the Cys29Trx -Cys89ArsC disulfide is less soft than Cys32Trx and Cys82ArsC, and Cys32Trx is softer than Cys82ArsC.[342] The high reactivity of Cys32Trx toward the Cys29Trx-Cys89ArsC disulfide is consistent with the lower softness of Cys32Trx compared to Cys82Trx. On the basis of the f_k^+ / f_k^-, it was found that Cys29Trx was more susceptible to nucleophilic attack than Cys89ArsC.[342]

Thus we can see that tests of different modeling techniques on large molecular systems are very encouraging indeed. However, there is no clear choice yet as to what the best Conceptual DFT based modeling technique for large chemical and biological systems is. A final and definitive statement of the regio-chemistry of large chemical and biological systems is yet to be laid down yet. Alternatively one can think of the situation as follows: developing a novel molecular modeling technique (justified by the relevant theory) to study the regioselectivity of large chemical and biological systems. DFT based local and global reactivity descriptors, augmented with this molecular modeling technique, will be adopted as a tool in the process of the study. More detailed study certainly will be required involving these ingredients as described above. Hence, our challenge is to develop a novel approach using DFT based local reactivity descriptors and molecular modeling techniques, which will be a systematic and cost effective way to predicting regioselectivity for large systems.

1.4. Organization of the Thesis

The primary motivation to embark on this study was to devise a novel modeling technique within the framework of Conceptual DFT based reactivity descriptors for predicting the regioselectivity of large chemical and biomolecular systems. In particular, we were interested in establishing a fragmentation based approach to investigate the reactivity of large molecular systems.

Chapter I (the present chapter) gives an overview of the research work, theoretical background, limitations, and advantages of the DFT-based global and local reactivity descriptors. We discuss in detail the recent developments and applications relevant to the objective of the thesis. This chapter also presents the objectives and the overall organization of the work.

In chapter II of the thesis, we describe an approach to cope with the extended systems. For this purpose a fragmentation approach within the framework of DFT has been implemented. In particular, we have attempted to carry out an extensive study on DNA. We develop a novel approach for modeling (named 'One-into-Many' model) a large system. We would like to highlight how an intra-molecular problem of a large system can be re-casted into an inter-molecular problem of individual fragments. Relevant calculations have been performed using local reactivity descriptors (i.e., local hardness). It is the first study of its kind in this area and was then extended to other biological systems.

The broader applicability of local hardness, ($\eta(\bar{r})$) as a reliable intermolecular reactivity descriptor primarily depends on the removal of the $\frac{1}{N}$ dependence, where N is the total number of electrons of the system. In chapter III, we critically illustrate the N-dependence problem of local hardness and propose a scheme is to evaluate the local hardness values.

To make 'One-into-Many' model widely applicable, it should be based on a descriptor, which has the essential quality of handling of intermolecular reactivity and at the same time resolving analytically N-dependence problem. In chapter IV we provide a simple formalism to obtain hardness potential ($h(\bar{r})$). We propose two variants of $h(k)$

and denote them as $\Delta^{-}h(k)$ and $\Delta^{+}h(k)$, which measure reactivities toward an approaching electrophilic (i.e., El^{+}) and nucleophilic (i.e., Nu^{-}) reagent, respectively.

In chapter V of the thesis contains the major conclusions of the work and a discussion of the future scope of work.

References and Notes

1. A. Hassner, *J. Org. Chem.*, 1968, **33**, 2684.
2. J. March, *Advanced Organic Chemistry: Reactions, Mechanisms, and Structure*, 4th edn, Wiley and Sons, NY, 1992.
3. K. Fukui, T. Yonezawa and H. Shingu, *J. Chem. Phys.*, 1952, **20**, 722.
4. K. Fukui, T. Yonezawa, C. Nagata and H. Shingu, *J. Chem. Phys.*, 1954, **22**, 1433.
5. K. Fukui, *Theory of orientation and stereoselection*, Springer, Berlin, Heidelberg, New York, 1973.
6. K. Fukui, *Science*, 1987, **218**, 747.
7. A. D. Becke and K. E. Edgecombe, *J. Chem. Phys.*, 1990, **92**, 5397.
8. A. Savin, O. Jepsen, J. Flad, O. Andersen, H. Preuss and H. von Schnering, *Angew. Chem. Int. Ed.*, 1992, **31**, 187.
9. R. Bonaccorsi, E. Scrocco and J. Tomasi, *J. Chem. Phys.*, 1970, **52**, 5270.
10. P. Politzer, *J. Chem. Phys.*, 1980, **72**, 3027.
11. P. Politzer, *J. Chem. Phys.*, 1980, **73**, 3264.
12. In *Chemical Applications of Atomic and Molecular Electrostatic Potentials*, ed. P. Politzer and D. G. Truhlar, Plenum: New York, 1981.
13. B. Pullman, *Int. J. Quantum Chem. Quantum Biol. Symp.*, 1990, **17**, 81.
14. P. Sjoberg and P. Politzer, *J. Phys. Chem.*, 1990, **94**, 3959.
15. J. Tomasi, R. Bonaccorsi and R. Cammi, *Theoretical Models of Chemical Bonding, Maksic, R., Ed. Springer: Berlin*, 1990, 230.
16. G. Naray-Szabo and G. G. Ferenczy, *Chem. Rev.*, 1995, **95**, 829.
17. P. Politzer, J. V. Burda, M. Concha, P. Lane and J. S. Murray, *J. Phys. Chem. A*, 2006, **110**, 756.
18. R. G. Pearson, *J. Am. Chem. Soc.*, 1963, **85**, 3533.
19. R. G. Pearson, *Hard and Soft Acids and Bases*, Dowden, Hutchinson & Ross, Stroudsberg, PA, 1973.
20. R. G. Parr and R. G. Pearson, *J. Am. Chem. Soc.*, 1983, **105**, 7512.
21. S. Krishnamurty, R. K. Roy, R. Vetrivel, S. Iwata and S. Pal, *J. Phys. Chem. A*, 1997, **101**, 7253.

22. R. C. Deka, R. Vetrivel and S. Pal, *J. Phys. Chem. A*, 1999, **103**, 5978.
23. R. T. Sanderson, *Science*, 1951, **114**, 670.
24. G. Klopman, *J. Chem. Phys.*, 1965, **43**, S124.
25. N. C. Baird, J. M. Sichel and M. A. Whitehead, *Theor. Chim. Acta.*, 1968, **11**, 38.
26. R. T. Sanderson, *Chemical Bonds and Bond Energy*, 2nd edn., Academic Press, New York, 1976.
27. N. K. Ray, L. Samuels and R. G. Parr, *J. Chem. Phys.*, 1979, **70**, 3680.
28. W. J. Mortier, S. K. Ghosh and S. Shankar, *J. Am. Chem. Soc.*, 1986, **108**, 4315.
29. P. Hohenberg and W. Kohn, *Phys. Rev. B*, 1964, **136**, 864.
30. W. Kohn and L. J. Sham, *Phys. Rev. A*, 1965, **140**, 1133.
31. A. S. Bamzai and B. M. Deb, *Rev. Mod. Phys.*, 1981, **53**, 95.
32. S. Ghosh and B. Deb, *Phys. Rep*, 1982, **92**, 1.
33. R. Dreizler and E. Gross, *Density functional theory: an approach to the quantum many-body problem*, Springer-Verlag Berlin, 1990.
34. T. Ziegler, *Chem. Rev.*, 1991, **91**, 651.
35. E. J. Baerends and O. V. Gritsenko, *J. Phys. Chem. A*, 1997, **101**, 5383.
36. W. Koch and M. Holthausen, *A chemist's guide to density functional theory*, Wiley-Vch Weinheim, 2000.
37. P. W. Ayers and W. Yang, in *Computational Medicinal Chemistry for Drug Discovery*, ed. P. Bultinck, H. De Winter, W. Langenaeker and J. P. Tollenaere, Marcel Dekker Inc: Basel, 2004, pp. 89.
38. R. G. Parr, *J. Chem. Sci.*, 2005, **117**, 613.
39. K. Capelle, *Brazilian Journal of Physics*, 2006, **36**, 1318.
40. A. J. Cohen, P. Mori-Sanchez and W. Yang, *Science*, 2008, **321**, 792.
41. J. P. Perdew, A. Ruzsinszky, L. A. Constantin, J. Sun and G. b. I. Csonka, *J. Chem. Theory Comput.*, 2009, **5**, 902.
42. R. G. Parr, R. A. Donnelly, M. Levy and W. E. Palke, *J. Chem. Phys.*, 1978, **68**, 3801.
43. R. G. Parr, *Annu. Rev. Phys. Chem.*, 1983, **34**, 631.
44. R. F. Nalewajski, *J. Phys. Chem.*, 1985, **89**, 2831.

45. R. G. Parr and W. Yang, *Density–Functional Theory of Atoms and Molecules*, Oxford University Press, New York, 1989.
46. R. G. Parr and W. Yang, *Annu. Rev. Phys. Chem.*, 1995, **46**, 701.
47. H. Chermette, *J. Comput. Chem.*, 1999, **20**, 129.
48. P. Geerlings and F. De Proft, *Int. J. Mol. Sci.*, 2002, **3**, 276.
49. P. Geerlings, F. De Proft and W. Langenaeker, *Chem. Rev.*, 2003, **103**, 1793.
50. M. H. Cohen and A. Wasserman, *J. Phys. Chem. A*, 2007, **111**, 2229.
51. J. L. Gázquez, *J. Mex. Chem. Soc.*, 2008, **52**, 3.
52. in *Chemical reactivity theory: a density functional view*, ed. P. K. Chattaraj, CRC, 2009.
53. R. G. Pearson, *J. Chem. Educ.*, 1987, **64**, 561.
54. R. G. Pearson, *Acc. Chem. Res.*, 1990, **23**, 1.
55. R. G. Parr and P. K. Chattaraj, *J. Am. Chem. Soc.*, 1991, **113**, 1854.
56. R. G. Pearson and W. E. Palke, *J. Phys. Chem.*, 1992, **96**, 3283.
57. S. Pal, N. Vaval and R. K. Roy, *J. Phys. Chem.*, 1993, **97**, 4404.
58. R. G. Pearson, *Acc. Chem. Res.*, 1993, **26**, 250.
59. P. W. Ayers and R. G. Parr, *J. Am. Chem. Soc.*, 2000, **122**, 2010.
60. A. K. Chandra and T. Uchimaru, *J. Phys. Chem. A*, 2001, **105**, 3578.
61. M. Torrent-Sucarrat, J. M. Luis, M. Duran and M. Sola, *J. Am. Chem. Soc.*, 2001, **123**, 7951.
62. K. R. S. Chandrakumar and S. Pal, *Int. J. Mol. Sci.*, 2002, **3**, 324.
63. M. Torrent-Sucarrat, J. M. Luis, M. Duran and M. Sola, *J. Chem. Phys.*, 2002, **117**, 10561.
64. P. W. Ayers, *Faraday Discuss.*, 2007, **135**, 161.
65. *Ab initio* was introduced in the quantum chemistry by David Craig in a letter to Robert G. Parr. Interview of Professor Robert G. Parr vide in http://www.quantum-chemistry-history.com/Parr1.htm.
66. S. Goedecker, *Rev. Mod. Phys.*, 1999, **71**, 1085.
67. W. Kohn, *Phys. Rev. Lett.*, 1996, **76**, 3168.
68. L. Greengard and V. Rokhlin, *J. Comput. Phys.*, 1987, **73**, 325.

69. C. A. White, B. G. Johnson, P. M. W. Gill and M. Head-Gordon, *Chem. Phys. Lett.*, 1994, **230**, 8.
70. R. Kutteh, E. Apra and J. Nichols, *Chem. Phys. Lett.*, 1995, **238**, 173.
71. M. Challacombe, E. Schwegler and J. Almlof, *J. Chem. Phys.*, 1996, **104**, 4685.
72. M. C. Strain, G. E. Scuseria and M. J. Frisch, *Science*, 1996, **271**, 51.
73. C. A. White, B. G. Johnson, P. M. W. Gill and M. Head-Gordon, *Chem. Phys. Lett.*, 1996, **253**, 268.
74. J. M. Pérez-Jorda and W. Yang, *J. Chem. Phys.*, 1997, **107**, 1218.
75. J. C. Burant, G. E. Scuseria and M. J. Frisch, *J. Chem. Phys.*, 1996, **105**, 8969.
76. E. Schwegler and M. Challacombe, *J. Chem. Phys.*, 1996, **105**, 2726.
77. M. Daw, *Phys. Rev. B*, 1993, **47**, 10895.
78. X. P. Li, R. W. Nunes and D. Vanderbilt, *Phys. Rev. B*, 1993, **47**, 10891.
79. S. Goedecker and L. Colombo, *Phys. Rev. Lett.*, 1994, **73**, 122.
80. S. Goedecker, *J. Comput. Phys.*, 1995, **118**, 261.
81. S. Goedecker and M. Teter, *Phys. Rev. B*, 1995, **51**, 9455.
82. Q. S. Zhao and W. Yang, *J. Chem. Phys.*, 1995, **102**, 9598.
83. T. S. Lee, D. M. York and W. Yang, *J. Chem. Phys.*, 1996, **105**, 2744.
84. J. M. Millam and G. E. Scuseria, *J. Chem. Phys.*, 1997, **106**, 5569.
85. X. Li, J. M. Millam, G. E. Scuseria, M. J. Frisch and H. B. Schlegel, *J. Chem. Phys.*, 2003, **119**, 7651.
86. W. Yang, *Phys. Rev. Lett.*, 1991, **66**, 1438.
87. W. Yang and T.-S. Lee, *J. Chem. Phys.*, 1995, **103**, 5674.
88. T. E. Exner and P. G. Mezey, *J. Phys. Chem. A*, 2004, **108**, 4301.
89. D. G. Fedorov and K. Kitaura, *J. Chem. Phys.*, 2004, **120**, 6832.
90. F. L. Gu, Y. Aoki, J. Korchowiec, A. Imamura and B. Kirtman, *J. Chem. Phys.*, 2004, **121**, 10385.
91. W. Li and S. Li, *J. Chem. Phys.*, 2004, **121**, 6649.
92. X. Chen, Y. Zhang and J. Z. H. Zhang, *J. Chem. Phys.*, 2005, **122**, 184105.
93. V. Deev and M. A. Collins, *J. Chem. Phys.*, 2005, **122**, 154102.
94. D. G. Fedorov and K. Kitaura, *J. Chem. Phys.*, 2005, **123**, 134103.
95. X. He and J. Z. H. Zhang, *J. Chem. Phys.*, 2005, **122**, 031103.

96. S. Li, W. Li and T. Fang, *J. Am. Chem. Soc.*, 2005, **127**, 7215.
97. W. Li and S. Li, *J. Chem. Phys.*, 2005, **122**, 194109.
98. Y. Mei, D. W. Zhang and J. Z. H. Zhang, *J. Phys. Chem. A*, 2005, **109**, 2.
99. R. P. A. Bettens and A. M. Lee, *J. Phys. Chem. A*, 2006, **110**, 8777.
100. X. H. Chen and J. Z. H. Zhang, *J. Chem. Phys.*, 2006, **125**, 044903.
101. M. A. Collins and V. A. Deev, *J. Chem. Phys.*, 2006, **125**, 104104.
102. V. Ganesh, R. K. Dongare, P. Balanarayan and S. R. Gadre, *J. Chem. Phys.*, 2006, **125**, 104109.
103. N. Jiang, J. Ma and Y. Jiang, *J. Chem. Phys.*, 2006, **124**, 114112.
104. W. Li, T. Fang and S. Li, *J. Chem. Phys.*, 2006, **124**, 154102.
105. T. Akama, M. Kobayashi and H. Nakai, *J. Comput. Chem.*, 2007, **28**, 2003.
106. D. G. Fedorov, T. Ishida, M. Uebayasi and K. Kitaura, *J. Phys. Chem. A*, 2007, **111**, 2722.
107. D. G. Fedorov and K. Kitaura, *J. Phys. Chem. A*, 2007, **111**, 6904.
108. M. Kobayashi, Y. Imamura and H. Nakai, *J. Chem. Phys.*, 2007, **127**, 074103.
109. A. M. Lee and R. P. A. Bettens, *J. Phys. Chem. A*, 2007, **111**, 5111.
110. W. Li, S. Li and Y. Jiang, *J. Phys. Chem. A*, 2007, **111**, 2193.
111. D. G. Fedorov, J. H. Jensen, R. C. Deka and K. Kitaura, *J. Phys. Chem. A*, 2008, **112**, 11808.
112. S. Li and W. Li, *Annu. Rep. Prog. Chem., Sec C (Physical Chemistry)*, 2008, **104**, 256.
113. S. Hirata, *Phys. Chem. Chem. Phys.*, 2009, **11**, 8397.
114. For Degenerate ground state, see (a) M. Levy, *Proc. Natl. Acad. Sci. U.S.A.*, 1979, **76**, 6062; (b) M. Levy, *Phys. Rev. A*, 1982, **26**, 1200.
115. E. J. Baerends and P. Ros, *Int. J. Quantum Chem., Quantum Chem. Symp.*, 1978, **12**, 169.
116. B. I. Dunlap, J. W. D. Connolly and J. R. Sabin, *J. Chem. Phys.*, 1979, **71**, 3396.
117. P. M. Boerrigter, G. Velde and E. J. Baerends, *Int. J. Quantum Chem.*, 1988, **33**, 87.
118. J. Andzelm and E. Wimmer, *J. Chem. Phys.*, 1992, **96**, 1280.
119. L. Pauling, *J. Am. Chem. Soc.*, 1932, **54**, 3570.

120. L. Pauling, *The Nature of the Chemical Bond*, Cornell University Press, Ithaca, NY, 1960.
121. R. P. Iczkowski and J. L. Margrave, *J. Am. Chem. Soc.*, 1961, **83**, 3547.
122. R. S. Mulliken, *J. Chem. Phys.*, 1934, **2**, 782.
123. E. P. Gyftopoulos and G. N. Hatsopoulos, *Proc. Natl. Acad. Sci. U.S.A.*, 1968, **60**, 786.
124. T. A. Koopmans, *Physica*, 1933, **1**, 104.
125. An alrenative to the use of Hartree-Fock orbitals is to use Kohn-Sham orbitals, as Janak[a] did. (a) J. F. Janak*, Phys. Rev. B,* 1978, **18**, 7165.
126. B. M. Deb, *J. Am. Chem. Soc.*, 1974, **96**, 2030.
127. B. M. Deb, P. N. Sen and S. K. Bose, *J. Am. Chem. Soc.*, 1974, **96**, 2044.
128. B. M. Deb, *J. Chem. Educ.*, 1975, **52**, 314.
129. R. G. Pearson, *Proc. Natl. Acad. Sci. U. S. A.*, 1986, **83**, 8440.
130. R. K. Roy and S. Pal, *J. Phys. Chem.*, 1995, **99**, 17822.
131. R. G. Pearson, *J. Chem. Educ.*, 1999, **76**, 267.
132. J. Simons and K. D. Jordan, *Chem. Rev.*, 1987, **87**, 535.
133. A. Z. Szarka, L. A. Curtiss and J. R. Miller, *J. Chem. Phys.*, 1999, **246**, 147.
134. A. Dreuw and L. S. Cederbaum, *Chem. Rev.*, 2002, **102**, 181.
135. R. T. Sanderson, *Science*, 1955, **121**, 207.
136. R. T. Sanderson, *Polar Covalence*, Academic Press, New York, 1983.
137. R. G. Pearson, *Coord. Chem. Rev.*, 1990, **100**, 403.
138. R. G. Pearson, *Chemical Hardness*, Wiley, New York, 1997.
139. S. Pal and K. R. S. Chandrakumar, *J. Am. Chem. Soc.*, 2000, **122**, 4145.
140. K. R. S. Chandrakumar and S. Pal, *J. Phys. Chem. A*, 2002, **106**, 5737.
141. W. Yang and R. G. Parr, *Proc. Natl. Acad. Sci. U.S.A.*, 1985, **82**, 6723.
142. The Factor of 2 is arbitrary, to create a symmetry between Mulliken's electronegativity and chemical hardness.
143. F. De Proft and P. Geerlings, *Chem. Rev.*, 2001, **101**, 1451.
144. K. L. Sebastian, *Chem. Phys. Lett.*, 1994, **231**, 40.
145. P. K. Chattaraj, G. H. Liu and R. G. Parr, *Chem. Phys. Lett.*, 1995, **237**, 171.
146. K. L. Sebastian, *Chem. Phys. Lett.*, 1995, **236**, 621.

147. P. K. Chattaraj, A. Cedillo and R. G. Parr, *Chem. Phys.*, 1996, **204**, 429.
148. P. Politzer, *J. Chem. Phys.*, 1987, **86**, 1072.
149. K. D. Sen, M. C. Bohm and P. C. Schmidt, in *Electronegativity (Structure and Bonding)*, ed. K. D. Sen and C. K. Jørgenson, Springer-Verlag, Berlin, Heidelberg, 1987, vol. 66, p. 99.
150. T. K. Ghanty and S. K. Ghosh, *J. Phys. Chem.*, 1993, **97**, 4951.
151. S. Hati and D. Datta, *J. Phys. Chem.*, 1994, **98**, 10451.
152. R. K. Roy, A. K. Chandra and S. Pal, *J. Phys. Chem.*, 1994, **98**, 10447.
153. Y. Simon-Manso and P. Fuentealba, *J. Phys. Chem. A*, 1998, **102**, 2029.
154. R. F. Nalewajski, *J. Am. Chem. Soc.*, 1984, **106**, 944.
155. R. G. Parr, L. von Szentpaly and S. Liu, *J. Am. Chem. Soc.*, 1999, **121**, 1922.
156. A. T. Maynard, M. Huang, W. G. Rice and D. G. Covell, *Proc. Natl. Acad. Sci. U.S.A.*, 1998, **95**, 11578.
157. M. Elango, R. Parthasarathi, G. K. Narayanan, A. M. Sabeelullah, U. Sarkar, N. S. Venkatasubramaniyan, V. Subramanian and P. K. Chattaraj, *J. Chem. Sci.*, 2005, **117**, 61.
158. P. Bagaria and R. K. Roy, *J. Phys. Chem. A*, 2008, **112**, 97.
159. P. Bagaria, S. Saha, S. Murru, V. Kavala, B. Patel and R. Roy, *Phys. Chem. Chem. Phys.*, 2009, **11**, 8306.
160. P. W. Ayers and R. G. Parr, *J. Am. Chem. Soc.*, 2001, **123**, 2007.
161. P. Jaramillo, P. Fuentealba and P. Pérez, *Chem. Phys. Lett.*, 2006, **427**, 421.
162. P. Jaramillo, P. Perez, R. Contreras, W. Tiznado and P. Fuentealba, *J. Phys. Chem. A*, 2006, **110**, 8181.
163. A. Cedillo, R. Contreras, M. Galvan, A. Aizman, J. Andres and V. S. Safont, *J. Phys. Chem. A*, 2007, **111**, 2442.
164. F. D. Vleeschouwer, V. V. Speybroeck, M. Waroquier, P. Geerlings and F. D. Proft, *Org. Lett.*, 2007, **9**, 2721.
165. P. W. Ayers, J. S. M. Anderson and L. J. Bartolotti, *Int. J. Quantum Chem.*, 2005, **101**, 520.
166. P. W. Ayers, J. S. M. Anderson, J. I. Rodriguez and Z. Jawed, *Phys. Chem. Chem. Phys.*, 2005, **7**, 1918

167. P. R. Campodónico, C. Pérez, M. Aliaga, M. Gazitúa and R. Contreras, *Chem. Phys. Lett.*, 2007, **447**, 375.

168. S. Liu, T. Li and P. W. Ayers, *J. Chem. Phys.*, 2009, **131**, 114106.

169. P. W. Ayers, S. Liu and T. Li, *Chem. Phys. Lett.*, 2009, **480**, 318.

170. The quantity $\rho(\bar{r})$ is of course of much interest, being directly accessible experimentally and readily visualizable - just the classical density of the electronic system. The properties of $\rho(\bar{r})$ should be mentioned: see Refs. 171-173,31,174-176.

171. R. F. W. Bader, *Acc. Chem. Res.*, 1975, **8**, 34.

172. V. H. Smith Jr and I. Absar, *Israel J. Chem.*, 1977, **16**, 87.

173. R. Bader, Y. Tal, S. Anderson and T. Nguyen-Dang, *Israel J. Chem*, 1980, **19**, 8.

174. R. F. W. Bader and C. Chang, *J. Phys. Chem.*, 1989, **93**, 2946.

175. R. F. W. Bader, *Atoms in Molecules: A Quantum Theory*, 1990.

176. P. Nasertayoob and S. Shahbazian, *J. Mol. Struct. (THEOCHEM)*, 2008, **869**, 53.

177. R. G. Parr and W. Yang, *J. Am. Chem. Soc.*, 1984, **106**, 4049.

178. An explicit expression for the Fukui function in terms of K-S orbital vide Ref 311.

179. P. W. Ayers and M. Levy, *Theor. Chem. Acc.*, 2000, **103**, 353.

180. Nalewajski and Parr[a] derived the Maxwell relations in DFT and they also discussed the physical implications of Maxwell relations in DFT. (a) R. F. Nalewajski and R. G. Parr, *J. Chem. Phys.*, 1982, **77**, 399.

181. W. T. Yang, Y. K. Zhang and P. W. Ayers, *Phys. Rev. Lett.*, 2000, **84**, 5172.

182. P. W. Ayers, *J. Math. Chem.*, 2008, **43**, 285.

183. J. P. Perdew, R. G. Parr, M. Levy and J. L. Balduz, Jr., *Phys. Rev. Lett.*, 1982, **49**, 1691.

184. W. Yang and W. J. Mortier, *J. Am. Chem. Soc.*, 1986, **108**, 5708.

185. R. K. Roy, S. Pal and K. Hirao, *J. Chem. Phys.*, 1999, **110**, 8236.

186. P. W. Ayers, *J. Chem. Phys.*, 2000, **113**, 10886.

187. P. Fuentealba, P. Perez and R. Contreras, *J. Chem. Phys.*, 2000, **113**, 2544.

188. R. F. Nalewajski and R. G. Parr, *Proc. Natl. Acad. Sci. U.S.A.*, 2000, **97**, 8879.

189. R. K. Roy, K. Hirao and S. Pal, *J. Chem. Phys.*, 2000, **113**, 1372.

190. R. K. Roy, K. Hirao, S. Krishnamurty and S. Pal, *J. Chem. Phys.*, 2001, **115**, 2901.
191. R. F. Nalewajski, *Phys. Chem. Chem. Phys.*, 2002, **4**, 1710.
192. J. Oláh, C. Van Alsenoy and A. B. Sannigrahi, *J. Phys. Chem. A*, 2002, **106**, 3885.
193. P. Bultinck, R. Carbo-Dorca and W. Langenaeker, *J. Chem. Phys.*, 2003, **118**, 4349.
194. M. Mandado, C. Van Alsenoy and R. A. Mosquera, *J. Phys. Chem. A*, 2004, **108**, 7050.
195. L. J. Bartolotti and P. W. Ayers, *J. Phys. Chem. A*, 2005, **109**, 1146.
196. M. Mandado, R. A. Mosquera, A. M. Grania and C. Van Alsenoy, *Tetrahedron*, 2005, **61**, 819.
197. M. Mandado, C. Van Alsenoy and R. A. Mosquera, *J. Phys. Chem. A*, 2005, **109**, 8624.
198. M. Mandado, C. Van Alsenoy and R. A. Mosquera, *Chem. Phys. Lett.*, 2005, **405**, 10.
199. R. G. Parr, P. W. Ayers and R. F. Nalewajski, *J. Phys. Chem. A*, 2005, **109**, 3957.
200. P. W. Ayers, *Theor. Chem. Acc.*, 2006, **115**, 370.
201. F. L. Hirshfeld, *Theor. Chim. Acta*, 1977, **44**, 129.
202. P. W. Ayers, R. C. Morrison and R. K. Roy, *J. Chem. Phys.*, 2002, **116**, 8731.
203. R. S. Mulliken, *J. Chem. Phys.*, 1955, **23**, 1833.
204. J. P. Foster and F. Weinhold, *J. Am. Chem. Soc.*, 1980, **102**, 7211.
205. A. B. Rives and F. Weinhold, *Int. J. Quantum Chem. Symp.*, 1980, **14**, 201.
206. A. E. Reed and F. Weinhold, *J. Chem. Phys.*, 1983, **78**, 4066.
207. A. E. Reed, R. B. Weinstock and F. Weinhold, *J. Chem. Phys.*, 1985, **83**, 735.
208. R. K. Roy, N. Tajima and K. Hirao, *J. Phys. Chem. A*, 2001, **105**, 2117.
209. P. W. Ayers, *Proc. Natl. Acad. Sci. USA*, 2000, **97**, 1959.
210. R. F. Nalewajski, *Int. J. Mol. Sci.*, 2002, **3**, 237.
211. P. Bultinck, C. Van Alsenoy, P. W. Ayers and R. Carbo-Dorca, *J. Chem. Phys.*, 2007, **126**, 144111.
212. S. Saha, R. Roy and P. Ayers, *Int. J. Quantum Chem.*, 2009, **109**, 1790.
213. P. W. Ayers, *Phys. Chem. Chem. Phys.*, 2006, **8**, 3387.

214. J. Melin, P. W. Ayers and J. V. Ortiz, *J. Phys. Chem. A*, 2007, **111**, 10017.
215. K. S. Min, A. G. DiPasquale, J. A. Golen, A. L. Rheingold and J. S. Miller, *J. Am. Chem. Soc.*, 2007, **129**, 2360.
216. J. Cioslowski, M. Martinov and S. T. Mixon, *J. Phys. Chem.*, 1993, **97**, 10948.
217. W. Yang, Y. Zhang and P. W. Ayers, *Phys. Rev. Lett.*, 2000, **84**, 84.
218. P. W. Ayers, F. De Proft, A. Borgoo and P. Geerlings, *J. Chem. Phys.*, 2007, **126**, 224108.
219. P. Bultinck, S. Fias, C. Van Alsenoy, P. W. Ayers and R. Carbo-Dorca, *J. Chem. Phys.*, 2007, **127**, 34102.
220. P. Bultinck and R. Carbo-Dorca, *J. Math. Chem.*, 2003, **34**, 67.
221. P. W. Ayers, C. Morell, F. De Proft and P. Geerlings, *Chem. Eur. J*, 2007, **13**, 8240.
222. A. K. Chandra and M. T. Nguyen, *Int. J. Mol. Sci.*, 2002, **3**, 310.
223. M. K. Harbola, P. K. Chattaraj and R. G. Parr, *Israel J. Chem.*, 1991, **31**, 395.
224. B. G. Baekelandt, A. Cedillo and R. G. Parr, *J. Chem. Phys.*, 1995, **103**, 8548.
225. M. H. Cohen, M. V. Ganduglia-Pirovano and J. Kudrnovsky, *J. Chem. Phys.*, 1995, **103**, 3543.
226. R. K. Roy, S. Krishnamurti, P. Geerlings and S. Pal, *J. Phys. Chem. A*, 1998, **102**, 3746.
227. M. Berkowitz and R. G. Parr, *J. Chem. Phys.*, 1988, **88**, 2554.
228. M. Berkowitz, S. K. Ghosh and R. G. Parr, *J. Am. Chem. Soc.*, 1985, **107**, 6811.
229. S. K. Ghosh and M. Berkowitz, *J. Chem. Phys.*, 1985, **83**, 2976.
230. P. K. Chattaraj, D. R. Roy, P. Geerlings and M. Torrent-Sucarrat, *Theor. Chem. Acc.*, 2007, **118**, 923.
231. R. G. Parr and J. L. Gazquez, *J. Phys. Chem.*, 1993, **97**, 3939.
232. S. K. Ghosh, *Chem. Phys. Lett.*, 1990, **172**, 77.
233. J. L. Gazquez, in *Struct. Bonding 80*, ed. K. D. Sen, Springer Verlag, Berlin, 1993, pp. 27.
234. W. Langenaeker, F. de Proft and P. Geerlings, *J. Phys. Chem.*, 1995, **99**, 6424.
235. S. Liu, F. De Proft and R. G. Parr, *J. Phys. Chem.*, 1997, **101**, 6991.
236. T. Gál, *J. Phys. A: Math. Gen.*, 2002, **35**, 5899.

237. M. Torrent-Sucarrat, M. Duran and M. Sola, *J. Phys. Chem. A*, 2002, **106**, 4632.

238. T. Gál, *J. Math. Chem.*, 2007, **42**, 661.

239. M. Torrent-Sucarrat, P. Salvador, P. Geerlings and M. Solà, *J. Comput. Chem.*, 2007, **28**, 574.

240. P. W. Ayers and R. G. Parr, *J. Chem. Phys.*, 2008, **128**, 184108.

241. M. Torrent-Sucarrat, P. Salvador, M. Solà and P. Geerlings, *J. Comput. Chem.*, 2008, **29**, 1064.

242. For unconstrained local hardness vide Ref 59.

243. P. Mignon, S. Loverix and P. Geerlings, *Chem. Phys. Lett.*, 2005, **401**, 40.

244. A. S. Ozen, F. De Proft, V. Aviyente and P. Geerlings, *J. Phys. Chem. A*, 2006, **110**, 5860.

245. M. Torrent-Sucarrat and P. Geerlings, *J. Chem. Phys.*, 2006, **125**, 244101.

246. A. Borgoo, M. Torrent-Sucarrat, F. D. Proft and P. Geerlings, *J. Chem. Phys.*, 2007, **126**, 234104.

247. P. Mignon, P. Geerlings and R. Schoonheydt, *J. Phys. Chem. C*, 2007, **111**, 12376.

248. M. Torrent-Sucarrat, F. De Proft, P. Geerlings and Paul W. Ayers, *Chem. - Eur. J.*, 2008, **14**, 8652.

249. M. Torrent-Sucarrat, F. D. Proft, P. W. Ayers and P. Geerlings, *Phys. Chem. Chem. Phys.* 2010, 1**2**, 1072.

250. S. Saha and R. K. Roy, *J. Phys. Chem. B*, 2007, **111**, 9664.

251. S. Saha and R. Roy, *Phys. Chem. Chem. Phys.*, 2008, **10**, 5591.

252. S. Saha and R. K. Roy, *J. Phys. Chem. B*, 2008, **112**, 1884.

253. In an approximate method (as is done in this report) when the Fukui function is computed using Kohn-Sham (KS) level of DFT (or HF level and other higher *ab initio* methods) and the local hardness is evaluated at the Thomas-Fermi-Dirac (TFD) level (i.e., TFD approximation of $F[\rho]$ to determine $\eta(\bar{r},\bar{r}')$) it will not necessarily be a constant quantity everywhere in the molecule. See Ref. 240.

254. P. Pérez, A. Toro-Labbé, A. Aizman and R. Contreras, *J. Org. Chem.*, 2002, **67**, 4747.

255. P. K. Chattaraj, B. Maiti and U. Sarkar, *J. Phys. Chem. A*, 2003, **107**, 4973.

256. P. K. Chattaraj, U. Sarkar and D. R. Roy, *Chem. Rev.*, 2006, **106**, 2065.
257. P. K. Chattaraj and D. R. Roy, *Chem. Rev.*, 2007, **107**, PR46.
258. P. K. Chattaraj and S. Giri, *Annu. Rep. Prog. Chem., Sec C (Physical Chemistry)*, 2009, **105**, 13.
259. R. K. Roy, *J. Phys. Chem. A*, 2004, **108**, 4934.
260. D. R. Roy, R. Parthasarathi, J. Padmanabhan, U. Sarkar, V. Subramanian and P. K. Chattaraj, *J. Phys. Chem. A*, 2006, **110**, 1084.
261. R. K. Roy, V. Usha, J. Paulovič and K. Hirao, *J. Phys. Chem. A*, 2005, **109**, 4601.
262. R. K. Roy, V. Usha, B. K. Patel and K. Hirao, *J. Comput. Chem.*, 2006, **27**, 773.
263. R. F. Nalewajski, *J. Chem. Phys.*, 1984, **81**, 2088.
264. G. S. Handler and N. H. March, *J. Chem. Phys.*, 1975, **63**, 438.
265. M. J. Stott and E. Zaremba, *Phys. Rev. A*, 1980, **21**, 12.
266. M. Galvan, A. Vela and J. L. Gazquez, *J. Phys. Chem.*, 1988, **92**, 6470.
267. T. K. Ghanty and S. K. Ghosh, *J. Phys. Chem.*, 1991, **95**, 6512.
268. M. Galvan and R. Vargas, *J. Phys. Chem.*, 1992, **96**, 1625.
269. T. K. Ghanty and S. K. Ghosh, *Inorg. Chem.*, 1992, **31**, 1951.
270. T. K. Ghanty and S. K. Ghosh, *J. Phys. Chem.*, 1994, **98**, 1840.
271. T. K. Ghanty and S. K. Ghosh, *J. Phys. Chem.*, 1994, **98**, 9197.
272. T. K. Ghanty and S. K. Ghosh, *J. Am. Chem. Soc.*, 1994, **116**, 8801.
273. T. K. Ghanty and S. K. Ghosh, *J. Am. Chem. Soc.*, 1994, **116**, 3943.
274. S. K. Ghosh, *Int. J. Quantum Chem.*, 1994, **49**, 239.
275. J. Cioslowski and M. Martinov, *J. Chem. Phys.*, 1995, **102**, 7499.
276. R. Vargas and M. Galván, *J. Phys. Chem.*, 1996, **100**, 14651.
277. R. Vargas, M. Galvan and A. Vela, *J. Phys. Chem. A*, 1998, **102**, 3134.
278. P. Pérez, J. Andrés, V. S. Safont, O. Tapia and R. Contreras, *J. Phys. Chem. A*, 2002, **106**, 5353.
279. J. Oláh, F. D. Proft, T. Veszprémi and P. Geerlings, *J. Phys. Chem. A*, 2004, **108**, 490.
280. E. Chamorro and P. Perez, *J. Chem. Phys.*, 2005, **123**, 114107.
281. E. Chamorro, F. D. Proft and P. Geerlings, *J. Chem. Phys.*, 2005, **123**, 154104.

282. F. De Proft, S. Fias, C. Van Alsenoy and P. Geerlings, *J. Phys. Chem. A*, 2005, **109**, 6335.
283. J. Oláh, T. Veszprémi and M. T. Nguyen, *Chem. Phys. Lett.*, 2005, **401**, 337.
284. E. Chamorro, P. Pérez, F. De Proft and P. Geerlings, *J. Chem. Phys.*, 2006, **124**, 044105.
285. J. Garza, R. Vargas, A. Cedillo, M. Galvan and P. K. Chattaraj, *Theor. Chem. Acc*, 2006, **115**, 257.
286. D. Guerra, R. Contreras, P. Pérez and P. Fuentealba, *Chem. Phys. Lett.*, 2006, **419**, 37.
287. J. Oláh, T. Veszprémi, F. D. Proft and P. Geerlings, *J. Phys. Chem. A*, 2007, **111**, 10815.
288. E. Chamorro, P. Perez, M. Duque, F. De Proft and P. Geerlings, *J. Chem. Phys.*, 2008, **129**, 064117.
289. P. Pérez, E. Chamorro and P. W. Ayers, *J. Chem. Phys.*, 2008, **128**, 204108.
290. B. Pinter, F. De Proft, T. Veszpremi and P. Geerlings, *J. Org. Chem.*, 2008, **73**, 1243.
291. T. Gál, P. W. Ayers, F. D. Proft and P. Geerlings, *J. Chem. Phys.*, 2009, **131**, 154114.
292. T. Gál, F. De Proft and P. Geerlings, *Arxiv preprint arXiv:0903.3271*, 2009.
293. T. Gál, P. Geerlings, A. Buchmann, A. Lühr, A. Saenz, R. Sewell, J. Dingjan, F. Baumgartner, I. Llorente-Garcia and S. Eriksson, *Arxiv preprint arXiv:0910.4782*, 2009.
294. D. Guerra, R. Contreras, A. Cedillo, A. Aizman and P. Fuentealba, *J. Phys. Chem. A*, 2009, **113**, 1390.
295. İ. Uğur, F. De Vleeschouwer, N. Tüzün, V. Aviyente, P. Geerlings, S. Liu, P. W. Ayers and F. De Proft, *J. Phys. Chem. A*, 2009, **113**, 8704.
296. M. H. Cohen, M. V. Ganduglia-Pirovano and J. Kudrnovsky, *J. Chem. Phys.*, 1994, **101**, 8988.
297. B. G. Baekelandt, *J. Chem. Phys.*, 1996, **105**, 4664.
298. F. De Proft, S. Liu and R. G. Parr, *J. Chem. Phys.*, 1997, **107**, 3000.
299. F. D. Proft, S. Liu and P. Geerlings, *J. Chem. Phys.*, 1998, **108**, 7549.

300. R. Balawender and P. Geerlings, *J. Chem. Phys.*, 2001, **114**, 682.
301. R. Balawender, F. D. Proft and P. Geerlings, *J. Chem. Phys.*, 2001, **114**, 4441.
302. P. Geerlings, F. D. Proft and R. Balawender, in *Reviews of Modern Quantum Chemistry:A celebration to the contributions of R. G. Parr*, ed. K. D. Sen, World Scientific, Singapore, 2002, p. 1053.
303. M. Torrent-Sucarrat, J. M. Luis, M. Duran, A. Toro-Labbe and M. Sola, *J. Chem. Phys.*, 2003, **119**, 9393.
304. E. Chamorro, F. D. Proft and P. Geerlings, *J. Chem. Phys.*, 2005, **123**, 084104.
305. C. Cárdenas, E. Chamorro, M. Galván and P. Fuentealba, *Int. J. Quantum Chem.*, 2007, **107**, 807.
306. C. Barrientos-Salcedo, D. Arenas-Aranda, F. Salamanca-Gomez, R. Ortiz-Muniz and C. Soriano-Correa, *J. Phys. Chem. A*, 2007, **111**, 4362.
307. R. Rosal, M. Pincus, P. W. Brandt-Rauf, R. L. Fine, J. Michl and H. Wang, *Biochemistry*, 2004, **43**, 1854.
308. M. Kanovsky, A. Raffo, L. Drew, R. Rosal, T. Do, F. K. Friedman, P. Rubinstein, J. Visser, R. Robinson, P. W. Brandt-Rauf, J. Michl, R. Fine and M. R. Pincus, *Proc. Natl. Acad. Sci. U.S.A.*, 2001, **98**, 12438.
309. Y. Li and J. N. S. Evans, *J. Am. Chem. Soc.*, 1995, **117**, 7756.
310. Y. Li and J. N. S. Evans, *Proc. Natl. Acad. Sci. U.S.A.*, 1996, **93**, 4612.
311. W. Yang, R. G. Parr and R. Pucci, *J. Chem. Phys.*, 1984, **81**, 2862.
312. M. Berkowitz, *J. Am. Chem. Soc.*, 1987, **109**, 4823.
313. P. K. Chattaraj, *J. Phys. Chem. A*, 2000, **105**, 511.
314. A. Baeten, F. De Proft and P. Geerlings, *Int. J. Quantum Chem.*, 1996, **60**, 931.
315. F. De Proft, W. Langenaeker and P. Geerlings, *J. Phys. Chem.*, 1993, **97**, 1826.
316. A. Baeten, F. De Proft, W. Langenaeker and P. Geerlings, *J. Mol. Struct. (Theochem)*, 1994, **306**, 203.
317. F. De Proft, S. Amira, K. Choho and P. Geerlings, *J. Phys. Chem.*, 1994, **98**, 5227.
318. W. Langenaeker, N. Coussement, F. De Proft and P. Geerlings, *J. Phys. Chem.*, 1994, **98**, 3010.
319. A. Baeten, F. De Proft and P. Geerlings, *Chem. Phys. Lett.*, 1995, **235**, 17.

320. S. Damoun, W. Langenaeker, G. Van de Woude and P. Geerlings, *J. Phys. Chem.*, 1995, **99**, 12151.
321. F. De Proft, W. Langenaeker and P. Geerlings, *Tetrahedron*, 1995, **51**, 4021.
322. F. D. Proft, W. Langenaeker and P. Geerlings, *Int. J. Quantum Chem.*, 1995, **55**, 459.
323. M. Huang, A. Maynard, J. A. Turpin, L. Graham, G. M. Janini, D. G. Covell and W. G. Rice, *J. Med. Chem.*, 1998, **41**, 1371.
324. A. Baeten, D. Maes and P. Geerlings, *J. Theor. Biol.*, 1998, **195**, 27.
325. P. Mignon, J. Steyaert, R. Loris, P. Geerlings and S. Loverix, *J. Biol. Chem.*, 2002, **277**, 36770.
326. R. T. Raines, *Chem. Rev.*, 1998, **98**, 1045.
327. J. Steyaert, *Eur. J. Biochem.*, 1997, **247**, 1.
328. S. Loverix and J. Steyaert, *Methods Enzymol.*, 2001, **341**, 305.
329. S. Krishnamurty and S. Pal, *J. Phys. Chem. A*, 2000, **104**, 7639.
330. P. Rivas, G. Zapata-Torres, J. Melin and R. Contreras, *Tetrahedron*, 2004, **60**, 4189.
331. G. Roos, P. Geerlings and J. Messens, *J. Phys. Chem. B*, 2009, **113**, 13465.
332. S. Shinkai, N. Honda, Y. Ishikawa and O. Manabe, *J. Am. Chem. Soc.*, 1985, **107**, 6286.
333. O. Tapia, R. Cardenas, J. Andres and F. Colonna-Cesari, *J. Am. Chem. Soc.*, 1988, **110**, 4046.
334. P. A. Karplus, M. J. Daniels and J. R. Herriott, *Science*, 1991, **251**, 60.
335. F. Fieschi, V. Nivière, C. Frier, J.-L. Décout and M. Fontecave, *J. Biol. Chem.*, 1995, **270**, 30392.
336. J. Andres, V. Moliner, V. S. Safont, L. R. Domingo and M. T. Picher, *J. Org. Chem.*, 1996, **61**, 7777.
337. J. J. Tanner, B. Lei, S.-C. Tu and K. L. Krause, *Biochemistry*, 1996, **35**, 13531.
338. R. Castillo, J. Andres and V. Moliner, *J. Am. Chem. Soc.*, 1999, **121**, 12140.
339. G. Roos, S. Loverix, F. De Proft, L. Wyns and P. Geerlings, *J. Phys. Chem. A*, 2003, **107**, 6828.

340. G. Roos, J. Messens, S. Loverix, L. Wyns and P. Geerlings, *J. Phys. Chem. B*, 2004, **108**, 17216.

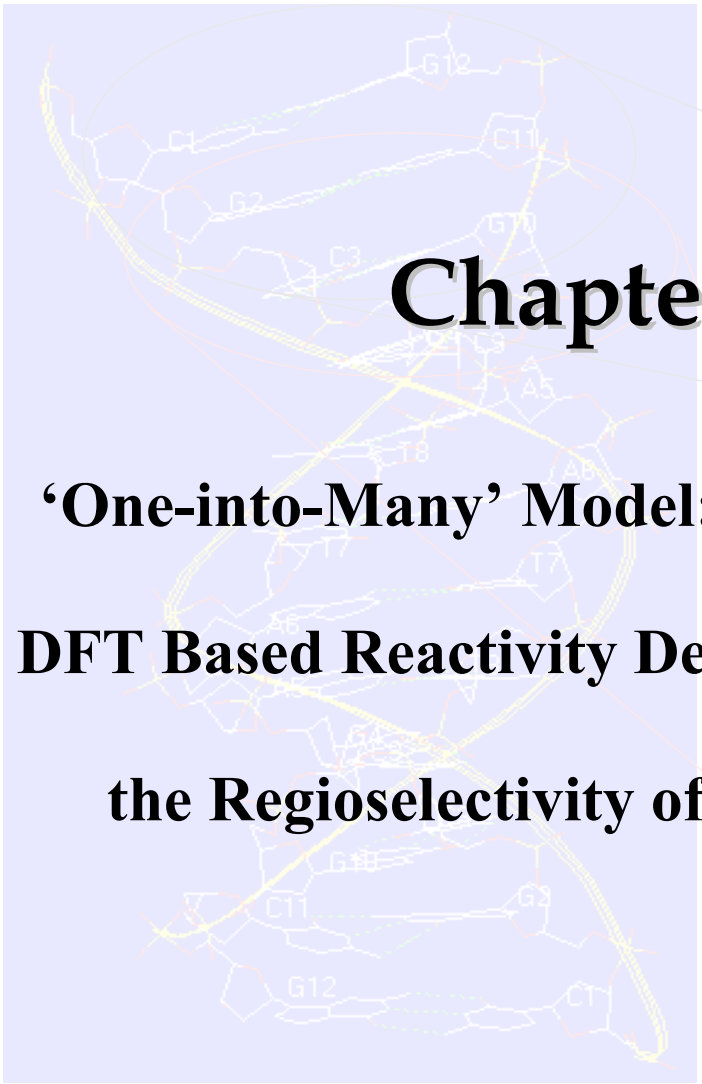
341. G. Roos, L. Buts, K. van Belle, E. Brosens, P. Geerlings, R. Loris, L. Wyns and J. Messens, *J. Mol. Biol.*, 2006, **360**, 826.

342. G. Roos, N. Foloppe, K. Van Laer, L. Wyns, L. Nilsson, P. Geerlings and J. Messens, *PLoS Comput. Biol.*, 2009, **5**, 1.

343. J. Messens, J. C. Martins, K. van Belle, E. Brosens, A. Desmyter, M. De Gieter, J. M. Wieruszeski, R. Willem, L. Wyns and I. Zegers, *Proc. Natl. Acad. Sci. U.S.A.*, 2002, **99**, 8506.

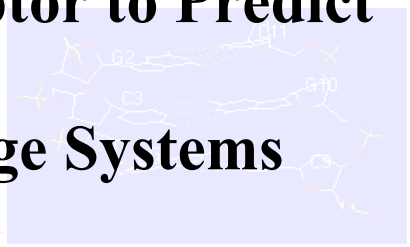
344. G. Roos, S. Loverix, E. Brosens, K. Van Belle, L. Wyns, P. Geerlings and J. Messens, *Chem. Bio. Chem.*, 2006, **7**, 981.

345. J. Messens, I. Van Molle, P. Vanhaesebrouck, M. Limbourg, K. van Belle, K. Wahni, J. C. Martins, R. Loris and L. Wyns, *J. Mol. Biol.*, 2004, **339**, 527.



Chapter II

'One-into-Many' Model: An Approach on DFT Based Reactivity Descriptor to Predict the Regioselectivity of Large Systems



2.1. Introduction

Obtaining properties of large molecules using a high level of theory has been a challenging task till date. Though the world has witnessed manifold rise in computational power, a direct calculation of the properties of a middle-sized protein from a good wave function is still considered to be almost impossible, at least in the foreseeable future, unless new methodologies are developed. Thus, detailed physical descriptions of large and complicated biological molecules for the purpose of understanding and modulating their biological functions require more intensive efforts than ever.

However, to overcome these limitations, different approaches on linear scaling-type algorithms have been reported in the literature.[1-21] In these approaches, the computational cost scales linearly with the size of the system. The divide-and-conquer (DC) method formulated by Yang,[1] and the density matrix minimization (DMM) method proposed by Li *et al*.[3] are the most popular ones among these approaches. The DC method simplifies the large-scale simulation surrounding the active sites of the concerned systems. As the name suggests, this method approximates the total electron density of the large systems as sum of the electron densities of subsystems, which are computationally trackable smaller fragments of the large system.[1] This was latter, on Yang and Lee[7] extended this method to density matrices. Of late, many workers widely applied the DC methodology to large systems.[8,10-12,14,15,20] A similar approach termed as the "molecular tailoring approach", was developed by Gadre *et al*.[4,21] to investigate the large chemical and biological systems. Analogus to DC method[1], this methods also involves breaking of a big system into the smaller ones, separate evaluation of the properties of the smaller fragments, followed by stitching together the suitably tailored smaller fragments. However, instead of electron density, the basic ingredients used by Gadre and co-workers[4,21] are the molecular electrostatic potential (MESP) and field (MEF). Li *et al*.[3] proposed the DMM method, wherein the diagonalization step in the SCF procedure is bypassed. Scuseria and collaborators[9,16] have applied the fast multipole methods (FMM) for linear scaling computation of the Coulomb integrals, a major constituent of the bottleneck in the HF and DFT calculations. Mezey and co-workers[5,6] have developed the Molecular Electron Density Lego Assembler (MEDLA) technique to construct *ab initio*

quality electron density distributions for large molecules from density distributions of small molecular fragments.

The purpose of the present chapter is to develop an alternate method to account for the regioselectivity (or better be called the 'site selectivity' in case of electrophilic and nucleophilic attack) of large chemical and biological systems. Accurate prediction of regioselectivity of large chemical and biological systems has its bottleneck in the computational limitation (because the computation is to be performed on the molecule as a whole). To overcome this problem we intend to break the large system into smaller fragments each having at least one reactive site and then comparing the intermolecular reactivity of the individual fragments using the Conceptual DFT[22-32] based local reactivity descriptor i.e., local hardness[33-35] (for the details of the Conceptual DFT based reactivity descriptors one can refer to chapter I). Theoretical justification in favour of using the local hardness based descriptor to compare the intermolecular reactivity sequences (here the fragments can be treated as individual molecules) is given in theoretical section (Section 2.2).

The newly developed method is tested on Watson-Crick double-stranded B-DNA (PDB ID: 1BNA).[36] For studying the properties of Watson-Crick double helical B-DNA, we break it into different smaller (i.e., base-pair) fragments and then compared the reactivity of the active sites in the smaller fragments with the help of the modified form of local hardness which was originally proposed by Berkowitz *et al*.[33]

In Section 2.2 we will explore the suitable reactivity descriptor for describing the intermolecular reactivity trend as well as its feasibility for computing large systems and we have discussed in brief the problems of defining the local hardness and how Langenaeker *et al*.[37] implemented a workable form. A novel extension has been proposed to evaluate the local hardness values of any active site and this approach bears some conceptual difference from that of Langenaeker *et al*.[37] In Section 2.3 we present the implementation on the Watson-Crick double helical B-DNA (PDB ID: 1BNA),[36] and also discuss various issues like break up into different base-pairs, important reactive sites on the individual bases towards electrophilic attack, etc., and the computational details. In sub-section 2.4.A. we have demonstrated how the local hardness values, evaluated in the proposed method, generate the inter-molecular reactivity trends of some substituted

benzenes. We have critically analyzed our results on B-DNA (PDB ID: 1BNA)[36] in sub-section 2.4.B. and shown whether the local hardness values of the reactive sites could explain the most reactive electrophilic sites of the whole DNA. Section 2.5 gives the summary of our investigation discussed in this chapter.

2.2. Theoretical Background:

A. (i) *On the Way of Detecting Suitable Intermolecular Reactivity Index*

In 1968, G. Klopman[38] attempted to quantify Pearson's HSAB principle[39] using polyelectronic perturbation theory and for that he defined two types of interactions, namely, orbital-controlled (i.e., soft-soft interaction) and charge-controlled (i.e., hard-hard interaction). Later on, with the development of Conceptual DFT based reactivity descriptors, it was realised that orbital-controlled reactivity descriptors include Fukui function index ($f(\bar{r})$ or $f(k)$),[40-43] local softness ($s(\bar{r})$ or $s(k)$),[42] philicity ($w(\bar{r})$ or $w(k)$),[44] 'relative electrophilicity'(s_k^+/s_k^-) and 'relative nucleophilicity' indices (s_k^-/s_k^+),[45] whereas local hardness, $\eta(\bar{r})$[33-35] (evaluated using Thomas-Fermi-Dirac (TFD) approach[46-48]), is an example of predominantly charge-controlled reactivity descriptors.

As regioselectivity is a local phenomenon, explaining the regioselectivity of any system by local reactivity descriptors are promised to be more reliable than the corresponding global reactivity descriptors. As we intend to search for suitable 'intermolecular' local reactivity index we have to look on to those orbital-controlled as well as charge-controlled reactivity descriptors.

The most useful orbital-controlled descriptor is Fukui function. Because of the normalization condition of Fukui function (i.e., $\int f(\bar{r})d\bar{r} = 1$ or $\sum_k^N f(k) = 1$)[25] it is applicable for explaining intramolecular reactivity (site-selectivity) trends and becomes less applicable for the study of intermolecular processes.

The most demanding local reactivity descriptor, which is believed to be a sustainable index for intermolecular reactivity trends, is local softness ($s(\bar{r})$ or $s(k)$), as

it describes the response of any particular site of a chemical species (in terms of change in electron density $\rho(\bar{r})$) to any global change in its chemical potential values. Furthermore, it also contains local as well as global information (see Eq (1.46)). However, local softness is (particularly individual values of s_k^+ and s_k^- are) strongly influenced by the basis set or correlation effects and because of the 'intensive' nature[49] of Fukui function ($f(\bar{r})$) the local softness parameter remains 'subintensive' (i.e., becomes smaller and smaller as the size of the system increases). For these two reasons local softness is a dubious choice as an intermolecular reactivity index.

Another reactivity descriptor which is partially orbital-controlled is philicity index ($w(\bar{r})$ or $w(k)$)[44] and is believed to be a reliable intra as well as intermolecular local reactivity index. However, $w(\bar{r})$ does not provide any extra information than that $s(\bar{r})$ or $f(\bar{r})$, as far as intramolecular reactivity is concerned.[50] Even though, individually this descriptor is 'extensive' (i.e., does not go to zero in the thermodynamic limit) in nature, the 'intensive' nature of $f(\bar{r})$ or $f(k)$ makes philicity ($w(\bar{r})$ or $w(k)$) indices applicable to limited cases[50-55] of intermolecular reactivities.

In Conceptual DFT the predominantly charge-controlled reactivity descriptor is local-hardness. The applicability of local hardness, ($\eta(\bar{r})$) as an intermolecular reactivity descriptor originates from the fact that it contains electronic part of the molecular electrostatic potential. We now present in brief, derivation of the working equations of $\eta(\bar{r})$ in order to account for its reliability as an intermolecular descriptor.

From Eqs (1.67) and (1.68) it is obvious that prescription of any routine calculation scheme for $\eta(\bar{r})$ is difficult, since the exact functional form for Hohenberg-Kohn functional $F[\rho(\bar{r})]$[56] is unknown. However, $F[\rho(\bar{r})]$ can be approximated based on the Thomas-Fermi-Dirac (TFD)[46-48] approach to DFT. If we neglect the nucleus-electron attraction in $F[\rho(\bar{r})]$, the following equation is obtained from the general form of the energy functional[25] $E^{TFD}[\rho(\bar{r})]$, without further approximations:

$$F_E^{TFD}[\rho(\bar{r})] = C_F \int \rho(\bar{r})^{5/3} d\bar{r} + \frac{1}{2}\iint \frac{\rho(\bar{r})\rho(\bar{r}')}{|\bar{r}-\bar{r}'|} d\bar{r}' d\bar{r} - C_X \int \rho(\bar{r})^{4/3} d\bar{r} \qquad (2.1)$$

Here, $C_F = \frac{3}{10}\left(3\pi^2\right)^{\frac{2}{3}} = 2.8712$ and $C_X = \frac{3}{4}\left(\frac{3}{\pi}\right)^{\frac{1}{2}} = 0.7386$ are the coefficients of the kinetic energy and exchange-energy functionals, respectively.[25]

Inserting Eqs (2.1) and (1.67), (where $\lambda[\rho(\bar{r}')]$ is replaced by $\rho(\bar{r}')$) Ghosh *et al*.[34] derived the expression of local hardness as (a detailed derivation of Eq (2.2) can be found in the Appendix A),

$$\tilde{\eta}_D^{TFD}(\bar{r}) = \frac{10}{9N} C_F \rho(\bar{r})^{2/3} - \frac{1}{N} V^{el}(\bar{r}) - \frac{4}{9N} C_X \rho(\bar{r})^{1/3} \qquad (2.2)$$

Starting from Eq (2.1) and taking into account the exponential falloff of the density in the outer regions of the system, the local hardness $\tilde{\eta}_D^{TFD'}(\bar{r})$ was approximated as,[33]

$$\tilde{\eta}_D^{TFD'}(\bar{r}) = -\frac{1}{N} V^{el}(\bar{r}) \qquad (2.3)$$

here, N is the total number of electrons of the system and $V^{el}(\bar{r})$ is the electronic part of the molecular electrostatic potential. However, Eq (2.3) can be derived by approximating just the coulombic contribution (i.e., only the middle term of Eq (2.1)).[33,57] It was shown that this approximated form of local hardness, (i.e., $-V^{el}(\bar{r})/N$) can be used as a reliable parameter for comparison of intermolecular reactivity sequences of any particular site in a series of molecules.[58-63] Recently, Geerlings and his collaborators also reported their study on the relative contributions of different energetic components to the global and local hardness values.[64,60,35]

From the preceding justification it is clear that local hardness is conducive to explain intermolecular properties. However, for predicting the overall reactivity sequence it will be more rational to consider both the charge-controlled as well as orbital-controlled contributions i.e., the descriptor should be dual in nature.[58] The argument in favour of the dual nature of the desired reactivity descriptor originates from the fact that during an electrophile-nucleophile interaction process, at the initial stage of a reaction, when two reactants approach each other, charge will play a major role in determining the reactivity. As the electrostatic forces operate from large distances, any hardness based (or charge-controlled) reactivity descriptors will be more suitable for explaining the reactivity at this stage (i.e., intermolecular reactivity sequence).[33] Once the reaction starts,

frontier orbitals play the major role in determining the reactivity of a particular site (or atom). This is because when an electrophile or a nucleophile approaches the substrate the preferable site of attack depends on the frontier orbitals of the substrate. Hence, one can argue that any softness based (orbital-controlled) reactivity descriptors (e.g., local softness ($s(\bar{r})$ or $s(k)$),[42] Fukui function index ($f(\bar{r})$ or $f(k)$),[40-43] philicity ($w(\bar{r})$ or $w(k)$)[44]) will be more suitable to describe the intramolecular reactivity. The findings of Klopman[38] also support this argument.

(ii) ***How to Recast an Intra-molecular Problem of a Large System to an Inter-molecular Problem between Smaller Sub-systems***

The bottleneck in predicting the intramolecular reactivity trends of large chemical and biological systems lies in the fact that the calculation needs to be performed on the whole system. This is because of the global softness part in the expression of local softness (i.e., $s^{\alpha}(\bar{r}) = f^{\alpha}(\bar{r})S$ where α = +, - and 0).[42] Thus, as the size of the system increases, one has to sacrifice the rigour of the theory thereby compromising the accuracy.To avoid this we intend to propose a scheme in which we break the larger system into different smaller ones, each having at least one reactive site, and then to study the reactivity of the required active sites in the individual fragments using hardness-based reactivity descriptors (e.g., local hardness). As argued in the previous sub-section [2.2.A.(i)], local hardness is mainly charge-controlled and poorly orbital-controlled. Since, it can take care of the long-range reactivity (i.e., intermolecular reactivity), in this way we can recast the intramolecular problem of a large system into an intermolecular problem of its smaller fragments and then predict the regioselectivity of the original large system. Thus, while the technique to be adopted is of DC type,[1] the local quantities (e.g., electron density, electronic contribution to the electrostatic potential) of the individual fragments are not extended to the original large system. Since the regioselectivity (or site selectivity) is a local phenomenon, it is assumed that the contribution to the local reactivity descriptor (e.g., 'local hardness' or 'atomic hardness') from the distant atoms or 'moiety' is less significant and thus can be neglected. However, contribution from nearly atoms or environment can be taken care of by careful

fragmentation process. This can be achieved (though not fully) by keeping some 'buffer zone' on both sides of the reactive site. Here, 'buffer zone' is that moiety of the chemical system which is common to two adjacent reactive sites (local hardness values of which are to be evaluated). This has been elaborated in Section 2.3 in case of '*Triple-Base-Pair Systems*'. Once, one can surmount this problem to a satisfactory level one can use local hardness descriptor (as mentioned in Section 2.2.A.(i)) for predicting reactivity trends of large chemical and biological systems.

(B) Evaluation of $\tilde{\eta}_D^{TFD}(\bar{r})$ in the Condensed Form

Based on the argument as stated above we have evaluated $\tilde{\eta}_D^{TFD}(\bar{r})$ from Eq (2.2). That is, instead of *a priori* neglecting kinetic energy functional and exchange correlation energy functional terms (i.e., the first and the third terms) in Eq (2.2) we choose to calculate them explicitly. Also, instead of evaluating electron density at a specific distance (i.e., $\rho(\bar{r})$) we consider the condensed value on a particular atom, i.e., the atomic population of any particular atom k ($P(k)$) replaces $\rho(\bar{r})$. Similarly, electronic contribution to the molecular electrostatic potential is evaluated at the atomic position (i.e., $V^{el}(\bar{r})$ is replaced by $V^{el}(k)$). Thus Eq (2.2) is modified to,

$$\tilde{\eta}_D^{TFD}(k) = \frac{10}{9N} C_F P(k)^{2/3} - \frac{1}{N} V^{el}(k) - \frac{4}{9N} C_X P(k)^{1/3} \qquad (2.4)$$

and Eq (2.3) becomes,

$$\tilde{\eta}_D^{TFD'}(k) = -\frac{1}{N} V^{el}(k) \qquad (2.5)$$

The reactivity descriptors as represented by Eqs (2.4) and (2.5) (i.e., $\tilde{\eta}_D^{TFD}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$) can be considered as a condensed form of local hardness i.e., 'atomic hardness'. It is worth mentioning here that because $\tilde{\eta}_D^{TFD}(\bar{r})$ (in Eq (2.4)) contains the nonlinear functions of $\rho(\bar{r})$ (the first and the third terms), the condensation of $\tilde{\eta}_D^{TFD}(\bar{r})$ (as it is in Eq (2.4)) is mathematically inexact.[65] Of course, any condensation inherently depends on the density partitioning scheme used and hence arbitrary.

Using Eqs (2.4) and (2.5), $\tilde{\eta}_D^{TFD}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$ values for the concerned active sites (i.e., atoms) in the individual fragments are evaluated. The fragment in which the active site has the highest value of $\tilde{\eta}_D^{TFD}(k)$ (or, $\tilde{\eta}_D^{TFD'}(k)$), is the most preferable site to be attacked by an electrophile (El^+). On the other hand, the fragment in which the active site has the lowest $\tilde{\eta}_D^{TFD}(k)$ (or, $\tilde{\eta}_D^{TFD'}(k)$) value is more prone to be attacked by a nucleophile (Nu^-). Since, the reactive sites in the individual fragments have satisfactory level of buffer zones around them (see sub-section 2.2.A.(ii) and Section 2.3. for more elaboration) one can easily locate the most preferable electrophilic and nucleophilic site of large-sized chemical or biological systems in the way as sketched above.

2.3. Computational Details

To implement the 'One-into-Many' model as outlined in previous section (subsection 2.2.A.(ii)) we have chosen a right-handed B-DNA (PDB ID: 1BNA)[36] as a model. The structure of the DNA is shown in Figure 2.1. Our model system consists of 12 base-pairs with the sequence d(CpGpCpGpApApTpTpCpGpCpG).[36] The literature survey on adduct formation indicated that the majority of known carcinogens react with DNA through N^2 and N7 positions of guanine.[66-70] Those positions are the most reactive sites towards electrophilic attack in double-stranded DNA. Apart from these positions, it is also reported that the exocyclic oxygen of guanine (O^6)[71,72] and the exocyclic oxygen of thymine (O^2)[72,73] residues are the reactive sites for electrophilic attack. In this chapter we present our calculations of $\tilde{\eta}_D^{TFD}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$ values for the reactive sites by considering one of the G·C base-pair and one of the A·T base-pair (given the name '*Single-Base-Pair Systems*') of DNA molecule (Figure 2.2). The geometries associated with these two base-pairs are taken as such from 1BNA.[36] However, for more reliable calculation we require a buffer zone around the reactive sites. So we have chosen three base-pairs at a time (named as '*Triple-Base-Pair Systems*') and evaluated the $\tilde{\eta}_D^{TFD}(k)$ (as given in Eq (2.4)) and also $\tilde{\eta}_D^{TFD'}(k)$ (as given in Eq (2.5)) values of the reactive sites in the middle base-pair. Here the base-pairs, which are on either side of the central base-pair

can created the buffer zones i.e., try to mimic the environment of the 1BNA.[36] Although, in reality, we would expect some contribution from base-pairs which are next to the adjacent base pairs, due to computational limitations we restrict our calculation to the '*Triple-Base-Pair Systems*' only. All possible combinations of '*Triple-Base-Pair Systems*' (Figure 2.3) are also taken from 1BNA[36] without any geometrical changes in the study.

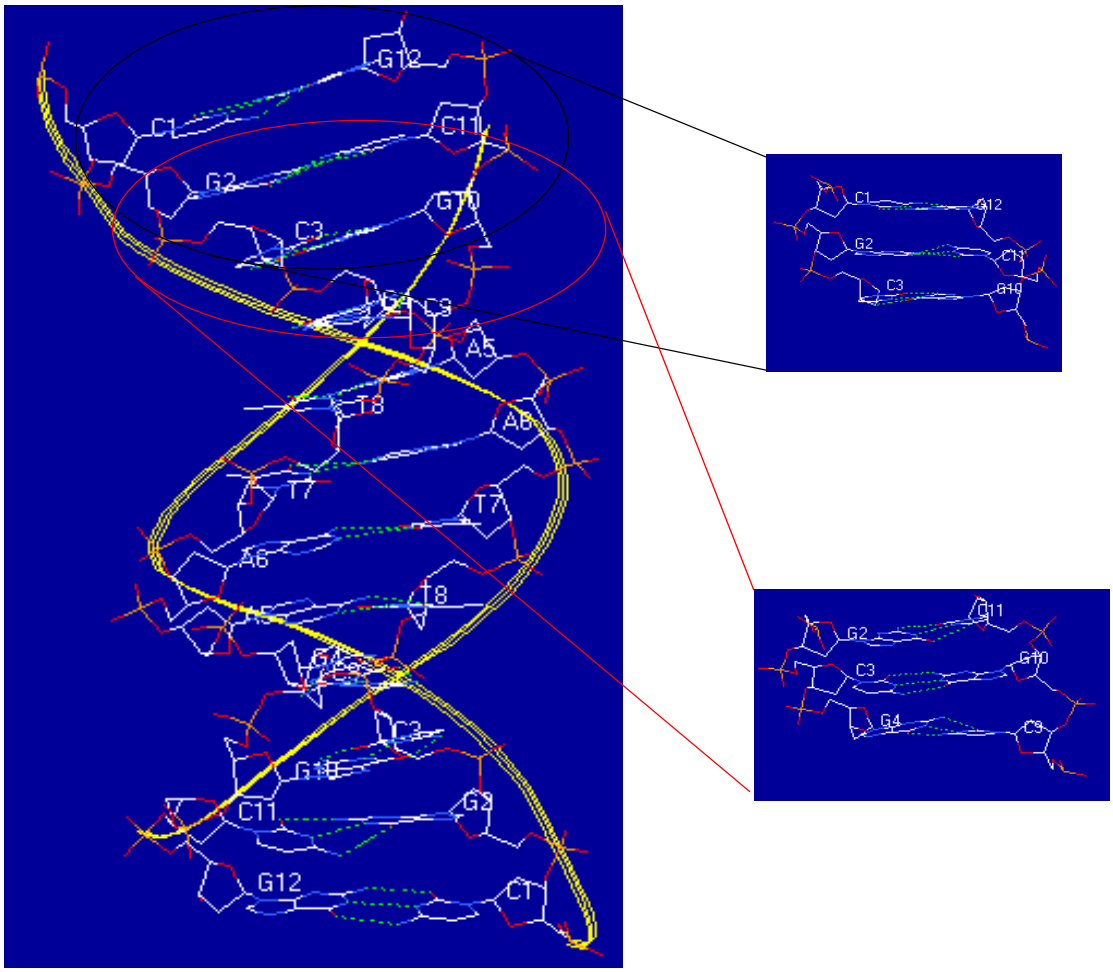
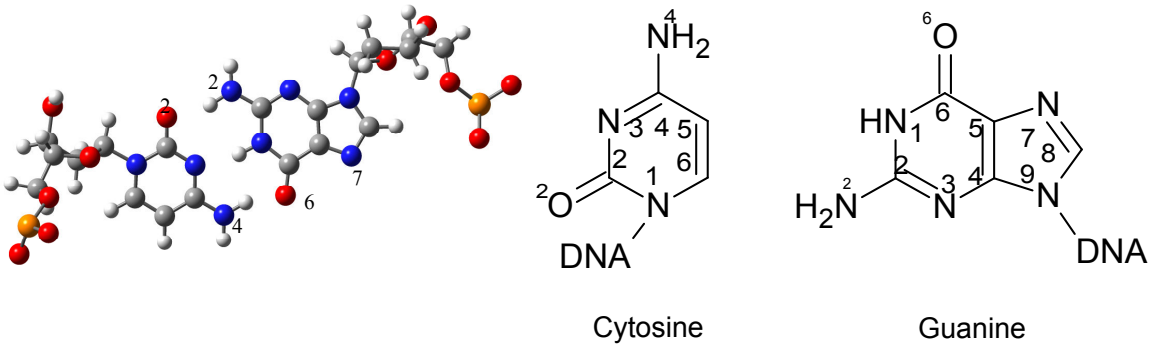
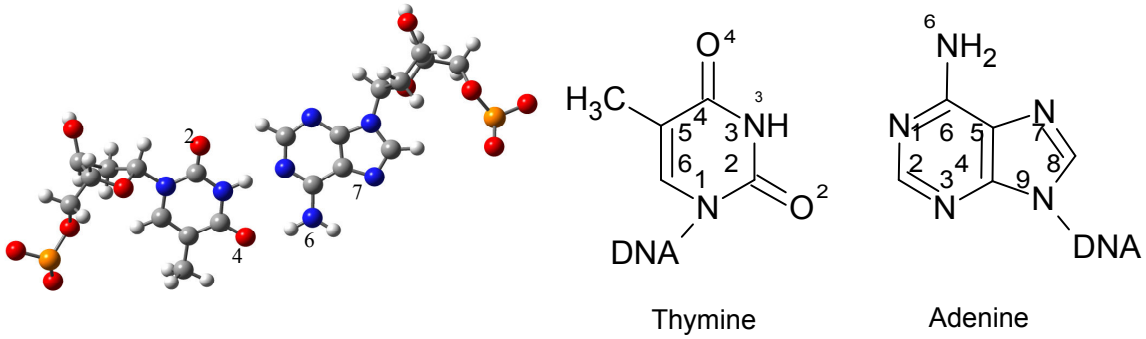


Figure 2.1: Fragmentation of Watson-Crick double-stranded B-DNA (PDB ID: 1BNA)[36] into *Triple-Base-Pair Systems*.



(a)



(b)

Figure 2.2: Two different *Single-Base-Pair Systems* (a) Cytosine-Guanine and (b) Thymine-Adenine

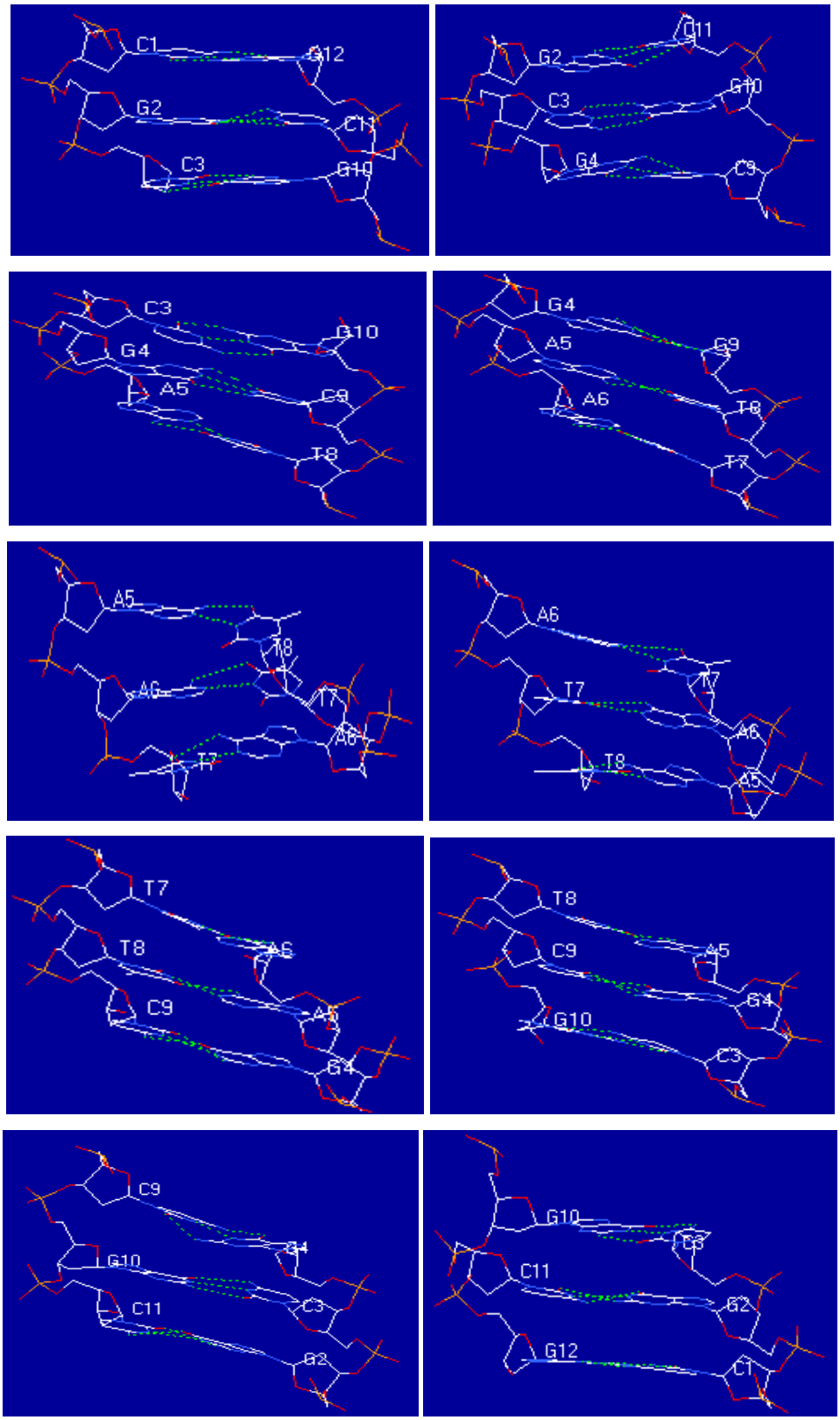


Figure 2.3: Ten different *Triple-Base-Pair Systems*.

As the total number of electrons in G·C base-pair is equal to the total number of electrons in A·T base-pair, the number of electrons (i.e., N) is constant for all the base-pairs (whether Single or Triple). Hence, we have neglected the factor $\frac{1}{N}$ while calculating $\tilde{\eta}_D^{TFD}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$, i.e., we further approximated Eqs (2.4) and (2.5) as

$$\tilde{\eta}_D^{TFD}(k) = \frac{10}{9} C_F P(k)^{2/3} - V^{el}(k) - \frac{4}{9} C_X P(k)^{1/3} \qquad (2.6)$$

where, C_F and C_X are 2.8712 and 0.7386 respectively (see Section 2.2.B.).[25]

and

$$\tilde{\eta}_D^{TFD'}(k) = -V^{el}(k) \qquad (2.7)$$

Single-point calculations were performed on the chosen base-pair systems using RHF/6-31G method. For *Single-Base-Pair Systems* we have also reported data generated from single-point calculation using superior methods like RHF/6-31G**, RHF/6-31++G**, RHF/cc-pVDZ, RHF/cc-pVTZ, LSDA/6-31G, LSDA/6-31G**, LSDA/6-31++G**, LSDA/cc-pVTZ, B3LYP/6-31++G**, B3LYP/cc-pVDZ and B3LYP/cc-pVTZ. Electron population (i.e., P_k) adopted in the present study are those obtained from Mulliken Population Analysis (MPA).[74] All the calculations have been performed by using Gaussian-98 program.[75] For fragmentation process we have used SwissPdb Viewer[76] software.

2. 4. Results and Discussion

A. Inter-molecular Reactivity Trends of Some Substituted Benzenes

To test the reliability of the $\tilde{\eta}_D^{TFD}(k)$ (Eq (2.6)) and $\tilde{\eta}_D^{TFD'}(k)$ (Eq (2.7)) we have generated those values at RHF/cc-pVTZ level for the o-, m- and p- positions of seven substituted benzenes (PhX, where X = -F, -OH, $-NH_2$, -CN, -CHO, $-CH=CH_2$ and - NO_2). These are the same systems chosen by Langenaeker *et al*.,[37] to verify the reliability of local hardness values (i.e., Eq (2.2)) generated at a distance of 4 a.u. from the concerned atomic site. The trend of experimental global hardness values is as follows[77]

$$PhF > PhOH > PhNH_2 > PhCN > PhCHO > PhCH=CH_2 > PhNO_2$$

The $\tilde{\eta}_D^{TFD}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$ values generated in our calculation (Table 2.1) also show the above trend when only o- positions are considered. It is obvious that the evaluated $\tilde{\eta}_D^{TFD}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$ values are not reliable in generating intra-molecular reactivity sequences because the substituents chosen here are not all o- directing only (e.g., - NO_2 is m- directing and so $\tilde{\eta}_D^{TFD}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$ should be highest in m- position in $PhNO_2$). Interestingly, these are the same observations made by Langenaeker *et al*.[37] from the generated local hardness values using Eq (2.3).

B. Intra-molecular Reactivity Trends of Right-handed B-DNA (PDB ID: 1BNA)[36]

In earlier papers[66-70] it has been reported that N^2 and N7 positions of guanine are the most reactive towards an electrophilic attack on DNA. Tomasz and co-worker[51,52] suggested that the cytosine of the G·C base-pair indirectly influences in the reaction between activated drug molecule (e.g., Mitomycine) and the 2-amino group (N^2) of guanine in DNA. This suggestion was based on the fact that varying the 5-substituent of cytosine from CH_3 to H to F in duplex DNA affected the rates of the alkylation of the guanine by the drug molecule (i.e., Mitomycine). Latter on Dannenberg *et al*.[69] reported that the experimental observations are due to a catalytic loan of basicity of the cytosine to guanine through one of the hydrogen bond.

The work of Muller *et al.*[78] and Moschel *et al*[79] showed that electron withdrawing groups on the N7 substitutent cause an increase in the rate of depurination of N7-alkylguanosins. Gates *et al*.,[70] reported later that the depurination is usually the predominant reaction observed for N7-alkylguanine residue in double-stranded DNA under physiological conditions and they summarized their results as follows: (i) electron withdrawing groups on the N7 substitutent facilitate depurination by improving the leaving group ability of the alkylated base and (ii) electron withdrawing substitutents (such as 2′*-OH*, 2′*-F* and 2′*-OMe*) on the sugar destabilized the oxocarbenium ion-like transition state leading to depurination, thus slowing the reaction.

Again, in several articles, it has been reported that the exocyclic oxygen of guanine (O^6)[71,72] and the exocyclic oxygen of thymine (O^2)[72,73] residues too, are important reactive sites, which are usually attacked by an incoming electrophile (El^+). Since, the oxygen of guanine (O^6) and the oxygen of thymine (O^2) greatly labilise the glycoside linkage, Singer[73] reported that O^2 of thymine in double-stranded DNA reacts to the same extent as does the O^6 of guanine. Also, preferable site of attack depends upon the nature of the incoming electrophile,[68,80] the mode of attack of that electrophile,[72,82,83,81] pH[81,82] and the nature of the solvent.[68,81,83] However, to the best knowledge of the authors till date there is no paper, either theoretical or experimental, which unambiguously discusses about the most reactive nucleophilic site among the above mentioned positions. Under such circumstances it will be more rational to make separate comparisons i.e., only between the exocyclic nitrogens, or between N7 positions, or between exocyclic oxygens of all the bases.

The results for different base-pair systems are tabulated as follows:

Single-Base-Pair Systems: The exocyclic nitrogens of DNA molecule are present in N^2 position of G, N^4 position of C and N^6 position of A, whereas the most important ring nitrogens are found in N7 position of G and A. The exocyclic oxygens of DNA are O^6 of G, O^2 of C and O^2 & O^4 of T. The first column (of Table 2.2.a, Table 2.3.a and Table 2.4.a) represents different methods adopted in our study. The second column gives the abbreviated sequence of *Single-Base-Pair Systems*. The names of the bases present in

different *Single-Base-Pair Systems* are given in the third column. The values of $\widetilde{\eta}_D^{TFD}(k)$ and $\widetilde{\eta}_D^{TFD'}(k)$ of different nucleophilic sites are shown in the subsequent columns.

Triple-Base-Pair Systems: In the case of *Triple-Base-Pair Systems*, as is mentioned in the earlier section (Section 2.3.), we have evaluated $\widetilde{\eta}_D^{TFD}(k)$ and $\widetilde{\eta}_D^{TFD'}(k)$ values of concerned atoms (by using only RHF/6-31G) of the middle base-pair. For example, in the triple base-pair systems like Cp**G**pC·Gp**C**pG, we evaluate $\widetilde{\eta}_D^{TFD}(k)$ and $\widetilde{\eta}_D^{TFD'}(k)$ values of the relevant atoms in the middle one i.e., G·C base-pair (shown in bold). Detail information for *Triple-Base-Pair Systems* are given column-wise in Table 2.2.b, Table 2.3.b and Table 2.4.b. First column represents the sequence of different *Triple-Base-Pair Systems*. The second column gives the sequence of middle base-pair, whereas the third one bears the name of bases which are present in the middle base-pair. The rest of the columns contain the same information as *Single-Base-Pair Systems* (i.e., Table 2.2.a, Table 2.3.a and Table 2.4.a).

(i) *Exocyclic Nitrogens (N^2 of G, N^4 of C and N^6 of A)*:

Single-Base-Pair Systems: For a *Single-Base-Pair Systems*, $\widetilde{\eta}_D^{TFD}(k)$ value (by using RHF/6-31G) of the exocyclic nitrogen (i.e., N^2) of guanine is found to be the largest (62.4980 a.u.) among all the exocyclic nitrogens present in DNA (Table 2.2.a). The second most reactive exocyclic nitrogen towards an electrophilic attack is N^4 position of cytosine (61.1286 a.u.) whereas the least reactive one appears to be the N^6 position of adenine (61.0051 a.u.). The trend remains unchanged when we go through higher basis sets like 6-31G**, cc-pVDZ, cc-pVTZ and also for different methods i.e., LSDA and B3LYP. It is noteworthy that the trend is similar when we calculate $\widetilde{\eta}_D^{TFD'}(k)$ (i.e., $-V^{el}(k)$) by using different basis sets as well as different methods. While comparing the exocyclic nitrogen positions by using RHF/6-31++G^{**}, LSDA/6-31++G^{**} and B3LYP/6-31++G^{**} methods, the highest $\widetilde{\eta}_D^{TFD}(k)$ values are observed to be on N^2 of G but the remaining trend is not similar to those of other methods.

Triple-Base-Pair Systems: The highest values of $\tilde{\eta}_D^{TFD}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$ (115.6112 a.u. and 103.5451 a.u. respectively), among all the 10 different *Triple-Base-Pair* combinations belong to the N^2 position of G (Table 2.2.b). The highest $\tilde{\eta}_D^{TFD}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$ values (112.2344 a.u. and 100.2253 a.u.) of N^6 of A is more than that of the corresponding values (110.8449 a.u. and 98.7451 a.u.) of N^4 position of C.

So, overall the observed trends can be summarized as follows:

For *Single-Base-Pair Systems*,

$$N^2 \ of \ G > N^4 \ of \ C > N^6 \ of \ A$$

for both $\tilde{\eta}_D^{TFD}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$

but, in the case of RHF/6-31++G**, LSDA/6-31++G**, B3LYP/6-31++G**

$$N^2 \ of \ G > N^6 \ of \ A > N^4 \ of \ C \qquad \text{for } \tilde{\eta}_D^{TFD}(k)$$

But for *Triple-Base-Pair Systems*,

$$N^2 \ of \ G > N^6 \ of \ A > N^4 \ of \ C$$

for both $\tilde{\eta}_D^{TFD}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$

(ii*) Nitrogens in 7 position (N7 of G and N7 of A)*:

Single-Base-Pair Systems: While comparing the N7 positions by using RHF/6-31G, we found that the highest $\tilde{\eta}_D^{TFD}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$ values (62.3170 a.u. and 50.7401 a.u. respectively) belong to that of G. This is true for all the methods used in our study i.e., RHF/6-31G**, RHF/6-31++G**, RHF/cc-pVDZ, RHF/cc-pVTZ, LSDA/6-31G, LSDA/6-31G**, LSDA/6-31++G**, LSDA/cc-pVTZ, B3LYP/6-31++G**, B3LYP/cc-pVDZ and B3LYP/cc-pVTZ (see Table 2.3.a).

Triple-Base-Pair Systems: For *Triple-Base-Pair Systems*, $\tilde{\eta}_D^{TFD'}(k)$ value (by using RHF/6-31G) of N7 of guanine is higher (99.2217 a.u.) than that of adenine (99.0867 a.u.) present in DNA (Table 2.3.b). Again the trend is same when $\tilde{\eta}_D^{TFD}(k)$

values are compared. The N7 position is found to be more nucleophilic in guanine (110.7587 a.u. in Gp**C**pG·Cp**G**pC) than in adenine.

A summary of the relative reactivity of various N7 positions in *Single and Triple-Base-Pair Systems* is as follows:

Single-Base-Pair Systems:

$N7\ of\ G > N7\ of\ A$ for both $\tilde{\eta}_D^{TFD}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$

Triple-Base-Pair Systems:

$N7\ of\ G > N7\ of\ A$ for both $\tilde{\eta}_D^{TFD}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$

(iii) *Exocyclic Oxygens (O^6 of G, O^2 of C and O^2 & O^4 of T)*:

Single-Base-Pair Systems: At RHF/6-31G level the highest $\tilde{\eta}_D^{TFD}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$ values (67.3282 a.u. and 54.5219 a.u. respectively) claim that the most reactive exocyclic oxygen is O^2 of cytosine (Table 2.4.a). The next most reactive exocyclic oxygen towards an electrophilic attack on it is O^2 position of thymine (66.8449 a.u. and 54.0904 a.u.) followed by O^6 position of guanine (65.8708 a.u. and 53.0815 a.u.) and O^4 position of thymine (65.6638 a.u. and 52.8764 a.u.). The same trend is observed in RHF/6-31G**, RHF/6-31++G**, RHF/cc-pVDZ, RHF/cc-pVTZ, LSDA/6-31G, LSDA/6-31G**, LSDA/6-31++G**, LSDA/cc-pVTZ, B3LYP/6-31++G**, B3LYP/cc-pVDZ and B3LYP/cc-pVTZ level for both $\tilde{\eta}_D^{TFD}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$.

Triple-Base-Pair Systems: The highest values of $\tilde{\eta}_D^{TFD}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$ are 119.6234 a.u. and 106.7696 a.u. respectively, which are on the O^2 of C (Table 2.4. b). The O^4 and O^2 of T, and O^6 of G have the next higher values of $\tilde{\eta}_D^{TFD}(k)$ (119.1068 a.u., 119.1058 a.u. and 117.5564 a.u. respectively). Although both $\tilde{\eta}_D^{TFD}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$ values are highest for O^2 of C, the trend varies for the remaining exocyclic oxygen atoms. The $\tilde{\eta}_D^{TFD'}(k)$ values of O^2 and O^4 of T exceed that of O^6 of G (106.3925 a.u., 106.3472 a.u. and 104.8922 a.u. respectively).

Thus, the observed overall trend is,

Single-Base-Pair Systems:

$$O^2\ of\ C > O^2\ of\ T > O^6\ of\ G > O^4\ of\ T$$

for both $\tilde{\eta}_D^{TFD}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$

Triple-Base-Pair Systems:

$O^2\ of\ C > O^2\ of\ T > O^4\ of\ T > O^6\ of\ G$ for $\tilde{\eta}_D^{TFD'}(k)$

but, $O^2\ of\ C > O^4\ of\ T > O^2\ of\ T > O^6\ of\ G$ for $\tilde{\eta}_D^{TFD}(k)$

(iv) *Comparison of Our Results with Earlier Investigations*:

In earlier studies,[66,68,69] the mode of action of carcinogens were investigated on the exocyclic nitrogen and it was concluded that the exocyclic amino group of G is the selective target for adduct formation by various carcinogenic agents. But there were no unambiguous reports on the relative reactivity of the remaining two exocyclic amino groups of DNA. However, it is encouraging to note that similar to previous experimental results our results also show that N^2 of G is the most reactive one.

Also, none of the earlier works[67,70] reported any comparison between N7 of G and of A. Most of these studies claim that when double stranded DNA is exposed to electrophiles the predominant reaction often occurs at N7 position of G. From our results based on *Single-Base-Pair Systems* as well as *Triple-Base-Pair Systems* N7 position of G is more reactive than that of A, as predicted by both $\tilde{\eta}_D^{TFD}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$.

In 1976, Singer[72] showed that the reactivity trend of exocyclic oxygens is $O^2\ of\ T = O^6\ of\ G > O^4\ of\ T > O^2\ of\ C$. However, our results differ from his observations. May be our results are also not full proof. There are at least two drawbacks in our calculations. The first one is that all of our calculations are in gas phase. But in reality DNA is a biological system, which has dielectric constant (ε_r) different from that of gas phase. Though the ε_r value of biological systems is very close to that of vacuum, still the trend of reactivities of the concerned atomic sites might change when the polarization effect of the biological medium is taken care of. Second, and probably the most, crucial loophole in our calculation might be the selection of basis set. Though, for

Single-Base-Pair Systems we have used higher basis sets and different methods (i.e., RHF/6-31G**, RHF/6-31++G**, RHF/cc-pVDZ, RHF/cc-pVTZ, LSDA/6-31G, LSDA/6-31G**, LSDA/6-31++G**, LSDA/cc-pVTZ, B3LYP/6-31++G**, B3LYP/cc-pVDZ and B3LYP/cc-pVTZ) these are computationally unaffordable for *Triple-Base-Pair Systems*.

However, in spite of these limitations, our calculations also confirm that the most reactive exocyclic amino group is that of N^2 position in G and the most reactive ring N-atom is that of N7 position of G as claimed in the earlier studies.[66-70]

2.5. Conclusion

In this chapter, we presented a new model to predict the regioselectivity of large chemical or biological systems. In this model a large system, right-handed B-DNA (PDB ID: 1BNA)[36] is broken into smaller fragments and then the local reactivity of the concerned atomic sites in the individual fragments are evaluated on the basis of Thomas-Fermi-Dirac (TFD) approach of density functional theory. To mimic the chemical environment buffer zones are considered surrounding the active site (in *Triple-Base-Pair Systems*).

The trends of atomic hardness values generated by the proposed model are as expected for exocyclic NH_2-groups and for ring N-atoms of the DNA base-pair systems. In the case of exocyclic nitrogen, $\tilde{\eta}_D^{TFD}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$ values of N^2 position of G is found to be the highest among all the exocyclic nitrogens present in DNA, which agrees with the previous experimental results.[66,68,69] While comparing the N7 position for *Single-Base-Pair Systems* as well as *Triple-Base-Pair Systems*, the highest $\tilde{\eta}_D^{TFD}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$ values belong to that of G, as also observed in the earlier studies.[67,70] Only for exocyclic O-atom in DNA base-pairs our method fails to generate expected trends of hardness values ($\tilde{\eta}_D^{TFD}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$). Singer[72] reported that O^6 of G and O^2 of T are the two significant reactive sites towards electrophilic attack in the double helical DNA. However, our results (both from $\tilde{\eta}_D^{TFD}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$ values) predict the O^2 postion of

C to be the most reactive site. This discrepancy may be due to the computational limitations to (i) take care of the dielectric effect of the biological medium and (ii) include polarization and diffuse functions in the basis set for *Triple-Base-Pair Systems*.

Also, the principle aim of the present work is to locate the preferred site (i.e., the atom) in DNA to be attacked by an external electrophilic (El^+) reagent. But DNA molecules have some special characteristics i.e., stacking. Recently Geerlings and co-worker[84] have shown that DFT based descriptor like local hardness ($\eta(\bar{r})$) can be used to verify the goodness of total interaction energy of stacked systems. Thus, more accurate the evaluated $\eta(\bar{r})$ values, better it can correlate with the total interaction energy of the stacked systems, (i.e., the structure of the energetically most favourable stacked DNA base-pairs). Thus once the computational limitations (as mentioned in the previous paragraph) are taken care of the proposed model ('One-into-Many' model) is expected to be used as a reliable tool to predict the regioselectivity of large systems as well as energetically favourable stacking of DNA base-pairs.

References

1. W. Yang, *Phys. Rev. Lett.*, 1991, **66**, 1438.
2. M. S. Daw, *Phys. Rev. B*, 1993, **47**, 10895.
3. X. P. Li, R. W. Nunes and D. Vanderbilt, *Phys. Rev. B*, 1993, **47**, 10891.
4. S. R. Gadre, R. N. Shirsat and A. C. Limaye, *J. Phys. Chem.*, 1994, **98**, 9165.
5. P. D. Walker and P. G. Mezey, *J. Am. Chem. Soc.*, 1994, **116**.
6. P. G. Mezey, *J. Math. Chem.*, 1995, **18**.
7. W. Yang and T. S. Lee, *J. Chem. Phys.*, 1995, **163**, 5674.
8. S. L. Dixon and K. M. Merz, Jr., *J. Chem. Phys.*, 1996, **104**, 6643.
9. M. C. Strain, G. E. Scuseria and M. Frisch, *J. Science*, 1996, **271**, 51.
10. R. E. Stratmann, G. E. Scuseria and M. Frisch, *J. Chem. Phys. Lett.*, 1996, **257**, 213.
11. S. L. Dixon and K. M. Merz, Jr., *J. Chem. Phys.*, 1997, **107**, 879.
12. J. P. Lewis, C. W. Carter, n. Jr, W. Pan, T. S. Lee and W. Yang, *J. Am. Chem. Soc.*, 1998, **120**, 5407.
13. M. Challacombe, *J. Chem. Phys.*, 1999, **110**, 2332.
14. M. D. Ermolaeva, A. van der Vaart and K. M. Merz, Jr., *J. Phys. Chem. A*, 1999, **103**, 1868.
15. V. Gogonea and K. M. Merz, Jr., *J. Phys. Chem. A*, 1999, **103**, 5171.
16. G. E. Scuseria, *J. Phys. Chem. A*, 1999, **103**, 4782.
17. A. Van der Vaart and K. M. Merz, Jr., *J. Phys. Chem. A*, 1999, **103**, 3321.
18. J. Khandogin, A. Hu and D. M. York, *J. Comput. Chem.*, 2000, **21**, 1562.
19. A. A. Bliznyuk, A. P. Rendell, T. W. Allen and S. H. Chung, *J. Phys. Chem.*, 2001, **105**, 12674.
20. C. F. Matta, *J. Phys. Chem. A*, 2001, **105**, 11088.
21. K. Babu and S. R. Gadre, *J. Comput. Chem.*, 2003, **24**, 484.
22. R. G. Parr, R. A. Donnelly, M. Levy and W. E. Palke, *J. Chem. Phys.*, 1978, **68**, 3801.
23. R. G. Parr, *Annu. Rev. Phys. Chem.*, 1983, **34**, 631.
24. R. F. Nalewajski, *J. Phys. Chem.*, 1985, **89**, 2831.

25. R. G. Parr and W. Yang, *Density–Functional Theory of Atoms and Molecules*, Oxford University Press, New York, 1989.
26. R. G. Parr and W. Yang, *Annu. Rev. Phys. Chem.*, 1995, **46**, 701.
27. H. Chermette, *J. Comput. Chem.*, 1999, **20**, 129.
28. P. Geerlings and F. De Proft, *Int. J. Mol. Sci.*, 2002, **3**, 276.
29. P. Geerlings, F. De Proft and W. Langenaeker, *Chem. Rev.*, 2003, **103**, 1793.
30. M. H. Cohen and A. Wasserman, *J. Phys. Chem. A*, 2007, **111**, 2229.
31. J. L. Gázquez, *J. Mex. Chem. Soc.*, 2008, **52**, 3.
32. in *Chemical reactivity theory: a density functional view*, ed. P. K. Chattaraj, CRC, 2009.
33. M. Berkowitz, S. K. Ghosh and R. G. Parr, *J. Am. Chem. Soc.*, 1985, **107**, 6811.
34. S. K. Ghosh and M. Berkowitz, *J. Chem. Phys.*, 1985, **83**, 2976.
35. P. K. Chattaraj, D. R. Roy, P. Geerlings and M. Torrent-Sucarrat, *Theor. Chem. Acc.*, 2007, **118**, 923.
36. H. R. Drew, R. M. Wing, T. Takano, C. Broka, S. Tanaka, K. Itakura and R. E. Dickerson, *Proc. Natl. Acad. Sci. U.S.A.*, 1981, **78**, 2179.
37. W. Langenaeker, F. De Proft and P. Geerlings, *J. Phys. Chem.*, 1995, **99**, 6424.
38. G. Klopman, *J. Am. Chem. Soc.*, 1968, **90**, 223.
39. R. G. Pearson, *J. Am. Chem. Soc.*, 1963, **85**, 3533.
40. (a) R. G. Parr and W. Yang, *J. Am. Chem. Soc.*, 1984, **106**, 4049; (b) W. Yang and W. J. Mortier, *J. Am. Chem. Soc.*, 1986, **108**, 5708.
41. W. Yang, R. G. Parr and R. Pucci, *J. Chem. Phys.*, 1984, **81**, 2862.
42. W. Yang and R. G. Parr, *Proc. Natl. Acad. Sci. U.S.A.*, 1985, **82**, 6723.
43. P. W. Ayers and M. Levy, *Theor. Chem. Acc.*, 2000, **103**, 353.
44. P. K. Chattaraj, B. Maiti and U. Sarkar, *J. Phys. Chem. A*, 2003, **107**, 4973.
45. R. K. Roy, S. Krishnamurti, P. Geerlings and S. Pal, *J. Phys. Chem. A*, 1998, **102**, 3746.
46. L. H. Thomas, *Proc. Cambridge Philos. Soc.*, 1927, **23**, 542.
47. E. Fermi, *Z. Phys.*, 1928, **48**, 73.
48. P. A. M. Dirac, *Proc. Cambridge Philos. Soc.*, 1930, **26**, 376.

49. P. W. Ayers, C. Morell, F. De Proft and P. Geerlings, *Chem. Eur. J*, 2007, **13**, 8240.
50. R. K. Roy, *J. Phys. Chem. A*, 2004, **108**, 4934.
51. A. Aizman, R. Contreras and P. Perez, *Tetrahedron*, 2005, **61**, 889.
52. R. K. Roy, V. Usha, J. Paulovič and K. Hirao, *J. Phys. Chem. A*, 2005, **109**, 4601.
53. R. K. Roy, P. Bagaria, S. Naik, V. Kavala and B. K. Patel, *J. Phys. Chem. A*, 2006, **110**, 2181.
54. R. K. Roy, V. Usha, B. K. Patel and K. Hirao, *J. Comput. Chem.*, 2006, **27**, 773.
55. P. Bagaria and R. K. Roy, *J. Phys. Chem. A*, 2008, **112**, 97.
56. P. Hohenberg and W. Kohn, *Phys. Rev. B*, 1964, **136**, 864.
57. S. Liu, F. De Proft and R. G. Parr, *J. Phys. Chem.*, 1997, **101**, 6991.
58. W. Langenaeker, F. de Proft and P. Geerlings, *J. Phys. Chem.*, 1995, **99**, 6424.
59. M. Torrent-Sucarrat, P. Geerlings, S. B. Liu and R. G. Parr, *Polish J. Chem.*, 1998, **72**, 1737.
60. M. Torrent-Sucarrat and P. Geerlings, *J. Chem. Phys.*, 2006, **125**, 244101.
61. S. Saha and R. K. Roy, *J. Phys. Chem. B*, 2007, **111**, 9664.
62. S. Saha and R. Roy, *Phys. Chem. Chem. Phys.*, 2008, **10**, 5591.
63. S. Saha and R. K. Roy, *J. Phys. Chem. B*, 2008, **112**, 1884.
64. M. Torrent-Sucarrat, J. M. Luis, M. Duran and M. Sola, *J. Mol. Struct. (Theochem)*, 2005, **727**, 139.
65. P. K. Chattaraj and S. Giri, *Annu. Rep. Prog. Chem., Sec C (Physical Chemistry)*, 2009, **105**, 13.
66. P. D. Lawley and P. Brookes, *Nature (London)*, 1961, **192**, 1081.
67. C. Nagata, A. Imamura, H. Salto and K. Fukui, *Gann*, 1963, **54**, 109.
68. A. Dipple, P. Brookes, D. S. Mackintosh and M. P. Rayman, *Biochemistry*, 1971, **10**, 4323.
69. J. J. Danneberg and M. Tomasz, *J. Am. Chem. Soc.*, 2000, **122**, 2062.
70. K. S. Gates, T. Nooner and S. Dutta, *Chem. Res. Toxicol.*, 2004, **17**, 839.
71. A. Loveless, *Nature (London)*, 1969, **223**, 206.
72. B. Singer, *Nature (London)*, 1976, **264**, 333.
73. B. Singer and J. M. Essigmann, *Carcinogenesis*, 1991, **12**, 949.

74. R. S. Mulliken, *J. Chem. Phys.*, 1955, **23**, 1833.

75. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, J. L. Andres, M. Head-Gordon, E. S. Replogle and J. A. Pople, *Gaussian 98, revision A.9*, Gaussian, Inc., Pittsburgh, PA, 1998.

76. N. Guex and M. C. Peitsch, *SWISS-MODEL and the Swiss-Pdb-Viewer: An environment for comparative protein modeling, Electrophoresis, Vol 18, 2714*, http://www.expasy.org/spdbv/, 1997.

77. R. G. Pearson, *J. Org. Chem.*, 1987, **54**, 1423.

78. N. Muller and G. Eisenbrand, *Chem. Biol. Interact.*, 1985, **53**, 173.

79. R. C. Moschel, W. R. Hudgins and A. Dipple, *J. Org. Chem.*, 1984, **49**, 363.

80. A. Dipple, *Carcinogenesis*, 1995, **16**, 437.

81. B. Singer, *Biochemistry*, 1972, **11**, 3939.

82. B. Singer, *FEBS Lett.*, 1976, **63**, 85.

83. R. Shapiro and S. J. Shiuey, *J. Org. Chem.*, 1976, **41**, 1597.

84. P. Mignon, S. Loverix, F. De Proft and P. Geerlings, *J. Phys. Chem. A*, 2004, **108**, 6038.

Table 2.1: $\widetilde{\eta}_D^{TFD}(k)$ and $\widetilde{\eta}_D^{TFD'}(k)$ values (in a.u.) in RHF/cc-pVTZ for substituted benzenes (for details see the text).

Bases	ortho		meta		para	
	$\widetilde{\eta}_D^{TFD}(k)$	$\widetilde{\eta}_D^{TFD'}(k)$	$\widetilde{\eta}_D^{TFD}(k)$	$\widetilde{\eta}_D^{TFD'}(k)$	$\widetilde{\eta}_D^{TFD}(k)$	$\widetilde{\eta}_D^{TFD'}(k)$
PhF	0.732064	0.527328	0.716504	0.514906	0.714841	0.512076
PhOH	0.729629	0.525911	0.716118	0.514217	0.714081	0.511416
$PhNH_2$	0.728713	0.524663	0.715231	0.513857	0.714352	0.510846
PhCN	0.681018	0.493414	0.668664	0.481844	0.666037	0.478887
PhCHO	0.664457	0.483647	0.650219	0.469804	0.646636	0.466371
$PhCHCH_2$	0.6633	0.482251	0.649625	0.469244	0.646559	0.465664
$PhNO_2$	0.604603	0.446624	0.585136	0.427306	0.580000	0.422445

Table 2.2.a: $\widetilde{\eta}_D^{TFD}(k)$ **and** $\widetilde{\eta}_D^{TFD'}(k)$ **values (in a.u.) in different methods for Exocyclic Nitrogens of *Single-Base-Pair Systems* (for details see the text).**

Methods	Base-pairs	Bases	N^2		N^6		N^4	
			$\widetilde{\eta}_D^{TFD}(k)$	$\widetilde{\eta}_D^{TFD'}(k)$	$\widetilde{\eta}_D^{TFD}(k)$	$\widetilde{\eta}_D^{TFD'}(k)$	$\widetilde{\eta}_D^{TFD}(k)$	$\widetilde{\eta}_D^{TFD'}(k)$
RHF/6-31G	G·C	Guanine	62.4980	50.3848	-	-	-	-
		Cytosine	-	-	-	-	61.1286	49.0444
	A·T	Adenine	-	-	61.0051	48.9105	-	-
RHF/6-31G**	G·C	Guanine	62.2928	50.3692	-	-	-	-
		Cytosine	-	-	-	-	60.9173	49.0319
	A·T	Adenine	-	-	60.7899	48.8943	-	-
RHF/6-31++G**	G·C	Guanine	62.1434	50.3605	-	-	-	-
		Cytosine	-	-	-	-	60.7356	49.0243
	A·T	Adenine	-	-	60.6910	48.8868	-	-
RHF/ cc-pVDZ	G·C	Guanine	61.6868	50.3868	-	-	-	-
		Cytosine	-	-	-	-	60.2959	49.0493
	A·T	Adenine	-	-	60.1786	48.9121	-	-
RHF/ cc-pVTZ	G·C	Guanine	61.7857	50.3865	-	-	-	-
		Cytosine	-	-	-	-	60.4289	49.0506
	A·T	Adenine	-	-	60.3157	48.9128	-	-
LSDA/6-31G	G·C	Guanine	62.1668	50.3522	-	-	-	-
		Cytosine	-	-	-	-	60.8516	49.0335
	A·T	Adenine	-	-	60.7345	48.9012	-	-
LSDA/6-31G**	G·C	Guanine	62.0448	50.3109	-	-	-	-
		Cytosine	-	-	-	-	60.7100	48.9907
	A·T	Adenine	-	-	60.5902	48.8555	-	-
LSDA/6-31++G**	G·C	Guanine	60.9925	49.8776	-	-	-	-
		Cytosine	-	-	-	-	60.3787	48.8728
	A·T	Adenine	-	-	60.4015	48.8394	-	-
LSAD/ cc-pVTZ	G·C	Guanine	61.5837	50.3088	-	-	-	-
		Cytosine	-	-	-	-	60.2619	48.9887
	A·T	Adenine	-	-	60.1529	48.8539	-	-
B3LYP/ 6-31++G**	G·C	Guanine	61.9532	50.3692	-	-	-	-
		Cytosine	-	-	-	-	60.6050	49.0317
	A·T	Adenine	-	-	60.5364	48.8918	-	-
B3LYP/ cc-pVDZ	G·C	Guanine	61.5098	50.4087	-	-	-	-
		Cytosine	-	-	-	-	60.1160	49.0684
	A·T	Adenine	-	-	59.9993	48.9288	-	-
B3LYP/ cc-pVTZ	G·C	Guanine	61.7171	50.4096	-	-	-	-
		Cytosine	-	-	-	-	60.3539	49.0715
	A·T	Adenine	-	-	60.2421	48.9321	-	-

Table 2.2.b: $\widetilde{\eta}_D^{TFD}(k)$ **and** $\widetilde{\eta}_D^{TFD'}(k)$ **values (in a.u.) in RHF/6-31G for Exocyclic Nitrogens of *Triple-Base-Pair Systems* (for details see the text).**

Base-pairs	Middle Base-pairs	Bases	N^2		N^6		N^4	
			$\widetilde{\eta}_D^{TFD}(k)$	$\widetilde{\eta}_D^{TFD'}(k)$	$\widetilde{\eta}_D^{TFD}(k)$	$\widetilde{\eta}_D^{TFD'}(k)$	$\widetilde{\eta}_D^{TFD}(k)$	$\widetilde{\eta}_D^{TFD'}(k)$
C·G **G·C** C·G	G·C	Guanine	111.9792	100.0030	-	-	-	-
		Cytosine	-	-	-	-	110.3045	98.2066
G·C **C·G** G·C	C·G	Cytosine	-	-	-	-	108.2889	96.1905
		Guanine	113.9762	101.8588	-	-	-	-
C·G **G·C** A·T	G·C	Guanine	111.9716	99.9706	-	-	-	-
		Cytosine	-	-	-	-	110.4205	98.3313
G·C **A·T** A·T	A·T	Adenine	-	-	110.9156	98.8376	-	-
A·T **A·T** T·A	A·T	Adenine	-	-	109.9442	97.9351	-	-
A·T **T·A** T·A	T·A	Adenine	-	-	109.3580	97.7641	-	-
T·A **T·A** C·G	T·A	Adenine	-	-	112.2344	100.2253	-	-
T·A **C·G** G·C	C·G	Cytosine	-	-	-	-	110.8449	98.7451
		Guanine	113.6772	101.6893	-	-	-	-
C·G **G·C** C·G	G·C	Guanine	115.6112	103.5451	-	-	-	-
		Cytosine	-	-	-	-	107.0084	94.9129
G·C **C·G** G·C	C·G	Cytosine	-	-	-	-	109.8515	97.7591
		Guanine	111.6059	99.5897	-	-	-	-

Table 2.3.a: $\widetilde{\eta}_D^{TFD}(k)$ and $\widetilde{\eta}_D^{TFD'}(k)$ values (in a.u.) in different methods for Nitrogens in 7 position (N7) of *Single-Base-Pair Systems* (for details see the text).

Methods	Base-pairs	Bases	N7	
			$\widetilde{\eta}_D^{TFD}(k)$	$\widetilde{\eta}_D^{TFD'}(k)$
RHF/ 6-31G	G·C	Guanine	62.3170	50.7401
	A·T	Adenine	61.8779	50.2518
RHF/ 6-31G**	G·C	Guanine	62.3798	50.7455
	A·T	Adenine	61.9480	50.2586
RHF/ 6-31++G**	G·C	Guanine	62.0712	50.7345
	A·T	Adenine	61.5823	50.2493
RHF/ cc-pVDZ	G·C	Guanine	62.1756	50.7606
	A·T	Adenine	61.7332	50.2740
RHF/ cc-pVTZ	G·C	Guanine	62.0698	50.7575
	A·T	Adenine	61.6506	50.2729
LSDA/ 6-31G	G·C	Guanine	62.1155	50.7231
	A·T	Adenine	61.6860	50.2602
LSDA/ 6-31G**	G·C	Guanine	62.2234	50.6913
	A·T	Adenine	61.8026	50.2268
LSDA/ 6-31++G**	G·C	Guanine	61.5215	50.3874
	A·T	Adenine	61.5232	50.2078
LSDA/ cc-pVTZ	G·C	Guanine	61.9063	50.6807
	A·T	Adenine	61.5060	50.2186
B3LYP/ 6-31++G**	G·C	Guanine	62.1018	50.7481
	A·T	Adenine	61.6347	50.2623
B3LYP/cc-pVDZ	G·C	Guanine	62.0544	50.7889
	A·T	Adenine	61.5942	50.3000
B3LYP/ cc-pVTZ	G·C	Guanine	62.0431	50.7829
	A·T	Adenine	61.6159	50.2973

Table 2.3.b: $\widetilde{\eta}_D^{TFD}(k)$ **and** $\widetilde{\eta}_D^{TFD'}(k)$ **values (in a.u.) in RHF/6-31G levels for Nitrogens in 7 position (N7) of *Triple-Base-Pair Systems* (for details see the text).**

Base-pairs	Middle Base-pairs	Bases	N7	
			$\widetilde{\eta}_D^{TFD}(k)$	$\widetilde{\eta}_D^{TFD'}(k)$
C·G **G·C** C·G	G·C	Guanine	110.4135	98.9130
G·C **C·G** G·C	C·G	Guanine	107.8791	96.2856
C·G **G·C** A·T	G·C	Guanine	109.7917	98.2793
G·C **A·T** A·T	A·T	Adenine	109.9205	98.2394
A·T **A·T** T·A	A·T	Adenine	109.4007	97.8233
A·T **T·A** T·A	T·A	Adenine	109.2854	97.2691
T·A **T·A** C·G	T·A	Adenine	110.7226	99.0867
T·A **C·G** G·C	C·G	Guanine	109.1509	97.6537
C·G **G·C** C·G	G·C	Guanine	106.0671	94.5296
G·C **C·G** G·C	C·G	Guanine	110.7587	99.2217

Table 2.4.a: $\widetilde{\eta}_D^{TFD}(k)$ and $\widetilde{\eta}_D^{TFD'}(k)$ values (in a.u.) in different methods for Exocyclic Oxygens of *Single-Base-Pair Systems* (for details see the text).

Methods	Base-pairs	Bases	O^6		O^2		O^4	
			$\widetilde{\eta}_D^{TFD}(k)$	$\widetilde{\eta}_D^{TFD'}(k)$	$\widetilde{\eta}_D^{TFD}(k)$	$\widetilde{\eta}_D^{TFD'}(k)$	$\widetilde{\eta}_D^{TFD}(k)$	$\widetilde{\eta}_D^{TFD'}(k)$
RHF/ 6-31G	G·C	Guanine	65.8708	53.0815	-	-	-	-
		Cytosine	-	-	67.3282	54.5219	-	-
	A·T	Thymine	-	-	66.8449	54.0904	65.6638	52.8764
RHF/ 6-31G**	G·C	Guanine	65.8995	53.0760	-	-	-	-
		Cytosine	-	-	67.3531	54.5166	-	-
	A·T	Thymine	-	-	66.8792	54.0846	65.6810	52.8688
RHF/ 6-31++G**	G·C	Guanine	65.7851	53.0642	-	-	-	-
		Cytosine	-	-	67.2646	54.5064	-	-
	A·T	Thymine	-	-	66.8317	54.0742	65.6284	52.8584
RHF/ cc-pVDZ	G·C	Guanine	65.7025	53.1033	-	-	-	-
		Cytosine	-	-	67.1671	54.5434	-	-
	A·T	Thymine	-	-	66.6956	54.1118	65.4894	52.8953
RHF/ cc-pVTZ	G·C	Guanine	65.7099	53.1025	-	-	-	-
		Cytosine	-	-	67.1904	54.5447	-	-
	A·T	Thymine	-	-	66.7394	54.1130	65.5329	52.8957
LSDA/ 6-31G	G·C	Guanine	65.5331	53.0230	-	-	-	-
		Cytosine	-	-	67.0120	54.4775	-	-
	A·T	Thymine	-	-	66.5631	54.0620	65.3692	52.8441
LSDA/ 6-31G**	G·C	Guanine	65.5984	52.9928	-	-	-	-
		Cytosine	-	-	67.0653	54.4467	-	-
	A·T	Thymine	-	-	66.6251	54.0303	65.4148	52.8098
LSDA/ 6-31++G**	G·C	Guanine	65.0487	52.6851	-	-	-	-
		Cytosine	-	-	66.8239	54.3064	-	-
	A·T	Thymine	-	-	66.5906	54.0094	65.4348	52.7904
LSDA/ cc-pVTZ	G·C	Guanine	65.4247	52.9946	-	-	-	-
		Cytosine	-	-	66.9287	54.4499	-	-
	A·T	Thymine	-	-	66.4937	54.0312	65.2890	52.8107
B3LYP/ 6-31++G**	G·C	Guanine	65.6708	53.0574	-	-	-	-
		Cytosine	-	-	67.0984	54.4978	-	-
	A·T	Thymine	-	-	66.6998	54.0713	65.5365	52.8535
B3LYP/ cc-pVDZ	G·C	Guanine	65.5361	53.1057	-	-	-	-
		Cytosine	-	-	66.9948	54.5427	-	-
	A·T	Thymine	-	-	66.5409	54.1177	65.3240	52.8989
B3LYP/ cc-pVTZ	G·C	Guanine	65.6087	53.1035	-	-	-	-
		Cytosine	-	-	67.0849	54.5444	-	-
	A·T	Thymine	-	-	66.6329	54.1140	65.4323	52.8990

Table 2.4.b: $\widetilde{\eta}_D^{TFD}(k)$ **and** $\widetilde{\eta}_D^{TFD'}(k)$ **values (in a.u.) in RHF/6-31G level for Exocyclic Oxygens of *Triple-Base-Pair Systems* (for details see the text).**

Base-pairs	Middle Base-pairs	Bases	O^6		O^2		O^4	
			$\widetilde{\eta}_D^{TFD}(k)$	$\widetilde{\eta}_D^{TFD'}(k)$	$\widetilde{\eta}_D^{TFD}(k)$	$\widetilde{\eta}_D^{TFD'}(k)$	$\widetilde{\eta}_D^{TFD}(k)$	$\widetilde{\eta}_D^{TFD'}(k)$
C·G	G·C	Guanine	116.7570	104.1631	-	-	-	-
G·C								
C·G		Cytosine	-	-	116.8022	103.9643	-	-
G·C	C·G	Cytosine	-	-	118.1496	105.3226	-	-
C·G								
G·C		Guanine	114.7804	101.9745	-	-	-	-
C·G	G·C	Guanine	117.5564	104.8922	-	-	-	-
G·C								
A·T		Cytosine	-	-	117.9860	105.1515	-	-
G·C	A·T	Thymine	-	-	118.1969	105.4565	114.8894	102.0863
A·T								
A·T								
A·T	A·T	Thymine	-	-	118.4327	105.7667	112.6297	99.8808
A·T								
T·A								
A·T	T·A	Thymine	-	-	119.1058	106.3925	111.1247	98.3704
T·A								
T·A								
T·A	T·A	Thymine	-	-	114.0554	101.2365	119.1068	106.3472
T·A								
C·G								
T·A	C·G	Cytosine	-	-	118.4255	105.6081	-	-
C·G								
G·C		Guanine	116.5885	103.9271	-	-	-	-
C·G	G·C	Guanine	112.6259	99.9286	-	-	-	-
G·C								
C·G		Cytosine	-	-	119.6234	106.7696	-	-
G·C	C·G	Cytosine	-	-	116.4802	103.6573	-	-
C·G								
G·C		Guanine	117.1832	104.4664	-	-	-	-

Chapter III

N-dependence Problem of Local Hardness

Parameter

3.1. Introduction

After the introducing a novel fragmentation based method, named as 'One-into-many',[1,2] to study the regioselectivity of large chemical and biological systems in the preceding chapter (i.e., chapter II), in this chapter we critically illustrated of the limitation of local hardness[3-5] $\eta(\bar{r})$, evaluated from two composite functions i.e., $\lambda[\rho(\bar{r}')] = \rho(\bar{r}')$[4,6-10,1,2] and $\lambda[\rho(\bar{r}')] = Nf(\bar{r}')$[11-18], when used for comparison of intermolecular reactivity trends in the systems of different sizes but having common reactive centers. The broader applicability of $\eta(\bar{r})$ primarily depends on the removal of its $\frac{1}{N}$ dependence. The $\frac{1}{N}$ term, in the expression of $\eta(\bar{r})$, sometimes alters the expected trend of reactivity of active sites. So, an attempt is made to illustrate the problems of evaluating $\eta(\bar{r})$ and an alternate scheme is proposed to evaluate the local hardness values. In Section 3.2 we have briefly discussed, the definition of the local hardness (the details of which have been discussed already in chapter I). In Section 3.3 the problems of evaluating local hardness $\eta(\bar{r})$, using the two composite functions $\rho(\bar{r})$ and $Nf(\bar{r}')$, are discussed. The implementation of the computational methodology is also elaborated in Section 3.3. In Section 3.4, a plausible scheme to bypass those problems of evaluating $\eta(\bar{r})$ is demonstrated. Finally, in the concluding section (Section 3.5) we have summarized the study of the present chapter.

3.2. Local Hardness Descriptor

We now turn to the definition of the local hardness $\eta(\bar{r})$.[3-5] Eq (1.65) shows that if we define local hardness by the formula:

$$\eta_{\lambda}(\bar{r}) = \left(\frac{\delta\mu}{\delta\rho(\bar{r})}\right)_{v(\bar{r})} = \frac{1}{N}\int \eta(\bar{r},\bar{r}')\lambda[\rho(\bar{r}')]d\bar{r}'$$

$$= \frac{1}{N}\int \frac{\delta^2 F_E}{\delta\rho(\bar{r}')\delta\rho(\bar{r})}\lambda[\rho(\bar{r}')]d\bar{r}' \qquad (3.1)$$

we will have two important choices of the composite function $\lambda[\rho(\bar{r}')]$ are $\rho(\bar{r}')$ (the electron density)[4,6-10,1,2] and $Nf(\bar{r}')$ (Fukui function multiplied by the total number of electrons of the system),[11-18]

$$\lambda[\rho(\bar{r}')] = \rho(\bar{r}') \qquad \text{yielding} \qquad \tilde{\eta}_D(\bar{r}) = \frac{1}{N}\int \eta(\bar{r},\bar{r}')\rho(\bar{r}')d\bar{r}' \tag{3.2}$$

and $$\lambda[\rho(\bar{r}')] = Nf(\bar{r}') \qquad \text{yielding} \qquad \tilde{\eta}_F(\bar{r}) = \int \eta(\bar{r},\bar{r}')f(\bar{r}')d\bar{r}' \tag{3.3}$$

The present chapter demonstrates a critical illustration of the limitation of $\eta(\bar{r})$ (evaluated from the above mentioned composite functions) when used for comparison of intermolecular reactivity trends between systems of different sizes but having common reactive centers.

3.3. Problems of Evaluating the Local Hardness $\eta(\bar{r})$ Using $\rho(\bar{r}')$ and $Nf(\bar{r}')$

A. Problem of Evaluating the Local Hardness $\eta(\bar{r})$ Using $\lambda[\rho(\bar{r}')] = \rho(\bar{r}')$

Starting from the Thomas-Fermi-Dirac (TFD)[19-21] approach of DFT and based on the general form of the energy functional[22] $E^{TFD}[\rho(\bar{r})]$ (and keeping in mind that the nucleus-electron attraction is not contained in $F_E[\rho(\bar{r})]$), Eq (3.4) can be obtained without further approximation (as we did in chapter II),

$$F_E^{TFD}[\rho(\bar{r})] = C_F\int \rho(\bar{r})^{5/3}d\bar{r} + \frac{1}{2}\iint \frac{\rho(\bar{r})\rho(\bar{r}')}{|\bar{r}-\bar{r}'|}d\bar{r}'d\bar{r} - C_X\int \rho(\bar{r})^{4/3}d\bar{r} \tag{3.4}$$

Here, $C_F = \frac{3}{10}\left(3\pi^2\right)^{\frac{2}{3}} = 2.8712$ and $C_X = \frac{3}{4}\left(\frac{3}{\pi}\right)^{\frac{1}{2}} = 0.7386$ are the coefficients of the kinetic energy and exchange-energy functionals, respectively.[22] Inserting Eq (3.4) in Eq (3.2) (where $\lambda[\rho(\bar{r}')]$ is replaced by $\rho(\bar{r}')$) Ghosh *et al*.[4] derived the expression of local hardness as,

$$\tilde{\eta}_D^{TFD}(\bar{r}) = \frac{10}{9N}C_F\rho(\bar{r})^{2/3} - \frac{1}{N}V^{el}(\bar{r}) - \frac{4}{9N}C_X\rho(\bar{r})^{1/3} \tag{3.5}$$

Starting from Eq (3.5) and taking into account the exponential falloff of the density in the outer regions of the system, the local hardness $\tilde{\eta}_D^{TFD'}(\bar{r})$ was approximated as,[3]

$$\tilde{\eta}_D^{TFD'}(\bar{r}) = -\frac{1}{N}V^{el}(\bar{r}) \tag{3.6}$$

here, N is the total number of electrons of the system and $V^{el}(\bar{r})$ is the electronic part of the molecular electrostatic potential.

Instead of evaluating electron density at a specific distance (i.e., $\rho(\bar{r})$) we consider the condensed value on a particular atom in chapter II [i.e., the atomic population of any particular atom k ($P(k)$) replaces $\rho(\bar{r})$]. Similarly, electronic contribution to the molecular electrostatic potential is evaluated at the atomic position (i.e., $V^{el}(\bar{r})$ is replaced by $V^{el}(k)$). Thus Eq (3.5) is modified to,

$$\tilde{\eta}_D^{TFD}(k) = \frac{10}{9N}C_F P(k)^{2/3} - \frac{1}{N}V^{el}(k) - \frac{4}{9N}C_X P(k)^{1/3} \tag{3.7}$$

and Eq (3.6) becomes,

$$\tilde{\eta}_D^{TFD'}(k) = -\frac{1}{N}V^{el}(k) \tag{3.8}$$

The reactivity descriptors as represented by Eqs (3.7) and (3.8) (i.e., $\tilde{\eta}_D^{TFD}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$) can be considered as condensed forms of local hardness i.e., 'atomic hardness'.[1,2] It is worth mentioning here that because $\tilde{\eta}_D^{TFD}(\bar{r})$ (in Eq. (3.5)) contains the nonlinear functions of $\rho(\bar{r})$ (the first and the third terms), the condensation of $\tilde{\eta}_D^{TFD}(\bar{r})$ to $\tilde{\eta}_D^{TFD}(k)$ (as it is in Eq (3.7)) is mathematically inexact.[23] Of course, any condensation inherently depends on the density partitioning scheme used and hence arbitrary. However, earlier study[1,2] reveals that a condensed-to-atom local hardness (i.e., $\tilde{\eta}_D^{TFD}(k)$), including all three terms, provides a reasonable good correlation with the experimental findings of biological systems (see chapter II for more elabortation).[1,2]

In the present study eleven carbonyl compounds are chosen to test whether the local hardness (or better be called 'atomic hardness') descriptor, as defined by Eqs (3.7) and (3.8), are suitable for comparison of intermolecular reactivity trends. These are, HCHO, CH_3CHO, CH_3CH_2CHO, C_6H_5CHO, CH_3COCH_3, $CH_3CH_2COCH_3$,

$CH_3CH_2CH_2COCH_3$, $C_6H_5COCH_3$, CH_2=CHCHO, CH_3CH=CHCHO and C_6H_5CH=CHCHO, which are of different sizes. These eleven carbonyl compounds are categorized into three groups.

In *Category I*, we have taken four simple aldehyde compounds: $HCHO$, CH_3CHO, CH_3CH_2CHO and C_6H_5CHO.

The systems we have chosen in the *Category II* are: CH_3COCH_3, $CH_3CH_2COCH_3$, $CH_3CH_2CH_2COCH_3$ *and* $C_6H_5COCH_3$.

Where as in *Category III*, we have grouped remaining three systems (i.e., $CH_2=CHCHO$, $CH_3CH=CHCHO$ and $C_6H_5CH=CHCHO$) having a $C=C$ bond in conjugation with a $C=O$ group.

The geometries are optimized using MP2/6-31G** method. Electron population (i.e., $P(k)$), adopted in the present study, are those obtained from Mulliken Population Analysis (MPA).[53] All the calculations have been performed by using Gaussian-98 program.[54] Normally, the carbon atom in the carbonyl group (i.e., $C_{C=O}$) is the most reactive site towards an electron donating species (i.e., nucleophile, Nu^-), whereas, the oxygen atom in the carbonyl group (i.e., $O_{C=O}$) is the most reactive site towards an electron attracting species (i.e., electrophile, El^+).

On the basis of calculated $\tilde{\eta}_D^{TFD}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$ (i.e., Eqs (3.7) and (3.8)) values of these reactive sites the intermolecular reactivity trend of the chosen carbonyl compounds are compared. The system for which the $O_{C=O}$ shows the highest value of $\tilde{\eta}_D^{TFD}(k)$ (and $\tilde{\eta}_D^{TFD'}(k)$) will be the most preferable site to be attacked by an electrophile (El^+). On the other hand, the system for which $C_{C=O}$ is having the lowest $\tilde{\eta}_D^{TFD}(k)$ (and $\tilde{\eta}_D^{TFD'}(k)$) value is more prone to be attacked by a nucleophile (Nu^-).

Category I: A summary of the relative reactivity of $C_{C=O}$ of the chosen systems in *Category I* towards an electron donating species (nucleophile, Nu^-) is as follows (Table 3.1),

$$C_6H_5CHO > CH_3CH_2CHO > CH_3CHO > HCHO$$

(from both $\tilde{\eta}_D^{TFD}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$).

And the observed overall trend of $O_{C=O}$ reactivity towards an electron attracting species (i.e., electrophile, El^+) can be summarized as follows (Table 3.1),

$$HCHO > CH_3CHO > CH_3CH_2CHO > C_6H_5CHO$$

(from both $\tilde{\eta}_D^{TFD}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$).

The reactivity of these compounds are highly dependent not only on the electronic charge (i.e., population) of reactive centers (i.e., $C_{C=O}$ and $O_{C=O}$) but also on the changes of these charges with the change of global number of electrons (i.e., N). However, when local hardness is evaluated using Eqs (3.7) and (3.8) it is the first factor which decides the reactivity (Section 3.3.B. contains discussion about the dependence of reactivity on the second factor). From Eqs (3.7) and (3.8) we can argue that more negative the $O_{C=O}$ is, the more reactive (i.e., more nucleophilic) it is [as more negative charge means more electronic population of $O_{C=O}$, which increases the numerical value of the individual terms in Eqs (3.7) and (3.8). Because $V^{el}(k)$ is negative valued and is dominant among the three, $\tilde{\eta}_D^{TFD}(k)$ increases with increasing negative charge of $O_{C=O}$]. Similarly, more positive the $C_{C=O}$ is, the more reactive (i.e., more electrophilic) it is. The electron donating (i.e., electron pushing or $+I$ effect) tendency of the moieties attached to the $-CHO$ group is, $C_6H_5- > CH_3CH_2- > CH_3- > H-$. Naturally, the positive charge on the $C_{C=O}$ will follow the reverse trend. Thus, the expected trend of the reactivity of $C_{C=O}$ towards an electron donating species (Nu^-) should be $HCHO > CH_3CHO > CH_3CH_2CHO > C_6H_5CHO$ for both $\tilde{\eta}_D^{TFD}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$.

On the other hand, the expected trend of reactivity of $O_{C=O}$ towards an electron attracting species (El^+) should be $C_6H_5CHO > CH_3CH_2CHO > CH_3CHO > HCHO$ for both $\tilde{\eta}_D^{TFD}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$. This is because more the positive charge on $C_{C=O}$ less the negative charge on $O_{C=O}$ causing the above trend.

Category II: The observed trend of reactivity of $C_{C=O}$ (towards a Nu^-) can be summarized as follows (Table 3.1),

$$C_6H_5COCH_3 > CH_3CH_2CH_2COCH_3 > CH_3CH_2COCH_3 > CH_3COCH_3$$

from both $\tilde{\eta}_D^{TFD}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$.

And a summary of the relative reactivity of $O_{C=O}$ (towards an El^+) in the chosen systems is as follows (Table 3.1),

$$CH_3COCH_3 > CH_3CH_2COCH_3 > CH_3CH_2CH_2COCH_3 > C_6H_5COCH_3$$

from both $\tilde{\eta}_D^{TFD}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$.

Here also, the trend of electronic effects [i.e., $+I$ and $+R$ (for C_6H_5- only) effects] exerted by the attached moieties on the $>C=O$ group is $C_6H_5- > CH_3CH_2CH_2- > CH_3CH_2- > CH_3-$. Thus, the expected trends of $C_{C=O}$ and $O_{C=O}$ reactivity in *Category II* are,

$$CH_3COCH_3 > CH_3CH_2COCH_3 > CH_3CH_2CH_2COCH_3 > C_6H_5COCH_3$$

(from $C_{C=O}$, towards a Nu^-)

and $C_6H_5COCH_3 > CH_3CH_2CH_2COCH_3 > CH_3CH_2COCH_3 > CH_3COCH_3$

(from $O_{C=O}$, towards an El^+).

Category III: A summary of the relative reactivity of $C_{C=O}$ in the chosen three systems in *Category III* (Table 3.1) is as follows:

$$C_6H_5CH=CHCHO > CH_3CH=CHCHO > CH_2=CHCHO$$

(from both $\tilde{\eta}_D^{TFD}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$).

And the overall observed trend of reactivity of $O_{C=O}$ towards an electron attracting species (electrophile, El^+) can be summarized as follows (Table 3.1):

$$CH_2=CHCHO > CH_3CH=CHCHO > C_6H_5CH=CHCHO$$

(from both $\tilde{\eta}_D^{TFD}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$).

In this case, the resonance effect plays a crucial role for predicting the reactivity sequence. The tend of resonance ($+R$) effect, exerted by the attached moieties on the $>C=O$, is, $C_6H_5CH=CH- > CH_3CH=CH- > CH_2=CH-$. Hence, the expected trends of reactivity of $C_{C=O}$ and $O_{C=O}$ in this category are,

$$CH_2=CHCHO > CH_3CH=CHCHO > C_6H_5CH=CHCHO$$

(from $C_{C=O}$, towards a Nu^-)

and $$C_6H_5CH=CHCHO > CH_3CH=CHCHO > CH_2=CHCHO$$

(from $O_{C=O}$, towards an El^+).

A careful analysis reveals the cause of unexpected intermolecular reactivity trends of the carbon and oxygen atoms (of carbonyl groups for all the three categories *I*, *II* and *III*) to be the $\frac{1}{N}$ term in the expressions of $\tilde{\eta}_D^{TFD}(k)$ or $\tilde{\eta}_D^{TFD'}(k)$. As the number of electrons increases with the size of the system, the $\frac{1}{N}$ factor alters the values of $\tilde{\eta}_D^{TFD}(k)$ or $\tilde{\eta}_D^{TFD'}(k)$ of $C_{C=O}$ and $O_{C=O}$, generating wrong trends of these parameters (this will be obvious from Table 3.3).

B. Problem of Evaluating the Local Hardness $\eta(\bar{r})$ Using $\lambda[\rho(\bar{r}')] = Nf(\bar{r}')$

Inserting Eq (3.4) in Eq (3.3) (where $\lambda[\rho(\bar{r}')]$ is replaced by $Nf(\bar{r}')$)[11-18] we get the frontier local hardness,

$$\tilde{\eta}_F^{TFD}(\bar{r}) = \frac{10}{9} C_F \int \rho_N^{\frac{1}{3}}(\bar{r}) \delta(\bar{r} - \bar{r}') f(\bar{r}') d\bar{r}' + \int \frac{1}{|\bar{r} - \bar{r}'|} f(\bar{r}') d\bar{r}'$$

$$- \frac{4}{9} C_X \int \rho_N^{\frac{2}{3}}(\bar{r}) \delta(\bar{r} - \bar{r}') f(\bar{r}') d\bar{r}' \qquad (3.9)$$

here, $f(\bar{r}')$ is the Fukui function,[24-27] (the local variation in the electron density of the system due to a change in the global number of electrons, for more details see chapter I) as,

$$f(\bar{r}) = \left(\frac{\partial^2 E}{\delta v(\bar{r}) \partial N} \right) = \left(\frac{\partial \rho(\bar{r})}{\partial N} \right)_{v(\bar{r})} \qquad (1.39)$$

Note that there is a discontinuity[24,28,29] in the derivative of the Fukui function just as there is for chemical potential.[30] When an electron is being added, one has $f^+(\bar{r})$; when it is being subtracted one has $f^-(\bar{r})$; one also has the average $f^0(\bar{r})$. Parr and Yang[24] have defined the left ($f^-(\bar{r})$), right ($f^+(\bar{r})$) and central ($f^0(\bar{r})$) derivatives of Eq (1.39). These three Fukui functions can be written by applying a finite difference approximation and the frontier-orbital theory of reactivity, as

$f^-(\bar{r}) \cong \rho_N(\bar{r}) - \rho_{N-1}(\bar{r})$ measures reactivity toward an electrophilic (El^+) reagent (1.40)

$f^+(\bar{r}) \cong \rho_{N+1}(\bar{r}) - \rho_N(\bar{r})$ measures reactivity toward a nucleophilic (Nu^-) reagent (1.41)

and

$f^0(\bar{r}) \cong \frac{1}{2}[\rho_{N+1}(\bar{r}) - \rho_{N-1}(\bar{r})]$ measures reactivity toward an innocuous (radical) reagent (1.42)

The condensed-to-atom Fukui functions are the Fukui indices (say for atom k) that are calculated using the procedure proposed by Yang and Mortier[31] as,

$$f^{-}(k) \cong P_N(k) - P_{N-1}(k) \tag{1.43}$$

$$f^{+}(k) \cong P_{N+1}(k) - P_N(k) \tag{1.44}$$

and $$f^{0}(k) \cong \frac{1}{2}[P_{N+1}(k) - P_{N-1}(k)] \tag{1.45}$$

(where $P(k)$ denotes the electronic population on atom k).

Applying similar idea conceived by Parr and Yang[24] to define Fukui function, we can split the local hardness $\tilde{\eta}_F^{TFD}(\bar{r})$ in two parts, $\tilde{\eta}_F^{TFD\ -}(\bar{r})$ and $\tilde{\eta}_F^{TFD\ +}(\bar{r})$, which measure reactivity toward an electrophilic reagent (i.e., El^+) and a nucleophilic reagent (i.e., Nu^-), respectively.

Although, $\tilde{\eta}_F^{TFD}(\bar{r})$[32] is equal to the global hardness η, at every point of space (Eq (1.69)), i.e.,

$$\eta = \tilde{\eta}_F^{TFD}(\bar{r})$$

it is not necessary that,

$$\eta = \tilde{\eta}_F^{TFD}(\bar{r}) = \tilde{\eta}_F^{TFD\ -}(\bar{r})$$

or $$\eta = \tilde{\eta}_F^{TFD}(\bar{r}) = \tilde{\eta}_F^{TFD\ +}(\bar{r})$$

Inserting Eq (1.40) in Eq (3.9) one can obtain the corresponding terms for the local hardness $\tilde{\eta}_F^{TFD\ -}(\bar{r})$:

from kinetic energy functional part,

$$\tilde{\eta}_K{}^{-}(\bar{r}) = \frac{10}{9} C_F \int \rho_N^{-\frac{1}{3}}(\bar{r}) \delta(\bar{r} - \bar{r}')[\rho_N(\bar{r}') - \rho_{N-1}(\bar{r}')] d\bar{r}'$$

$$\Rightarrow \tilde{\eta}_K{}^{-}(\bar{r}) = \frac{10}{9} C_F \rho_N^{\frac{2}{3}}(\bar{r}) - \frac{10}{9} C_F \rho_N^{-\frac{1}{3}}(\bar{r}) \rho_{N-1}(\bar{r}) \tag{3.10}$$

from Coulomb contribution part,

$$\tilde{\eta}_J^{-}(\bar{r}) = \int \frac{1}{|\bar{r} - \bar{r}'|}[\rho_N(\bar{r}') - \rho_{N-1}(\bar{r}')] d\bar{r}'$$

$$\Rightarrow \tilde{\eta}_J^{-}(\bar{r}) = -V_N^{el}(\bar{r}) + V_{N-1}^{el}(\bar{r}) \tag{3.11}$$

from exchange-correlation part,

$$\tilde{\eta}_X^{\,-}(\bar{r}) = -\frac{4}{9}C_X \int \rho_N^{-\frac{2}{3}}(\bar{r})\delta(\bar{r}-\bar{r}')[\rho_N(\bar{r}') - \rho_{N-1}(\bar{r}')]d\bar{r}'$$

$$\Rightarrow \tilde{\eta}_X^{\,-}(\bar{r}) = -\frac{4}{9}C_X \rho_N^{\frac{1}{3}}(\bar{r}) + \frac{4}{9}C_X \rho_N^{-\frac{2}{3}}(\bar{r})\rho_{N-1}(\bar{r}) \qquad (3.12)$$

Adding Eqs (3.10), (3.11) and (3.12) we get (for detail derivations of Eqs (3.10), (3.11) and (3.12) one can refer to appendix B),

$$\tilde{\eta}_F^{TFD\,-}(\bar{r}) = \tilde{\eta}_K^{\,-}(\bar{r}) + \tilde{\eta}_J^{\,-}(\bar{r}) + \tilde{\eta}_X^{\,-}(\bar{r})$$

$$\Rightarrow \tilde{\eta}_F^{TFD\,-}(\bar{r}) = \frac{10}{9}C_F \rho_N^{\frac{2}{3}}(\bar{r}) - \frac{10}{9}C_F \rho_N^{-\frac{1}{3}}(\bar{r})\rho_{N-1}(\bar{r}) - V_N^{el}(\bar{r}) + V_{N-1}^{el}(\bar{r})$$

$$-\frac{4}{9}C_X \rho_N^{\frac{1}{3}}(\bar{r}) + \frac{4}{9}C_X \rho_N^{-\frac{2}{3}}(\bar{r})\rho_{N-1}(\bar{r}) \qquad (3.13)$$

Similarly, one can obtain the corresponding terms for the local hardness $\tilde{\eta}_F^{TFD\,+}(\bar{r})$, by inserting Eq (1.41) in Eq (3.9),

from kinetic energy functional part,

$$\tilde{\eta}_K^{\,+}(\bar{r}) = \frac{10}{9}C_F \int \rho_N^{-\frac{1}{3}}(\bar{r})\delta(\bar{r}-\bar{r}')[\rho_{N+1}(\bar{r}') - \rho_N(\bar{r}')]d\bar{r}'$$

$$\Rightarrow \tilde{\eta}_K^{\,+}(\bar{r}) = \frac{10}{9}C_F \rho_{N+1}(\bar{r})\rho_N^{-\frac{1}{3}}(\bar{r}) - \frac{10}{9}C_F \rho_N^{\frac{2}{3}}(\bar{r}) \qquad (3.14)$$

from Coulomb contribution part,

$$\tilde{\eta}_J^{+}(\bar{r}) = \int \frac{1}{|\bar{r}-\bar{r}'|}[\rho_{N+1}(\bar{r}') - \rho_N(\bar{r}')]d\bar{r}'$$

$$\Rightarrow \tilde{\eta}_J^{+}(\bar{r}) = -V_{N+1}^{el}(\bar{r}) + V_N^{el}(\bar{r}) \qquad (3.15)$$

from exchange-correlation part,

$$\tilde{\eta}_X^{\,+}(\bar{r}) = -\frac{4}{9}C_X \int \rho_N^{-\frac{2}{3}}(\bar{r})\delta(\bar{r}-\bar{r}')[\rho_{N+1}(\bar{r}') - \rho_N(\bar{r}')]d\bar{r}'$$

$$\Rightarrow \tilde{\eta}_X^{\,+}(\bar{r}) = -\frac{4}{9}C_X \rho_{N+1}(\bar{r})\rho_N^{-\frac{2}{3}}(\bar{r}) + -\frac{4}{9}C_X \rho_N^{\frac{1}{3}}(\bar{r}) \qquad (3.16)$$

Hence $\tilde{\eta}_F^{TFD\ +}(\bar{r})$ is [adding Eqs (3.14), (3.15) and (3.16)] (detailed derivations of Eqs (3.14), (3.15) and (3.16) can be found in appendix B),

$$\tilde{\eta}_F^{TFD\ +}(\bar{r}) = \tilde{\eta}_K^{\ +}(\bar{r}) + \tilde{\eta}_J^{+}(\bar{r}) + \tilde{\eta}_X^{\ +}(\bar{r})$$

$$\Rightarrow \tilde{\eta}_F^{TFD\ +}(\bar{r}) = \frac{10}{9}C_F\rho_{N+1}(\bar{r})\rho_N^{-\frac{1}{3}}(\bar{r}) - \frac{10}{9}C_F\rho_N^{\frac{2}{3}}(\bar{r}) - V_{N+1}^{el}(\bar{r}) + V_N^{el}(\bar{r})$$

$$-\frac{4}{9}C_X\rho_{N+1}(\bar{r})\rho_N^{-\frac{2}{3}}(\bar{r}) + \frac{4}{9}C_X\rho_N^{\frac{1}{3}}(\bar{r}) \qquad (3.17)$$

The condensed-to-atom local hardness i.e., $\tilde{\eta}_F^{TFD\ -}(k)$ and $\tilde{\eta}_F^{TFD\ +}(k)$ can be written as

$$\tilde{\eta}_F^{TFD\ -}(k) = \frac{10}{9}C_F P_N^{\frac{2}{3}}(k) - \frac{10}{9}C_F P_N^{-\frac{1}{3}}(k)P_{N-1}(k) - V_N^{el}(k) + V_{N-1}^{el}(k)$$

$$-\frac{4}{9}C_X P_N^{\frac{1}{3}}(k) + \frac{4}{9}C_X P_N^{-\frac{2}{3}}(k)P_{N-1}(k) \qquad (3.18)$$

and

$$\tilde{\eta}_F^{TFD\ +}(k) = \frac{10}{9}C_F P_{N+1}(k)P_N^{-\frac{1}{3}}(k) - \frac{10}{9}C_F P_N^{\frac{2}{3}}(k) - V_{N+1}^{el}(k) + V_N^{el}(k)$$

$$-\frac{4}{9}C_X P_{N+1}(k)P_N^{-\frac{2}{3}}(k) + \frac{4}{9}C_X P_N^{\frac{1}{3}}(k) \qquad (3.19)$$

As mentioned in chapter I (see Ref 253 of chapter I), when we consider $\lambda[\rho(\bar{r}')] = Nf(\bar{r}')$ the local hardness $\tilde{\eta}_F^{TFD}(\bar{r})$ (at every point of space) becomes equal to global hardness η (Eq (1.69)).[32] However, in an approximate method (as is done in this study) when the Fukui function is computed using Kohn-Sham (KS) level of DFT (or HF level and other higher *ab initio* methods) and the local hardness is evaluated at the Thomas-Fermi-Dirac (TFD)[22] level (i.e., TFD approximation of $F[\rho]$ to determine $\eta(\bar{r},\bar{r}')$) it will not necessarily be a constant quantity everywhere in the molecule.[37a]

We have calculated $\tilde{\eta}_F^{TFD\ -}(k)$ and $\tilde{\eta}_F^{TFD\ +}(k)$ values of $O_{C=O}$ and $C_{C=O}$ and then compared the intermolecular reactivity trend of the chosen carbonyl compounds. The system for which $O_{C=O}$ shows the highest value of $\tilde{\eta}_F^{TFD\ -}(k)$ (i.e., $\tilde{\eta}_F^{TFD\ -}(O_{C=O})$) will be the most reactive one towards an electrophile (El^+). Similarly, the system having

the highest $\tilde{\eta}_F^{TFD\ +}(k)$ (i.e., $\tilde{\eta}_F^{TFD\ +}(C_{C=O})$) value on the $C_{C=O}$ is the most reactive towards a nucleophile (Nu^-).

Category I: The observed trend of reactivity of $O_{C=O}$ (towards El^+) can be summarized as follows (Table 3.2),

$$HCHO > CH_3CHO > CH_3CH_2CHO > C_6H_5CHO$$

from $\tilde{\eta}_F^{TFD\ -}(k)$.

And a summary of the relative reactivity of $C_{C=O}$ in the chosen systems is as follows (Table 3.2),

$$HCHO > CH_3CH_2CHO > CH_3CHO > C_6H_5CHO$$

from $\tilde{\eta}_F^{TFD\ +}(k)$.

Category II: A summary of the relative reactivity of $O_{C=O}$ of the chosen systems in *Category II* towards an electron attracting species (i.e., electrophile, El^+) is as follows (Table 3.2),

$$CH_3COCH_3 > CH_3CH_2COCH_3 > CH_3CH_2CH_2COCH_3 > C_6H_5COCH_3$$

from $\tilde{\eta}_F^{TFD\ -}(k)$.

And the observed overall trend of $C_{C=O}$ reactivity towards an electron donating species (nucleophile, Nu^-) can be summarized as follows (Table 3.2),

$$CH_3COCH_3 > CH_3CH_2COCH_3 > CH_3CH_2CH_2COCH_3 > C_6H_5COCH_3$$

from $\tilde{\eta}_F^{TFD\ +}(k)$.

Category III: A summary of the relative reactivity of $O_{C=O}$ of the chosen systems in *Category III* towards an electron attracting species (electrophile, El^+) is as follows (Table 3.2),

$$CH_2 = CHCHO > CH_3CH = CHCHO > C_6H_5CH = CHCHO$$

from $\tilde{\eta}_F^{TFD\ -}(k)$.

And, the observed overall trend of $C_{C=O}$ reactivity towards an electron donating species (nucleophile, Nu^-) can be summarized as follows (Table 3.2),

$$CH_2 = CHCHO > CH_3CH = CHCHO > C_6H_5CH = CHCHO$$

from $\tilde{\eta}_F^{TFD\ +}(k)$.

Here again the observed trends of reactivity for $O_{C=O}$ (for the categories *I*, *II* and *III*) are different from the expected ones [see subsection 3.3.A], although, these are much better for $C_{C=O}$ (except the unexpected trend between CH_3CHO and CH_3CH_2CHO in the *Category I* and that is also at fourth decimal point).

After critically analyzing our data, we understand that it is the Fukui function [$f(\bar{r})$] term which is altering the expected trend. The $f^-_{O_{C=O}}$ and $f^+_{C_{C=O}}$ values from fifth and sixth columns in Table 3.2 clearly confirm our claim. Although, using the composite function $\lambda[\rho(\bar{r}')] = Nf(\bar{r}')$, apparently we could remove the $\frac{1}{N}$ problem from the expression of local hardness $\tilde{\eta}_F^{TFD}(\bar{r})$ [Eq (3.9)], N-dependence is re-introduced indirectly through the normalization condition of Fukui function (i.e., $\int f(\bar{r})d\bar{r} = 1$ or $\sum_k^N f(k) = 1$).[22] This shows that because of the 'intensive' nature[33] of Fukui function[17] [$f(\bar{r})$] the local hardness parameter remain 'subintensive' (i.e., becomes smaller and smaller as the size of the system increases). It should also be mentioned in this context that 'intensive' nature of $f(\bar{r})$ or $f(k)$ makes local softness[26] [$s(\bar{r})$ or $s(k)$] and philicity[34] [$w(\bar{r})$ or $w(k)$] indices applicable to limited cases[35-41] of intermolecular reactivities even though, individually, these two descriptors are 'extensive' (i.e., do not go to zero in the thermodynamic limit) in nature.

3.4. How to Remove the N -dependence in Local Hardness $\eta(\bar{r})$?

From the discussion in Section 3.3. we can argue that none of the two composite functions (i.e., $\lambda[\rho(\bar{r}')] = \rho(\bar{r}')$ or $\lambda[\rho(\bar{r}')] = Nf(\bar{r}')$) is able to solve the $\frac{1}{N}$ problem in local hardness $\eta(\bar{r})$ parameter. It is also true that the broader applicability of $\eta(\bar{r})$ as a reliable intermolecular reactivity descriptor primarily depends on the removal of its $\frac{1}{N}$ dependence. This is also justified from the conceptual point of view. Here the comparison of intermolecular reactivity trends is mainly based on the local hardness values of those particular sites (or atoms in the condensed form), electronic or any other effects exerted by the rest of the system already incorporated. But this $\frac{1}{N}$ factor creates an impression as if the whole system does equally contribute to the reactivity of that particular site or atom. In reality, parts (or moieties) of the system, far from the site of interest, may have very little or no effect on the reactivity of that particular site. Geerlings and collaborators also raised a similar argument in some of their earlier studies.[42] So, the best way to incorporate the electronic (or any other) effects of the rest of the system, without overemphasizing $\frac{1}{N}$ factor, is to consider only the active site (or atoms or group) for which the number of electrons is same. For example, in the chosen systems if only $C = O$ moiety is considered N will be same (i.e., $N = 14$) for all the systems and we can use the modified form of Eqs (3.7) and (3.8) as,

$$\tilde{\eta}_D^{TFD}(k) = \frac{10}{9} C_F P(k)^{2/3} - V^{el}(k) - \frac{4}{9} C_X P(k)^{1/3} \tag{3.20}$$

$$\tilde{\eta}_D^{TFD'}(k) = -V^{el}(k) \tag{3.21}$$

By using Eqs (3.20) and (3.21) we observe the following trends (Table 3.3):

Category I:

$$HCHO > CH_3CHO > CH_3CH_2CHO > C_6H_5CHO$$

for the $C_{C=O}$ towards an electron donating species (nucleophile, Nu^-)

and $$C_6H_5CHO > CH_3CH_2CHO > CH_3CHO > HCHO$$

for the $O_{C=O}$ towards an electron attracting species (electrophile, El^+).

Category II:

$$CH_3COCH_3 > CH_3CH_2COCH_3 > CH_3CH_2CH_2COCH_3 > C_6H_5COCH_3$$

for the $C_{C=O}$ towards an electron donating species (nucleophile, Nu^-)

and $$C_6H_5COCH_3 > CH_3CH_2CH_2COCH_3 > CH_3CH_2COCH_3 > CH_3COCH_3$$

for the $O_{C=O}$ towards an electron attracting species (electrophile, El^+).

Category III:

$$CH_2 = CHCHO > CH_3CH = CHCHO > C_6H_5CH = CHCHO$$

for the $C_{C=O}$ towards an electron donating species (nucleophile, Nu^-),

and $$C_6H_5CH = CHCHO > CH_3CH = CHCHO > CH_2 = CHCHO$$

for the $O_{C=O}$ towards an electron attracting species (electrophile, El^+).

These are exactly similar to the expected ones [subsection 3.3. A.].

3.5. Conclusion

In the present chapter we have discussed the applicability of 'local hardness' parameter in predicting intermolecular reactivity sequences of carbonyl compounds. We have found that none of the two composite functions (i.e., $\lambda[\rho(\bar{r}')] = \rho(\bar{r}')$ or $\lambda[\rho(\bar{r}')] = Nf(\bar{r}')$) are able to solve the $\frac{1}{N}$ problem in local hardness $\eta(\bar{r})$ parameter. But, the broader applicability of $\eta(\bar{r})$ as a reliable intermolecular reactivity descriptor primarily depends on the removal of its $\frac{1}{N}$ dependence. This suggests that the best way to incorporate the electronic (or any other) effects of the rest of the system, without overemphasizing $\frac{1}{N}$ factor, is to consider only the active site (or atoms or group) for which the number of electrons is same.

Applying this idea we can calculate the preferable regioselectivity of large chemical and biological systems using the 'One-into-Many' model proposed earlier.[1, 2] In this model a large system is proposed to be broken into different fragments and in that way an intra-molecular problem of a large system can be re-casted into an inter-molecular problem of individual fragments. The fragmentation should be done in such a way that each part contains at least one reactive site. Then evaluating the local hardness of the active sites (or atoms) of the smaller fragments separately inter-molecular reactivity trends can be compared. However, if the fragments are of different sizes (i.e., having different number of electrons, N) the above model may not always be reliable in predicting site selectivity of large systems, even through the active sites (or moieties e.g., $-CHO$, $>C=O$ or $-NH_2$) in the individual fragments are identical. But the local hardness values of the active sites (or atoms), calculated through Eqs (3.20) and (3.21), seems to resolve this problem to a large extent because it considers the electronic effect exerted by the rest of the system without over-emphasizing the $\frac{1}{N}$ factor.

References

1. S. Saha and R. K. Roy, *J. Phys. Chem. B*, 2007, **111**, 9664.
2. S. Saha and R. K. Roy, *J. Phys. Chem. B*, 2008, **112**, 1884.
3. M. Berkowitz, S. K. Ghosh and R. G. Parr, *J. Am. Chem. Soc.*, 1985, **107**, 6811.
4. S. K. Ghosh and M. Berkowitz, *J. Chem. Phys.*, 1985, **83**, 2976.
5. P. K. Chattaraj, D. R. Roy, P. Geerlings and M. Torrent-Sucarrat, *Theor. Chem. Acc.*, 2007, **118**, 923.
6. W. Langenaeker, F. de Proft and P. Geerlings, *J. Phys. Chem.*, 1995, **99**, 6424.
7. P. Mignon, S. Loverix and P. Geerlings, *Chem. Phys. Lett.*, 2005, **401**, 40.
8. A. S. Ozen, F. De Proft, V. Aviyente and P. Geerlings, *J. Phys. Chem. A*, 2006, **110**, 5860.
9. A. Borgoo, M. Torrent-Sucarrat, F. D. Proft and P. Geerlings, *J. Chem. Phys.*, 2007, **126**, 234104.
10. P. Mignon, P. Geerlings and R. Schoonheydt, *J. Phys. Chem. C*, 2007, **111**, 12376.
11. S. K. Ghosh, *Chem. Phys. Lett.*, 1990, **172**, 77.
12. S. Liu, F. De Proft and R. G. Parr, *J. Phys. Chem.*, 1997, **101**, 6991.
13. M. Torrent-Sucarrat, M. Duran and M. Sola, *J. Phys. Chem. A*, 2002, **106**, 4632.
14. M. Torrent-Sucarrat and P. Geerlings, *J. Chem. Phys.*, 2006, **125**, 244101.
15. M. Torrent-Sucarrat, P. Salvador, P. Geerlings and M. Solà, *J. Comput. Chem.*, 2007, **28**, 574.
16. M. Torrent-Sucarrat, F. De Proft, P. Geerlings and Paul W. Ayers, *Chem. - Eur. J.*, 2008, **14**, 8652.
17. M. Torrent-Sucarrat, P. Salvador, M. Solà and P. Geerlings, *J. Comput. Chem.*, 2008, **29**, 1064.
18. M. Torrent-Sucarrat, F. D. Proft, P. W. Ayers and P. Geerlings, *Phys. Chem. Chem. Phys.* 2010, 1**2**, 1072.
19. L. H. Thomas, *Proc. Cambridge Philos. Soc.*, 1927, **23**, 542.
20. E. Fermi, *Z. Phys.*, 1928, **48**, 73.
21. P. A. M. Dirac, *Proc. Cambridge Philos. Soc.*, 1930, **26**, 376.

22. R. G. Parr and W. Yang, *Density–Functional Theory of Atoms and Molecules*, Oxford University Press, New York, 1989.
23. P. K. Chattaraj and S. Giri, *Annu. Rep. Prog. Chem., Sec C (Physical Chemistry)*, 2009, **105**, 13.
24. R. G. Parr and W. Yang, *J. Am. Chem. Soc.*, 1984, **106**, 4049.
25. W. Yang, R. G. Parr and R. Pucci, *J. Chem. Phys.*, 1984, **81**, 2862.
26. W. Yang and R. G. Parr, *Proc. Natl. Acad. Sci. U.S.A.*, 1985, **82**, 6723.
27. P. W. Ayers and M. Levy, *Theor. Chem. Acc.*, 2000, **103**, 353.
28. W. T. Yang, Y. K. Zhang and P. W. Ayers, *Phys. Rev. Lett.*, 2000, **84**, 5172.
29. P. W. Ayers, *J. Math. Chem.*, 2008, **43**, 285.
30. J. P. Perdew, R. G. Parr, M. Levy and J. L. Balduz, Jr., *Phys. Rev. Lett.*, 1982, **49**, 1691.
31. W. Yang and W. J. Mortier, *J. Am. Chem. Soc.*, 1986, **108**, 5708.
32. M. K. Harbola, P. K. Chattaraj and R. G. Parr, *Israel J. Chem.*, 1991, **31**, 395.
33. P. W. Ayers, C. Morell, F. De Proft and P. Geerlings, *Chem. Eur. J*, 2007, **13**, 8240.
34. P. K. Chattaraj, B. Maiti and U. Sarkar, *J. Phys. Chem. A*, 2003, **107**, 4973.
35. D. Babic and N. Trinajstic, *J. Mol. Struct. (THEOCHEM)*, 1994, **120**, 321.
36. R. K. Roy, *J. Phys. Chem. A*, 2004, **108**, 4934.
37. A. Aizman, R. Contreras and P. Perez, *Tetrahedron*, 2005, **61**, 889.
38. R. K. Roy, V. Usha, J. Paulovič and K. Hirao, *J. Phys. Chem. A*, 2005, **109**, 4601.
39. R. K. Roy, P. Bagaria, S. Naik, V. Kavala and B. K. Patel, *J. Phys. Chem. A*, 2006, **110**, 2181.
40. R. K. Roy, V. Usha, B. K. Patel and K. Hirao, *J. Comput. Chem.*, 2006, **27**, 773.
41. P. Bagaria and R. K. Roy, *J. Phys. Chem. A*, 2008, **112**, 97.
42. P. Mignon, S. Loverix, F. De Proft and P. Geerlings, *J. Phys. Chem. A*, 2004, **108**, 6038.

Table 3.1. $\widetilde{\eta}_D^{TFD}(k)$ **and** $\widetilde{\eta}_D^{TFD'}(k)$ **values (in a.u.) of** $C_{C=O}$ **and** $O_{C=O}$ **in MP2/6-31G** in the carbonyl compounds (for both** $C_{C=O}$ **and** $O_{C=O}$ **).**

Category	Systems	Atoms	Total number of electrons (N) of the system	$\widetilde{\eta}_D^{TFD}(k)$	$\widetilde{\eta}_D^{TFD'}(k)$
I	**HCHO**	$C_{C=O}$	16	1.7940	1.1904
		$O_{C=O}$	16	2.3742	1.5886
	CH_3CHO	$C_{C=O}$	24	1.2883	0.8925
		$O_{C=O}$	24	1.6521	1.1264
	CH_3CH_2CHO	$C_{C=O}$	32	1.0147	0.7182
		$O_{C=O}$	32	1.2751	0.8807
	C_6H_5CHO	$C_{C=O}$	56	0.6405	0.4714
		$O_{C=O}$	56	0.7818	0.5559
II	**CH_3COCH_3**	$C_{C=O}$	32	1.0347	0.7423
		$O_{C=O}$	32	1.2914	0.8958
	$CH_3CH_2COCH_3$	$C_{C=O}$	40	0.8661	0.6323
		$O_{C=O}$	40	1.0696	0.7530
	$CH_3CH_2CH_2COCH_3$	$C_{C=O}$	48	0.7431	0.5483
		$O_{C=O}$	48	0.9107	0.6468
	$C_6H_5COCH_3$	$C_{C=O}$	64	0.5936	0.4481
		$O_{C=O}$	64	0.7112	0.5130
III	**CH_2=CHCHO**	$C_{C=O}$	30	1.0695	0.7527
		$O_{C=O}$	30	1.3584	0.9369
	CH_3CH=CHCHO	$C_{C=O}$	38	0.8716	0.6217
		$O_{C=O}$	38	1.0971	0.7640
	C_6H_5CH=CHCHO	$C_{C=O}$	70	0.5159	0.3802
		$O_{C=O}$	70	0.6350	0.4541

Table 3.2: $\tilde{\eta}_F^{TFD\ -}(O_{C=O})$, $\tilde{\eta}_F^{TFD\ +}(C_{C=O})$, $f_{O_{C=O}}^{-}$ **and** $f_{C_{C=O}}^{+}$ **values (in a.u.) in MP2/6-31G** for carbonyl compounds (for details see the text).**

Category	Systems	Atoms	$\tilde{\eta}_F^{-}(O_{C=O})$	$\tilde{\eta}_F^{+}(C_{C=O})$	$f_{O_{C=O}}^{-}$	$f_{C_{C=O}}^{+}$
I	**HCHO**	$C_{C=O}$	-	0.8835	-	0.2946
		$O_{C=O}$	1.2150	-	0.5112	-
	CH_3CHO	$C_{C=O}$	-	0.8549	-	0.2826
		$O_{C=O}$	1.1511	-	0.4863	-
	CH_3CH_2CHO	$C_{C=O}$	-	0.8561	-	0.2905
		$O_{C=O}$	1.1256	-	0.4756	-
	C_6H_5CHO	$C_{C=O}$	-	0.5565	-	0.1674
		$O_{C=O}$	1.0168	-	0.4345	-
II	**CH_3COCH_3**	$C_{C=O}$	-	0.7799	-	0.2456
		$O_{C=O}$	1.1076	-	0.4697	-
	$CH_3CH_2COCH_3$	$C_{C=O}$	-	0.7748	-	0.2465
		$O_{C=O}$	1.0940	-	0.4658	-
	$CH_3CH_2CH_2COCH_3$	$C_{C=O}$	-	0.7740	-	0.2477
		$O_{C=O}$	1.0897	-	0.4651	-
	$C_6H_5COCH_3$	$C_{C=O}$	-	0.5287	-	0.1560
		$O_{C=O}$	1.0055	-	0.4303	-
III	**CH_2=CHCHO**	$C_{C=O}$	-	0.5619	-	0.1565
		$O_{C=O}$	1.0857	-	0.4605	-
	CH_3CH=CHCHO	$C_{C=O}$	-	0.5600	-	0.1594
		$O_{C=O}$	1.0611	-	0.4519	-
	C_6H_5CH=CHCHO	$C_{C=O}$	-	0.3961	-	0.0939
		$O_{C=O}$	0.9938	-	0.4280	-

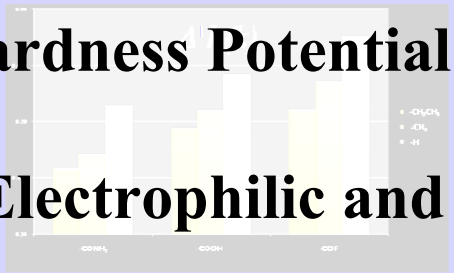
Table 3.3: $\widetilde{\eta}_D^{TFD}(k)$ and $\widetilde{\eta}_D^{TFD'}(k)$ values (in a.u.) in MP2/6-31G for carbonyl compounds (for details see the text).**

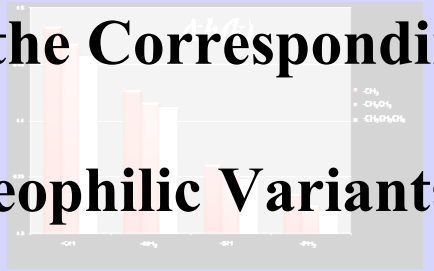
Category	Systems	Atoms	$\widetilde{\eta}_D^{TFD}(k)$	$\widetilde{\eta}_D^{TFD}(k)$	$\widetilde{\eta}_D^{TFD'}(k)$	$\widetilde{\eta}_D^{TFD'}(k)$
I	**HCHO**	$C_{C=O}$	28.7037	-	19.0467	-
		$O_{C=O}$	-	37.9869	-	25.4180
	CH_3CHO	$C_{C=O}$	30.9198	-	21.4203	-
		$O_{C=O}$	-	39.6501	-	27.0348
	CH_3CH_2CHO	$C_{C=O}$	32.4693	-	22.9813	-
		$O_{C=O}$	-	40.8028	-	28.1819
	C_6H_5CHO	$C_{C=O}$	35.8700	-	26.3966	-
		$O_{C=O}$	-	43.7817	-	31.1301
II	**CH_3COCH_3**	$C_{C=O}$	33.1102	-	23.7543	-
		$O_{C=O}$	-	41.3237	-	28.6660
	$CH_3CH_2COCH_3$	$C_{C=O}$	34.6444	-	25.2920	-
		$O_{C=O}$	-	42.7837	-	30.1191
	$CH_3CH_2CH_2COCH_3$	$C_{C=O}$	35.6700	-	26.3165	-
		$O_{C=O}$	-	43.7121	-	31.0466
	$C_6H_5COCH_3$	$C_{C=O}$	37.9913	-	28.6785	-
		$O_{C=O}$	-	45.5197	-	32.8330
III	**CH_2=CHCHO**	$C_{C=O}$	32.0852	-	22.5802	-
		$O_{C=O}$	-	40.7510	-	28.1078
	CH_3CH=CHCHO	$C_{C=O}$	33.1223	-	23.6243	-
		$O_{C=O}$	-	41.6901	-	29.0330
	C_6H_5CH=CHCHO	$C_{C=O}$	36.1113	-	26.6135	-
		$O_{C=O}$	-	44.4480	-	31.7875

Chapter IV

$$\frac{\delta \mu(r)}{\delta \rho(\bar{r})} = \eta(r) = N\eta(\bar{r})$$

Hardness Potential and the Corresponding Electrophilic and Nucleophilic Variants





4.1. Introduction

In the previous chapter, we have illustrated the limitation of $\eta(\bar{r})$ (evaluated from the two composite functions i.e., $\lambda[\rho(\bar{r}')] = \rho(\bar{r}')$ [1-8] and $\lambda[\rho(\bar{r}')] = Nf(\bar{r}')$ [9-16]) when used for comparison of intermolecular reactivity trends between systems of different sizes but having common reactive centers.[17] After a careful analysis it was revealed that as the number of electrons increases (with increasing size of the system), the $\frac{1}{N}$ factor alters the expected trends of $\tilde{\eta}_D^{TFD}(\bar{r})$ or $\tilde{\eta}_D^{TFD'}(\bar{r})$ (i.e., when the composite function $\lambda[\rho(\bar{r}')] = \rho(\bar{r}')$) values. It was also shown that when the composite function, $\lambda[\rho(\bar{r}')] = Nf(\bar{r}')$, although the $\frac{1}{N}$ problem apparently is solved, N-dependence appears implicitly through the normalization condition of the Fukui function. So, the broader applicability of $\eta(\bar{r})$ as a reliable intermolecular reactivity descriptor necessitates the removal of its N-dependence. This is because the comparison of intermolecular reactivity trends is mainly based on the local hardness values of those particular sites (or atoms in the condensed form), electronic or any other effects exerted by the rest of the system having already been incorporated. But the $\frac{1}{N}$ factor creates an impression as if the whole system does equally contribute to the reactivity of that particular site or atom. In reality, parts (or moieties) of the system, far from the site of interest, may have very little or no effect on the reactivity of that particular site. Geerlings and collaborators also raised a similar argument in some of their earlier studies.[18] Therefore, the best way to incorporate the electronic (or other) effects of the rest of the system, without overemphasizing $\frac{1}{N}$ factor, is to consider only the active site for which the number of electrons is same. Thus the modified form of Eqs (3.5) and (3.6) could be written as

$$\tilde{\eta}_D^{TFD}(\bar{r}) = \frac{10}{9} C_F \rho(\bar{r})^{2/3} - V^{el}(\bar{r}) - \frac{4}{9} C_X \rho(\bar{r})^{1/3} \qquad (3.20a)$$

and

$$\tilde{\eta}_D^{TFD'}(\bar{r}) = -V^{el}(\bar{r}) \qquad (3.21b)$$

For example we have shown that if only $C = O$ moiety is considered (N will be same, i.e., $N = 14$, for all the chosen carbonyl systems) and one can use the modified 'condensed-to-atom'[7, 8, 17] form of Eqs (3.20a) and (3.21b)) as,

$$\tilde{\eta}_D^{TFD}(k) = \frac{10}{9} C_F P(k)^{2/3} - V^{el}(k) - \frac{4}{9} C_X P(k)^{1/3} \tag{3.20}$$

$$\tilde{\eta}_D^{TFD'}(k) = -V^{el}(k) \tag{3.21}$$

However, we could not provide any analytical justification of resolving the N-dependence problem of local hardness in the way mentioned above. In the present chapter it is shown how this limitation is eliminated by practical implementation of two variants of a different reactivity descriptor i.e., hardness potential $h(\bar{r})$.[19]

4.2. Theoretical Background

Starting from the definition of hardness potential $h(\bar{r})$,[19] proposed by Parr and Gázquez, we can write

$$h(r) = \int \frac{\delta^2 F[\rho]}{\delta\rho(\bar{r})\delta\rho(\bar{r}')} \rho(\bar{r}')d\bar{r}' = N\eta(\bar{r}) \quad \text{(from Eqs (1.57) and (1.58))} \tag{4.1}$$

It is important to note that unlike local hardness ($\eta(\bar{r})$), $h(\bar{r})$ does not contain $\frac{1}{N}$ factor, so it is expected that $h(\bar{r})$ can resolve the N-dependence problem.

To obtain complete mathematical definition of $h(\bar{r})$, one needs to approximate the Hohenberg-Kohn functional $F[\rho(\bar{r})]$.[20] $F[\rho(\bar{r})]$ can be approximated based on the Thomas-Fermi-Dirac (TFD)[46-48] approach to DFT. If we neglect the nucleus-electron attraction in $F[\rho(\bar{r})]$, the following equation is obtained from the general form of the energy functional $E^{TFD}[\rho(\bar{r})]$,[24] without further approximations:

$$F_E^{TFD}[\rho(\bar{r})] = C_F \int \rho(\bar{r})^{5/3} d\bar{r} + \frac{1}{2} \iint \frac{\rho(\bar{r})\rho(\bar{r}')}{|\bar{r} - \bar{r}'|} d\bar{r}' d\bar{r} - C_X \int \rho(\bar{r})^{4/3} d\bar{r}$$

(4.2)

Here, $C_F = \frac{3}{10}\left(3\pi^2\right)^{\frac{2}{3}} = 2.8712$ and $C_X = \frac{3}{4}\left(\frac{3}{\pi}\right)^{\frac{1}{2}} = 0.7386$ are the coefficients of the kinetic energy and exchange-energy functionals, respectively.[24]

Inserting Eq (4.2) in Eq (4.1), one may find the expression of hardness potential, $h(\bar{r})$ as,

$$h(\bar{r}) = \frac{10}{9} C_F \rho(\bar{r})^{2/3} - V^{el}(\bar{r}) - \frac{4}{9} C_X \rho(\bar{r})^{1/3} \tag{4.3}$$

Like $\eta(\bar{r})$,[2,7,8,17] the hardness potential, $h(\bar{r})$ can be evaluated though Eq (4.3). Since the electronic part of the molecular electrostatic potential,[19] $V^{el}(\bar{r})$, is usually dominant when compared to the other two terms (i.e., kinetic and exchange energy terms).[25] Moreover, the negative sign in the third term (i.e., the exchange energy term) cancels out, to some extent, the effect of the first term (i.e., the kinetic energy term). Hence, one may further approximate $h(\bar{r})$ as,

$$h(\bar{r}) = -V^{el}(\bar{r}) \tag{4.4}$$

Instead of evaluating electronic contribution to the molecular electrostatic potential at a specific site (i.e., $V^{el}(\bar{r})$) we will consider the condensed value on a particular atom k (i.e., $V^{el}(k)$), so, $h(\bar{r})$ is replaced by $h(k)$ as,

$$h(k) = -V^{el}(k) \tag{4.5}$$

It is worth mentioning here that because $h(\bar{r})$ (in Eq (4.3)) contains nonlinear functions of $\rho(\bar{r})$ (the first and the third terms), the condensation of $h(\bar{r})$ (as it is in Eq (4.3)) is mathematically inexact.[26] However, earlier studies in chapters III and IV reveal that a condensed-to-atom local hardness (i.e., $\eta(k)$), including all three terms, provides a reasonable good correlation with the experimental findings of biological systems as well as chemical systems.[7,8,17]

4.3. Methodology

To implement Eq (4.5) for evaluating 'condensed hardness potential', we chose homologous series of chemical systems containing some common functional groups, *viz*, *–COOH, –COF, –CONH₂, –OH, –SH, –NH₂, –PH₂*. The carbon atom in the first three functional groups (i.e., *–COOH, –COF* and *–CONH₂*) is expected to be the most reactive site toward an electron donating species (i.e., nucleophile, Nu^-) whereas *O*, *S*, *N*, *P* atoms in the *–OH*, *–SH*, *–NH₂*, *–PH₂* are, respectively, the most reactive sites toward an electron accepting species (i.e., electrophile, El^+). Thus, based on the nature of the functional groups the chosen chemical systems can be grouped as electrophiles (systems having functional groups *–COOH, –COF* and *–CONH₂*) and nucleophiles (systems having functional groups *–OH*, *–SH*, *–NH₂*, *–PH₂*). The homologous series generated from above functional groups are as follows:

Category A: Series generated from carboxylic acid and its derivatives

(i) ***H*****C***OOH, CH₃****C***OOH, CH₃****C***H₂COOH*

(ii) ***H*****C***OF, CH₃****C***OF, CH₃CH₂****C***OF*

(iii) ***H*****C***ONH₂, CH₃****C***ONH₂, CH₃CH₂****C***ONH₂*

The individual members of the generated series differ in the number of electrons, N, making them suitable choice to test the N-dependence problem of $h(k)$. We will be interested in assessing the $h(k)$ values of the bold-faced carbon atoms. From Eq (4.5) it is clear that $h(k)$ is a positive quantity (because $V^{el}(k)$ is a negative quantity). Thus, it can be argued that in a particular series the system for which the bold-faced carbon atom has the lowest $h(k)$ value will be the most reactive one towards a nucleophile (Nu^-), i.e., that system will be the most electrophilic in nature.

***Category B*:** Series generated from nucleophilic functional groups:

(i) $CH_3\mathbf{O}H$, $CH_3CH_2\mathbf{O}H$, $CH_3CH_2CH_2\mathbf{O}H$

(ii) $CH_3\mathbf{S}H$, $CH_3CH_2\mathbf{S}H$, $CH_3CH_2CH_2\mathbf{S}H$

(iii) $CH_3\mathbf{N}H_2$, $CH_3CH_2\mathbf{N}H_2$, $CH_3CH_2CH_2\mathbf{N}H_2$

(iv) $CH_3\mathbf{P}H_2$, $CH_3CH_2\mathbf{P}H_2$, $CH_3CH_2CH_2\mathbf{P}H_2$

Thus we see that general forms of individual members belonging to these series are either $CH_3(CH_2)_n\mathbf{X}H$ or $CH_3(CH_2)_n\mathbf{Y}H_2$ (here $n = 0, 1, 2$; $X = O, S$ and $Y = N, P$). As the electronegative atom (printed in bold) is the atom of our interest; we will systematically examine how its $h(k)$ value varies with n. On the basis of calculated $h(k)$ (i.e., Eq (4.5)) values of these reactive sites the intermolecular reactivity trend of the chosen chemical systems can be compared. In a series the system for which the electronegative atom (printed in bold) shows the highest value of $h(k)$ (because $h(k) = -V^{el}(k)$ and $V^{el}(k)$ is a negative quantity) will be the most reactive one towards an electrophile (El^+).

Geometry optimizations as well as subsequent single point calculations of these compounds are carried out at MP2(FC)/6-31G** (here 'FC' stands for 'frozen core') and B3LYP/6-31G** levels. All the calculations (i.e., evaluation of $V^{el}(k)$) have been performed using Gaussian-98 program.[27] We confirmed that there is no imaginary frequency at any of the optimized geometries.

4.4. Results Generated by *h(k)*

Category A: The main factor, possibly the only factor, which causes difference in the calculated $h(k)$ values among the individual members belonging to a particular series in this category is the difference in inductive effect (i.e., $+I$ effect). As the number of intermediate $-CH_2-$ moiety increases, the electron density on electrophilic centres (i.e., bold-faced carbon atoms) should also increase resulting in higher $h(k)$ values. Conversely, decreasing electron density on these electrophilic carbon atoms lowers the $\left|V^{el}(k)\right|$ value and hence, $h(k)$ value also decreases. Based on this argument the expected intermolecular reactivity (i.e., electrophilicity) orders within a particular series (as obtained by comparing $h(k)$ values) should be as follows,

(i) $H\mathbf{C}OOH > CH_3\mathbf{C}OOH > CH_3CH_2\mathbf{C}OOH$;

(ii) $H\mathbf{C}OF > CH_3\mathbf{C}OF > CH_3CH_2\mathbf{C}OF$;

(iii) $H\mathbf{C}ONH_2 > CH_3\mathbf{C}ONH_2 > CH_3CH_2\mathbf{C}ONH_2$;

The generated $h(k)$ values satisfy this expected trend at both MP2 and B3LYP levels (Table 4.1.a).

However, irregularities are observed when systems belonging to different series of category A (and having the alkyl moiety same) are compared. For example, on the basis of electron donation power exerted by *–OH*, *–F* and $–NH_2$ groups on the *C=O* moiety, we expect the trend of reactivity (electrophilicity) of $\mathbf{C}_{C=O}$ to be as,

$R\mathbf{C}OF > R\mathbf{C}OOH > R\mathbf{C}ONH_2$; where $R = –H, –CH_3, –CH_2CH_3$

The hardness potential $h(k)$ gives results contrary to our expectation for these molecules in both the methods (Table 4.1.a).

***Category B*:** In general terms nucleophilicity can be appreciated by considering the availability of the electrons in the nucleophile. More available the electrons are, more nucleophilic the corresponding atom is. When 'condensed hardness potential' is evaluated using Eq (4.5) the electron density, indirectly (i.e., through $V^{el}(k)$), controls the reactivity of the chosen compound. The systems chosen in a particular series differ from each other by the presence of an extra intervening $–CH_2–$ moiety. When the number of such $–CH_2–$ moieties increases, the electron density on the electronegative atom (in the functional group) is expected to increase (because of increased electron donation tendency, i.e., electron pushing or *+I* effect) resulting in an increase in $h(k)$ values. Thus, expected trends of $h(k)$ values on the electronegative atom (printed in bold) for systems in different series should be as follows,

(i) $CH_3\mathbf{O}H < CH_3CH_2\mathbf{O}H < CH_3CH_2CH_2\mathbf{O}H$;

(ii) $CH_3\mathbf{S}H < CH_3CH_2\mathbf{S}H < CH_3CH_2CH_2\mathbf{S}H$;

(iii) $CH_3\mathbf{N}H_2 < CH_3CH_2\mathbf{N}H_2 < CH_3CH_2CH_2\mathbf{N}H_2$;

(iv) $CH_3\mathbf{P}H_2 < CH_3CH_2\mathbf{P}H_2 < CH_3CH_2CH_2\mathbf{P}H_2$.

The generated trends of $h(k)$ values are as expected in both MP2(FC)/6-31G** and B3LYP/6-31G** methods (see Table 4.1.b).

However, as in category A (see above), here also generated trends of $h(k)$ values are not as per expectation when systems belonging to different series are compared. Specifically, hardness potential (i.e., $h(k)$) fails to follow the expected trend for $CH_3(CH_2)_n\mathbf{O}H$ and $CH_3(CH_2)_n\mathbf{N}H_2$ and for $CH_3(CH_2)_n\mathbf{S}H$ and $CH_3(CH_2)_n\mathbf{P}H_2$. The expected trends of $h(k)$, based on the nucleophilicity of the functional groups should be as follows:

$CH_3(CH_2)_n\mathbf{O}H < CH_3(CH_2)_n\mathbf{N}H_2$; where, $n = 0, 1, 2$

and $CH_3(CH_2)_n\mathbf{S}H < CH_3(CH_2)_n\mathbf{P}H_2$; where, $n = 0, 1, 2$

The argument in favour of the above expected trend is that because of lower electronegativity of the *N* atom (when compared to that of *O* atom), the $-NH_2$ group will be more nucleophilic than $-OH$ group. Similarly, on the basis of reduced electronegativity of *P* atom (than that of *S* atom), it is expected that $-PH_2$ group is more reactive (i.e., nucleophilic) than $-SH$ group. The trends we observe, however, are exactly opposite to what we have just hypothesized (Table 4.1.b).

4.5. A Critical Analysis of the Observed Results and Proposition of Electrophilic and Nucleophilic Hardness Potentials

The results from the previous section show that trends of electrophilicity (in **Category A** systems) and nucleophilicity (in **Category B** systems) generated by $h(k)$ values of the strongest electrophilic and nucleophilic atoms, respectively, are as expected when the systems belong to the same homologous series. However, irregularities are observed when systems belonging to different series (in either category) are compared. A probable reason of this unexpected trend may be the fact that $h(k)$ can not take into account the difference of effective electronegativities of the concerned electrophilic and nucleophilic atoms in systems belonging to different series (either in **Category A** or **B**). For example, effective electronegativity of the $\mathbf{C}_{C=O}$ in *HCOOH* and *HCOF* will be different. Similarly, in **Category B** series the electronegativity values of *O*, *S* and *N* atoms in $CH_3\mathbf{O}H$, $CH_3\mathbf{S}H$ and $CH_3\mathbf{N}H_2$ are different. As defined by Eq (4.5), $h(k)$ is a property of the system, say *M,* in neutral state; it reflects only the charge distribution of *M* in that

state. The changes that occur in *M* as it begins to interact (e.g., polarization and charge transfer) with some approaching species can not be taken care by $h(k)$ itself. In reality this approaching electrophile (or nucleophile) induces polarization of the electron density of the concerned target species. But this effect is not taken into account here since $h(\bar{r})$ (or $h(k)$) corresponds to the electron density of the unperturbed neutral system.

To take care of the response of the system (in terms of changing electron density) as an electrophile or nucleophile approaches the system, we can invoke an idea similar to the one conceived by Parr and Yang[28] to define electrophilic and nucleophilic Fukui function (vide Eqs (1.40) and (1.41) in chapter I). We would like to propose two variants of $h(k)$ and denote them as $\Delta^{-}h(k)$ and $\Delta^{+}h(k)$, which measure reactivities toward an approaching electrophilic (i.e., El^{+}) and nucleophilic (i.e., Nu^{-}) reagent, respectively. This seems to be particularly rational here because the response of reactive centres, having different effective electronegativities, towards an approaching electrophile or nucleophile will be different.

Similar to the approximate condensed atomic Fukui function indices,[29] one may write

(i) ***condensed hardness potential descriptor for studies of nucleophilic attack on the system***

$$\Delta^{+}h(k) = h_{N+1}(k) - h_N(k)$$

or, $$\Delta^{+}h(k) = -V^{el}_{N+1}(k) - [-V^{el}_{N}(k)] \qquad \text{(using Eq (4.5))}$$

or, $$\Delta^{+}h(k) = -[V^{el}_{N+1}(k) - V^{el}_{N}(k)] \qquad (4.6)$$

(ii) ***condensed hardness potential descriptor for studies of electrophilic attack on the system***

$$\Delta^{-}h(k) = h_N(k) - h_{N-1}(k)$$

or, $$\Delta^{-}h(k) = -V^{el}_{N}(k) - [-V^{el}_{N-1}(k)] \qquad \text{(using Eq (4.5))}$$

or, $$\Delta^{-}h(k) = -[V^{el}_{N}(k) - V^{el}_{N-1}(k)] \qquad (4.7)$$

4.6. Interpretation of Reactivities Based on Electrophilic [Eq (4.6)] and Nucleophilic [Eq (4.7)] Hardness Potentials

A. (i) *Intermolecular Electrophilicity Trends for Systems Belonging to the Same Homologous Series in Category A*

In this case when a nucleophile (Nu^-) attacks the electrophilic centre (i.e., $\boldsymbol{C}_{C=O}$) the electron density increases on it resulting in higher $V_{N+1}^{el}(k)$ value (numerically) than that of $V_N^{el}(k)$. So the quantity inside the third-bracket in Eq (4.6) will be negative (as $V^{el}(k)$ is a negative quantity) making $\Delta^+h(k)$ a positive quantity. Now as the number of intervening $-CH_2-$ moiety increases electron density (and hence $V_{N+1}^{el}(k)$) on the $\boldsymbol{C}_{C=O}$ does not enhance as expected in the event of an approaching nucleophile because of the existence of already pushing $+I$ effect. So, as the size of the alkyl moiety increases the quantity inside the third-bracket in Eq (4.6) becomes less negative and so $\Delta^+h(k)$ less positive. This means as the size of the system in a particular homologous series increases the electrophilicity of $\boldsymbol{C}_{C=O}$ decreases and so also does $\Delta^+h(k)$ values. From Table 4.2 (a) and Figures 4.1.(a) & (b) we observe the trend of $\Delta^+h(k)$ values as,

(i) $CH_3CH_2\boldsymbol{C}OOH < CH_3\boldsymbol{C}OOH < H\boldsymbol{C}OOH$;

(ii) $CH_3CH_2\boldsymbol{C}OF < CH_3\boldsymbol{C}OF < H\boldsymbol{C}OF$;

(iii) $CH_3CH_2\boldsymbol{C}ONH_2 < CH_3\boldsymbol{C}ONH_2 < H\boldsymbol{C}ONH_2$;

which is just as expected.

(ii) *Intermolecular Electrophilicity Trends for Systems Belonging to Different Homologous Series in Category A*

Here reactivity (i.e., electrophilicity) difference of $\boldsymbol{C}_{C=O}$ arises due to the difference of electron withdrawing power of attached groups (i.e., of $-OH$, $-F$ and $-NH_2$). As electron withdrawing power changes in the sequence $-F > -OH > -NH_2$ (in fact the last two groups are electron pushing), the relative increase of electron density on $\boldsymbol{C}_{C=O}$ in the $N+1$ electron system (i.e., assuming complete one electron transfer from the approaching nucleophile to the electrophile) also follows the same trend. Thus numerical value of $V_{N+1}^{el}(k)$ for $\boldsymbol{C}_{C=O}$ will follow the trend,

$R\mathbf{C}OF > R\mathbf{C}OOH > R\mathbf{C}ONH_2$; where $R = -H, -CH_3, -CH_2CH_3$

Considering the negative sign of $V^{el}(k)$ from Eq (4.6) we get the trend of positive $\Delta^{+}h(k)$ values of $\mathbf{C}_{C=O}$ as,

$R\mathbf{C}OF > R\mathbf{C}OOH > R\mathbf{C}ONH_2$; where $R = -H, -CH_3, -CH_2CH_3$

Results from Table 4.2 (a) (Figures 4.1.(a) & (b)) confirm the above trend in both the methods (with the sole exception that in B3LYP/6-31G** method the $\Delta^{+}h(k)$ value of $CH_3CH_2\mathbf{C}OF$ is lower than that of $CH_3CH_2\mathbf{C}OOH$, see Table 4.2.(a) and Figure 4.1.(b)). Thus, higher is the $\Delta^{+}h(k)$ value of $\mathbf{C}_{C=O}$, higher is its reactivity (i.e., electrophilicity) in systems belonging to different homologous series.

B. (i) ***Intermolecular Nucleophilicity Trends for Systems Belonging to the Same Homologous Series In Category B***

During an El^{+} attack electron density over the nucleophilic atom (printed in bold) is normally reduced. As a result the $V^{el}(k)$ value (numerical) of the nucleophilic atom is expected to decrease after transfer of some electron density to the electrophile. Now considering the fact that $V^{el}(k)$ is a negative quantity and assuming a full one electron transfer from nucleophile to electrophile we get a positive value of $\Delta^{-}h(k)$ from Eq (4.7). However, this reduction in electron density on the nucleophilic atom (printed in bold) is increasingly compensated as the number of intervening $-CH_2-$ moiety increases (because of increasing $+I$ effect). So, numerically $V_{N-1}^{el}(k)$ value of the nucleophilic atom increases in a particular homologous series as, *e.g.*,

$CH_3CH_2CH_2\mathbf{O}H > CH_3CH_2\mathbf{O}H > CH_3\mathbf{O}H$

This is true for other homologous series also in **Category B**. Here it is worth mentioning that with increasing number of intervening $-CH_2-$ moiety both $V_N^{el}(k)$ and $V_{N-1}^{el}(k)$ of the nucleophilic atom increase. But what exactly happens is that the relative increase of $V_{N-1}^{el}(k)$ is more than that of $V_N^{el}(k)$. This is understandable also because the nucleophilic atom loses maximum electron to the approaching electrophile and hence compensation is also more with increasing $+I$ effect of the intervening $-CH_2-$ moieties. Now considering the negative sign of $V^{el}(k)$ from Eq (4.7) we can argue that lower the

positive value of $\Delta^{-}h(k)$ higher is the nucleophilicity of the corresponding nucleophilic atom. Based on this argument the expected $\Delta^{-}h(k)$ values of the nucleophilic atom of systems belonging to different homologous series of **Category B** should be as follows,

(i) $CH_3CH_2CH_2\mathbf{O}H < CH_3CH_2\mathbf{O}H < CH_3\mathbf{O}H$

(ii) $CH_3CH_2CH_2\mathbf{S}H < CH_3CH_2\mathbf{S}H < CH_3\mathbf{S}H$;

(iii) $CH_3CH_2CH_2\mathbf{N}H_2 < CH_3CH_2\mathbf{N}H_2 < CH_3\mathbf{N}H_2$;

(iv) $CH_3CH_2CH_2\mathbf{P}H_2 < CH_3CH_2\mathbf{P}H_2 < CH_3\mathbf{P}H_2$.

This is what we have observed in both the methods (see Table 4.2 (b) and Figures 4.2.(a) and (b)).

(ii) ***Intermolecular Nucleophilicity Trends for Systems Belonging to Different Series in Category B***

Here, the electronegativity parameter of *O*, *S*, *N* and *P* plays the deciding role. The electronegativity values change in the sequence $P < S < N < O$. Because of the electronegativity difference the value of $\Delta^{-}h(k)$ (i.e., $-[V_N^{el}(k) - V_{N-1}^{el}(k)]$) is also not same for two systems having same alkyl moiety *R*- but belonging to different homologous series. Since ***O*** atom has a stronger hold on its electron than ***N*** (say) $V_N^{el}(k)$ of ***O*** in *R*-***O****H* will be higher (numerically) than that of ***N*** in *R*-***N***H_2. However, relative compensation (due to $+I$ effect of *R*- moiety) after the loss of one electron will be more for ***N*** in *R*-***N***H_2 than ***O*** in *R*-***O****H*. This is presumably because ***N*** atom in neutral *R*-***N***H_2 does not pull that much electron as ***O*** does in *R*-***O****H*. So the left out electron density in the *R*- moiety is higher in *R*-***N***H_2 than in *R*-***O****H*. This left-out electron density will be pulled by the ***N*** atom in *R*-***N***H_2 when the electrophile attacks on it. As a result $V_{N-1}^{el}(k)$ for ***N*** in *R*-***N***H_2 will be higher (numerically) than that of ***O*** in *R*-***O****H*. Again considering the negative sign of $V^{el}(k)$ we get higher positive $\Delta^{-}h(k)$ value for ***O*** in *R*-***O****H* than ***N*** in *R*-***N***H_2. Thus lower the nucleophilicity higher is the $\Delta^{-}h(k)$ value of the nucleophilic atom.

The above argument holds true for systems of all four homologous series in **Category B** and our observed trend of $\Delta^{-}h(k)$ values (Figures 4.2.(a) & (b) and Table 4.2.(b)) in both the methods is as follows,

$R\text{-}\mathbf{O}H > R\text{-}\mathbf{N}H_2 > R\text{-}\mathbf{S}H > R\text{-}\mathbf{P}H_2$;

where $R = -CH_3, -CH_2CH_3, -CH_2\,CH_2CH_3$

There is only one exception and that is in B3LYP/6-31G** method the $\Delta^- h(k)$ value of $CH_3CH_2CH_2\mathbf{N}H_2$ is little higher than that of $CH_3CH_2CH_2\mathbf{O}H$, (see Table 4.2.(b) and Figure 4.2.(b)). This is similar to the trend of electronegativity values of the nucleophilic atoms and, as expected, opposite to the trend of their nucleophilicity.

4.6. Conclusion

The present chapter shows how hardness potential can effectively solve the $\frac{1}{N}$ problem in local hardness $\eta(\vec{r})$ parameter. While the hardness potential is proven to be an effective technique for analyzing and predicting the reactivity behavior for the systems belonging to the same homologous series, an intrinsic limitation upon its use is also recognized. Hardness potential $h(k)$, in its original working definition (Eq (4.5)), is unable to explain the reactivity sequence when the systems belong to different homologous series or when the reactive centres in the chemical systems vary from each other. To overcome this limitation we have proposed electrophilic and nucleophilic variants of the original hardness potential as $\Delta^+ h(k)$ and $\Delta^- h(k)$, respectively (Eqs (4.6) and (4.7)). The superiority of $\Delta^+ h(k)$ and $\Delta^- h(k)$ over $h(k)$ stems from the fact that they take care of the response (i.e., the changing electron density scenario) of the reactive centre toward an electrophilic (El^+) or nucleophilic (Nu^-) attack on it. It is interesting to observe that both of these two new descriptors correlate very well with the expected reactivity trends. It is also worth mentioning here that the electronegativity differences of the reactive atoms are also well taken care by these new descriptors. Keeps future applications in mind, $\Delta^+ h(k)$ and $\Delta^- h(k)$ have the potential to be used as efficient tools for studying the regio-chemistry of biological systems.

References and Notes

1. S. K. Ghosh and M. Berkowitz, *J. Chem. Phys.*, 1985, **83**, 2976.
2. W. Langenaeker, F. de Proft and P. Geerlings, *J. Phys. Chem.*, 1995, **99**, 6424.
3. P. Mignon, P. Geerlings and R. Schoonheydt, *J. Phys. Chem. C*, 2007, **111**, 12376.
4. P. Mignon, S. Loverix and P. Geerlings, *Chem. Phys. Lett.*, 2005, **401**, 40.
5. A. S. Ozen, F. De Proft, V. Aviyente and P. Geerlings, *J. Phys. Chem. A*, 2006, **110**, 5860.
6. A. Borgoo, M. Torrent-Sucarrat, F. D. Proft and P. Geerlings, *J. Chem. Phys.*, 2007, **126**, 234104.
7. S. Saha and R. K. Roy, *J. Phys. Chem. B*, 2007, **111**, 9664.
8. S. Saha and R. K. Roy, *J. Phys. Chem. B*, 2008, **112**, 1884.
9. M. Torrent-Sucarrat and P. Geerlings, *J. Chem. Phys.*, 2006, **125**, 244101.
10. S. K. Ghosh, *Chem. Phys. Lett.*, 1990, **172**, 77.
11. S. Liu, F. De Proft and R. G. Parr, *J. Phys. Chem.*, 1997, **101**, 6991.
12. M. Torrent-Sucarrat, M. Duran and M. Sola, *J. Phys. Chem. A*, 2002, **106**, 4632.
13. M. Torrent-Sucarrat, P. Salvador, P. Geerlings and M. Solà, *J. Comput. Chem.*, 2007, **28**, 574.
14. M. Torrent-Sucarrat, P. Salvador, M. Solà and P. Geerlings, *J. Comput. Chem.*, 2008, **29**, 1064.
15. M. Torrent-Sucarrat, F. D. Proft, P. W. Ayers and P. Geerlings, *Phys. Chem. Chem. Phys.* 2010, 1**2**, 1072.
16. M. Torrent-Sucarrat, F. De Proft, P. Geerlings and Paul W. Ayers, *Chem. - Eur. J.*, 2008, **14**, 8652.
17. S. Saha and R. Roy, *Phys. Chem. Chem. Phys.*, 2008, **10**, 5591.
18. P. Mignon, S. Loverix, F. De Proft and P. Geerlings, *J. Phys. Chem. A*, 2004, **108**, 6038.
19. R. G. Parr and J. L. Gazquez, *J. Phys. Chem.*, 1993, **97**, 3939.
20. P. Hohenberg and W. Kohn, *Phys. Rev. B*, 1964, **136**, 864.
21. L. H. Thomas, *Proc. Cambridge Philos. Soc.*, 1927, **23**, 542.
22. E. Fermi, *Z. Phys.*, 1928, **48**, 73.

23. P. A. M. Dirac, *Proc. Cambridge Philos. Soc.*, 1930, **26**, 376.

24. R. G. Parr and W. Yang, *Density–Functional Theory of Atoms and Molecules*, Oxford University Press, New York, 1989.

25. In the literature several authors have shown that the Coulomb term is usually dominant[a-d] in the evaluation of local hardness although the kinetic term is not negligible and can be compared to Coulomb term[e-g]: (a) Ref 28; (b) P. K. Chattaraj, A. Cedillo and R. G. Parr, *J. Chem. Phys.*, 1995, **103**, 10621; (b) Ref 11; (c) L. F. Pacios, *Chem. Phys. Lett.*, 1997, **276**, 381; (d) Ref 6; (e) Ref 9; (f) Ref 13; (g) Ref 14.

26. P. K. Chattaraj and S. Giri, *Annu. Rep. Prog. Chem., Sec C (Physical Chemistry)*, 2009, **105**, 13.

27. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, J. L. Andres, M. Head-Gordon, E. S. Replogle and J. A. Pople, *Gaussian 98, revision A.9*, Gaussian, Inc., Pittsburgh, PA, 1998.

28. R. G. Parr and W. Yang, *J. Am. Chem. Soc.*, 1984, **106**, 4049.

29. W. Yang and W. J. Mortier, *J. Am. Chem. Soc.*, 1986, **108**, 5708.

Table 4.1 (a): **Hardness potential (*h(k)*) values of $C_{C=O}$ (carbon atom of the *C=O* moiety; shown in bold) in the chosen systems of the Category A. In each boxes the upper value is generated by MP2(FC)/6-31G** method and the lower one by *B3LYP/6-31G*** method. Eq (4.5) (see text) is used to generate the values.**

Series	Chosen Systems for *Electrophilic Centre*	*h(k)* (a.u.)
(i)	-**C**OOH	
	H**C**OOH	21.9416 *22.0142*
	CH_3**C**OOH	24.2924 *24.3498*
	CH_3CH_2**C**OOH	25.8425 *25.8863*
(ii)	-**C**OF	
	H**C**OF	22.0632 *22.1504*
	CH_3**C**OF	24.4086 *24.4854*
	CH_3CH_2**C**OF	25.9538 *26.0098*
(iii)	-**C**$CONH_2$	
	H**C**$CONH_2$	21.7720 *21.8289*
	CH_3**C**$CONH_2$	24.0987 *24.1446*
	CH_3CH_2**C**$CONH_2$	25.6450 *25.6235*

Table 4.1 (b): **Hardness potential (*h(k)*) values of the nucleophilic atoms (printed in bold) in the chosen systems of the Category B. In each boxes the upper value is generated by MP2(FC)/6-31G** method and the lower one by *B3LYP/6-31G*** method. Eq (4.5) (see text) is used to generate the values.**

Series	Chosen Systems for *Nucleophilic Centre*	*h(k)* (a.u.)
(i)	**-O**H	
	CH_3**O**H	25.8630 *25.8712*
	CH_3CH_2**O**H	27.5066 *27.5015*
	CH_3CH_2 CH_2**O**H	28.5045 *28.4997*
(ii)	**-S**H	
	CH_3**S**H	62.0199 *61.9678*
	CH_3CH_2**S**H	63.4636 *63.3899*
	CH_3CH_2 CH_2**S**H	64.6378 *64.5469*
(iii)	**-N**H_2	
	CH_3**N**H_2	22.3281 *22.3292*
	CH_3CH_2**N**H_2	23.9245 *23.9136*
	CH_3CH_2 CH_2**N**H_2	24.9703 *24.9557*
(iv)	**-P**H_2	
	CH_3**P**H_2	57.2628 *57.2151*
	CH_3CH_2**P**H_2	58.6389 *58.5748*
	CH_3CH_2 CH_2**P**H_2	59.5843 *59.5138*

Table 4.2 (a): **Electrophilic hardness potential ($\Delta^{+}h(k)$, Eq (4.6)) values of the electrophilic centres (shown in bold) in the chosen chemical systems of the Category A. In each boxes the upper value is generated by MP2(FC)/6-31G** method and the lower one by *B3LYP/6-31G*** method.**

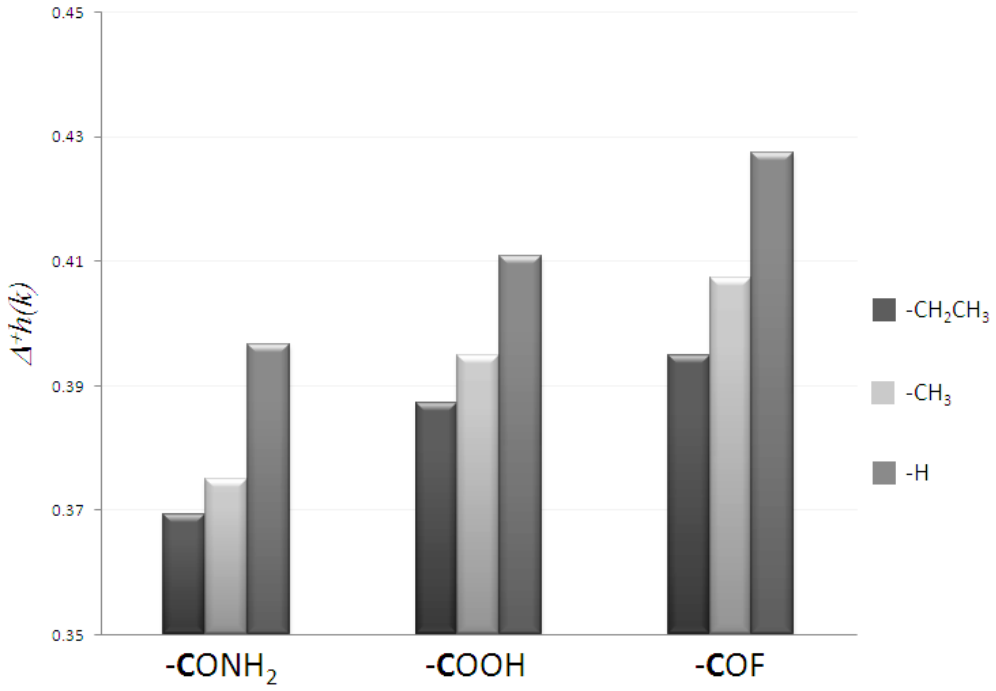
Series	Chosen Systems for *Electrophilic Centre*	$\Delta^{+}h(k)$ (a.u.)
(i)	-**C**OOH	
	H**C**OOH	0.4108 *0.3907*
	CH_3**C**OOH	0.3948 *0.3605*
	CH_3CH_2**C**OOH	0.3871 *0.3526*
(ii)	-**C**OF	
	H**C**OF	0.4275 *0.4095*
	CH_3**C**OF	0.4073 *0.3752*
	CH_3CH_2**C**OF	0.3949 *0.3480*
(iii)	-**C**$CONH_2$	
	H**C**$CONH_2$	0.3966 *0.3746*
	CH_3**C**$CONH_2$	0.3750 *0.3349*
	CH_3CH_2**C**$CONH_2$	0.3694 *0.3332*

Table 4.2 (b): Nucleophilic hardness potential (*Δh(k)*, Eq (4.7)) values of the nucleophilic centres (printed in bold) in the chosen chemical systems of Category B. In each boxes the upper value is generated by MP2(FC)/6-31G method and the lower one by *B3LYP/6-31G*** method.**

Series	Chosen Systems for *Nucleophilic Centre*	*Δh(k)* (a.u.)
(i)	**-O**H	
	CH_3**O**H	0.4830 *0.4472*
	CH_3CH_2**O**H	0.4681 *0.4004*
	CH_3CH_2 CH_2**O**H	0.4571 *0.3753*
(ii)	**-S**H	
	CH_3**S**H	0.3610 *0.3587*
	CH_3CH_2**S**H	0.3518 *0.3453*
	CH_3CH_2 CH_2**S**H	0.3467 *0.3347*
(iii)	**-N**H_2	
	CH_3**N**H_2	0.4273 *0.4133*
	CH_3CH_2**N**H_2	0.4156 *0.3897*
	CH_3CH_2 CH_2**N**H_2	0.4112 *0.3789*
(iv)	**-P**H_2	
	CH_3**P**H_2	0.3352 *0.3338*
	CH_3CH_2**P**H_2	0.3245 *0.3148*
	CH_3CH_2 CH_2**P**H_2	0.3194 *0.3007*

Figure 4.1: Bar diagram presentation of the electrophilic hardness potential ($\Delta^{+}h(k)$) values of the electrophilic atom (shown in bold) in the chosen chemical systems of Category A **(a)** at MP2(FC)/6-31G** **(b)** B3LYP/6-31G**.

(a)



(b)

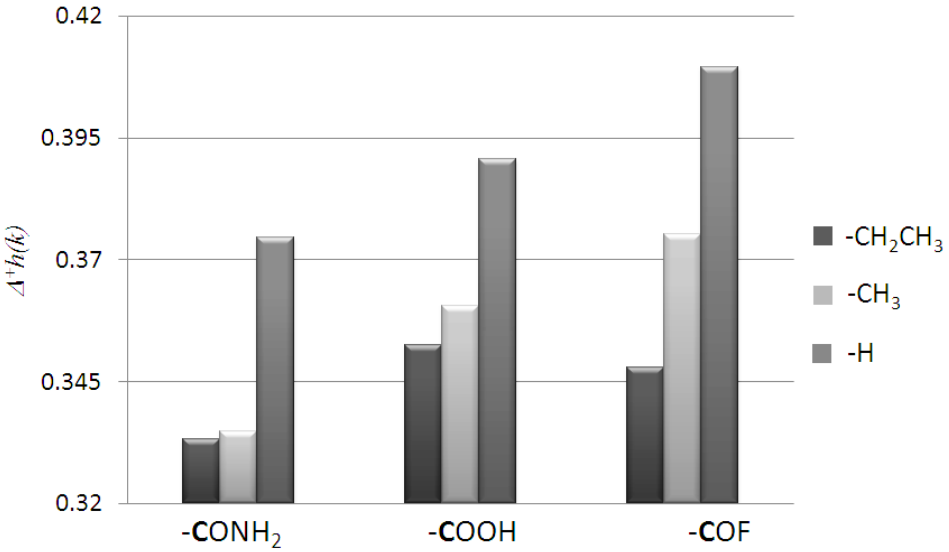
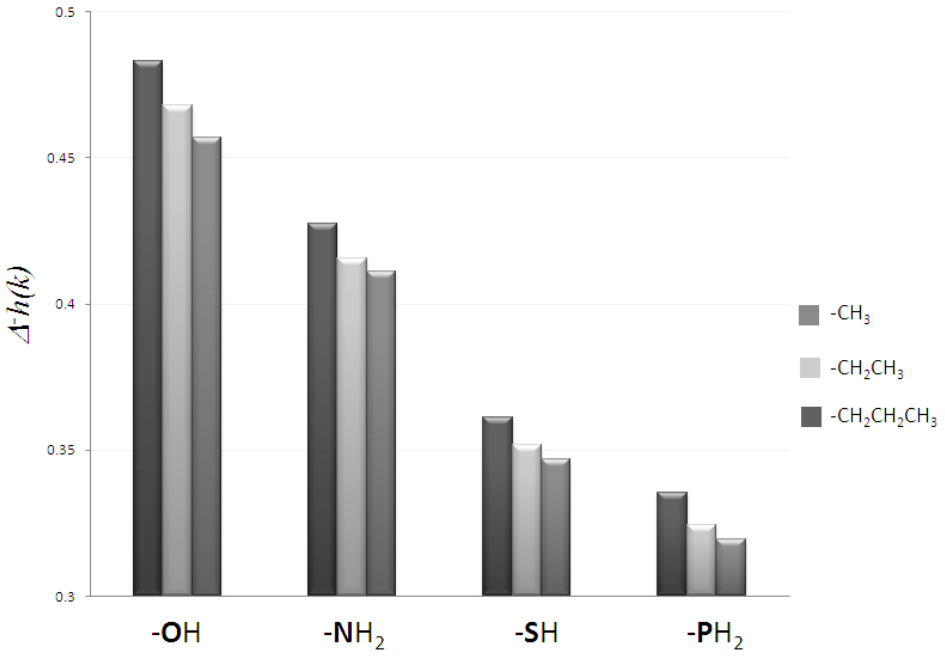
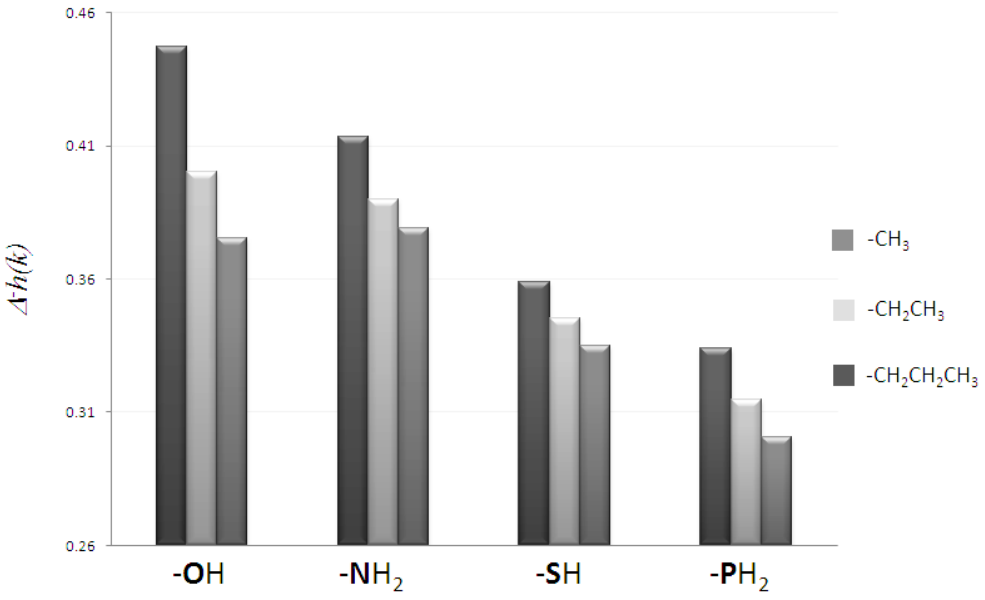


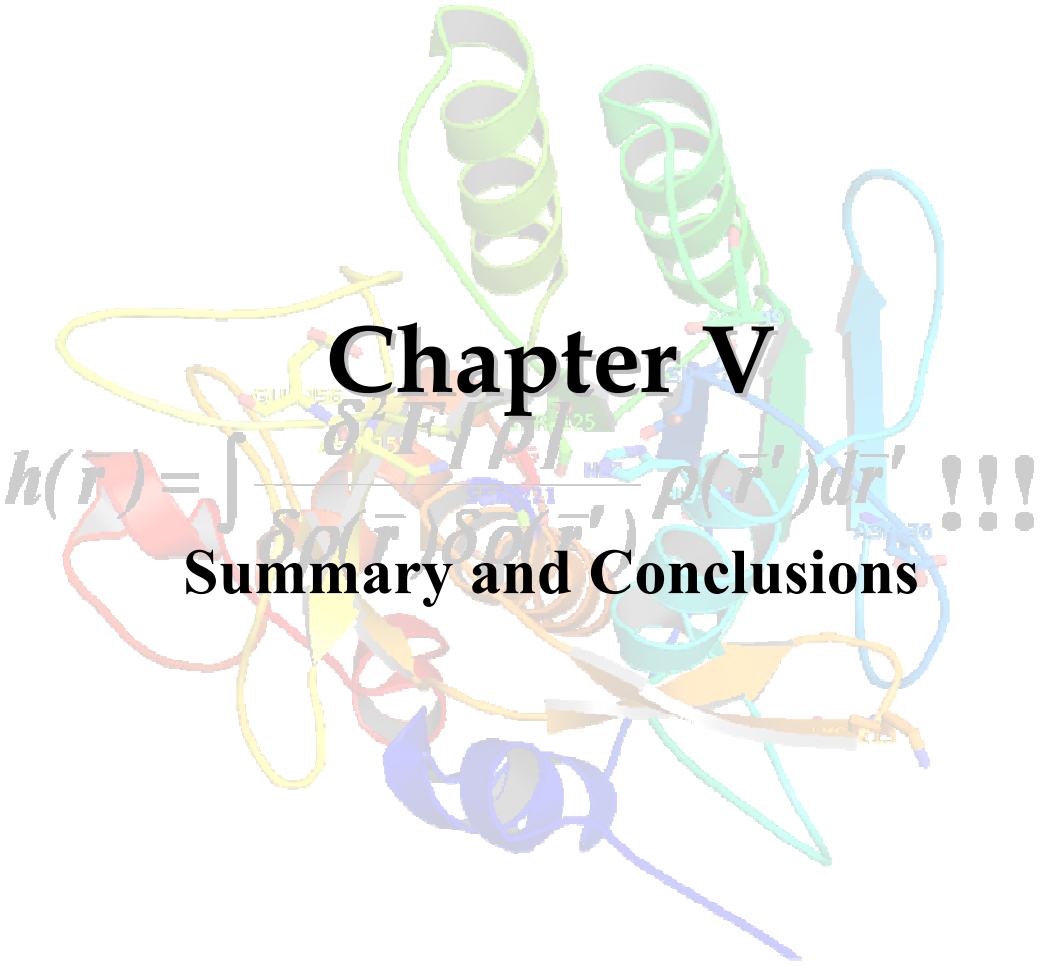
Figure 4.2: Bar diagram presentation of the nucleophilic hardness potential ($\Delta^{-}h(k)$) values of the nucleophilic atom (printed in bold) in the chosen chemical systems of Category B **(a)** at MP2(FC)/6-31G** **(b)** B3LYP/6-31G**.

(a)



(b)





Chapter V

Summary and Conclusions

5.1. General Conclusions

The work presented in this thesis is an attempt to develop a cost-effective and reliable approach for predicting the regioselectivity of large chemical and biological systems using Conceptual Density Functional Theory (DFT)[1-11] based reactivity descriptors. The last few decades have wintnessed spectacular developments in the computational methodologies. However, despite these developments, an accurate quantum chemical modeling of large molecules and extended systems, such as biological systems, had always been a challenge for computational chemists. In this thesis, we have presented a novel approach using DFT based reactivity descriptors, namely local hardness ($\eta(\bar{r})$) and molecular modeling techniques to make a systematic and computationally less expensive way of predicting regioselectivity for large systems. We have proposed a model, named as 'One-into-Many' model,[12,13] in which one can break the larger system into different smaller ones, each having at least one reactive site and then to study the reactivity of the required active sites in the individual fragments using hardness-based reactivity descriptors (e.g., local hardness,[14-16] $\eta(\bar{r})$). The argument in favour of using local hardness $\eta(\bar{r})$ originates from the fact that the regioselectivity (or site selectivity) is a local phenomenon. It is assumed that the contribution to the local reactivity descriptor (e.g., 'local hardness', evaluated on the basis of Thomas-Fermi-Dirac (TFD)[17-19] approach of density functional theory) from the distance atoms or moieties are less significant and thus can be neglected. Moreover, local hardness is mainly a charge-controlled descriptor as it contains electronic contribution to the molecular electrostatic potential. Thus it can take care of the long-range reactivity (i.e., intermolecular reactivity) and hence, one can recast the intramolecular problem of a large system into an intermolecular problem of its smaller fragments and then predict the regioselectivity of the original large system. However, contribution from closer atoms or environment can be taken care of by careful fragmentation process. This is to be achieved, even if not fully, by keeping some 'buffer zone' on both sides of the reactive site. Here the 'buffer zone' refers to that moiety of the chemical system, which is common to two adjacent reactive sites (local hardness values of which are to be evaluated). It is notable that $\eta(\bar{r})$ (or better it's condensed form, $\eta(k)$) suffers from one

severe limitation and that is it's N-dependence problem.[20] The broader applicability of $\eta(\bar{r})$ as a reliable intermolecular reactivity descriptor necessitates the removal of its N-dependence. Therefore, we have provided a simple formalism to obtain the hardness potential,[21] $h(\bar{r})$, and introduced electrophilic and nucleophilic variants of $h(\bar{r})$[22] for characterising the bio-molecular systems.

5.2. Specific Conclusions

The study of regioselectivity remains a very important topic in the chemical literature. Thousand of papers have appeared on this subject, and interest in it is accelerating. Furthermore, not only are the methodologies for experimentation steadily improving, but also the content of the theory is still evolving. In DFT, the big advantage is that the electron number, *N*, has a central place in the theory. A great strength of the density functional language -augured Conceptual DFT- is its appropriateness for defining and elucidating important universal concepts of molecular structure and molecular reactivity. By now there is powerful evidence that the Conceptual DFT here used, and here extended, provides not only a correct quantitative description of molecular electronic structure but also is a description generally in agreement with previous studies in the chemical literature.

Accurate prediction of regioselectivity of large chemical and biological systems has its bottleneck in the computational limitation (because the computation is to be performed on the molecule as a whole). To allow for the treatment of larger molecular systems it is often necessary to adopt fragmentation based approach. This improvement will allow for a theoretical investigation of large molecules such as biomolecules, an area which has been so far limited to the application of semiempirical or force-field approaches. However, since the reliability of these methods is unclear and they do not offer procedures to systematically improve the level of accuracy, it is highly desirable to apply reliable methods to predict the regioselectivity of large chemical and biological systems. Thus, detailed physical descriptions of large and complicated biological molecules for the purpose of understanding and modulating their biological functions require more intensive efforts than ever.

Motivated by its potential importance and guided by the insights from Conceptual DFT, we have given one simple fragmentation ('One-into-Many' model[12,13]) approach. This model may be considered as the first one, which describes how Conceptual DFT based reactivity descriptor can be used to systematically address the regioselectivity problem of large chemical and biological systems. Large chemical and biological systems with multiple reactive sites are proposed to be broken into small fragments having at least one reactive site in each fragment. The environment around each reactive site is mimicked by incorporating a buffer zone. To implement the 'One-into-Many'[12,13] model, we have chosen right-handed B-DNA (PDB ID: 1BNA)[23] as a model system. In this model system there are 12 base-pairs with the sequence d(CpGpCpGpApApTpTpCpGpCpG).[23] Detail literature study on adduct formation, indicates that the majority of known carcinogens react with DNA through N^2 and N7 positions of guanine.[24-28] Those positions are the most reactive sites towards electrophilic attack in double-stranded DNA. Apart from these positions, it is also reported that the exocyclic oxygen of guanine (O^6)[29,30] and the exocyclic oxygen of thymine (O^2)[30,31] residues are the reactive sites for electrophilic attack. The reactivity descriptor, used in 'One-into-Many' model as a key tool, is local hardness[14-16] ($\eta(\bar{r})$) because its dominant component is the electronic part of the molecular electrostatic potential (MESP). MESP has a long distant effect, thus making it suitable for predicting intermolecular reactivity and so fitting the proposed model. We have calculated the 'atomic condensed' local hardness (i.e., $\widetilde{\eta}_D^{TFD}(k)$ and $\widetilde{\eta}_D^{TFD'}(k)$) values of those reactive sites by considering one of the G·C (i.e., Guanine-Cytosine) base-pair and one of the A·T (i.e., Adenine -Thymine) base-pair (given the name '*Single-Base-Pair Systems*') of DNA molecule. However, to generate more reliable data a buffer zone around the reactive sites is required. So, three base-pairs were chosen at a time (named as '*Triple-Base-Pair Systems*') and the $\widetilde{\eta}_D^{TFD}(k)$ and also $\widetilde{\eta}_D^{TFD'}(k)$ values of the reactive sites in the middle base-pair were evaluated. Here the base-pairs, which are on either side of the central base-pair, form the buffer zone to mimicking the environment of the 1BNA.[23]All possible combinations of '*Triple-Base-Pair Systems*' were also taken from 1BNA.[23] We have concluded[12, 13] that the trends of atomic hardness values generated by the proposed model are as expected for exocyclic

NH_2-groups and for ring N-atoms of the DNA base-pair systems. Only for exocyclic O-atom in DNA base-pairs, the method proposed by us fails to generate expected trends of hardness values ($\tilde{\eta}_D^{TFD}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$). We have reported[12, 13] that (both from $\tilde{\eta}_D^{TFD}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$ values) O^2 position of cytosine is the most reactive site even though Singer[30] has observed that O^6 of guanine and O^2 of thymine are the two significant reactive sites towards electrophilic attack in the double helical DNA. This failure is possibly attributable to our inability (due to lack of computational facilities) to (i) take care of the dielectric effect of the biological medium and (ii) include polarization and diffuse functions in the basis set for *Triple-Base-Pair Systems*. However, the method developed here appears to be promising once the computational limitations are taken care of.

Also, we have attempted[20] to explain the problems of evaluating local hardness (i.e., $\eta(\bar{r})$)[14-16] choosing relative reactivity aspects of some carbonyl compounds. To inspect the relative reactivity of carbonyl compounds as model systems. we have found that none of the two composite functions (i.e., $\lambda[\rho(\bar{r}')] = \rho(\bar{r}')$ [15,32-36,12,13] or $\lambda[\rho(\bar{r}')] = Nf(\bar{r}')$ [37-44]) is able to overcome the N-dependence problem of local hardness $\eta(\bar{r})$ parameter.[20] After a careful analysis we have revealed that as the number of electrons increases with the size of the system, the $\frac{1}{N}$ factor alters the expected trends of $\tilde{\eta}_D^{TFD}(\bar{r})$ or $\tilde{\eta}_D^{TFD'}(\bar{r})$ (i.e., when $\lambda[\rho(\bar{r}')]$ is $\rho(\bar{r}')$) values. And in the case of $\tilde{\eta}_F^{TFD}(\bar{r})$ (i.e., by using $\lambda[\rho(\bar{r}')] = Nf(\bar{r}')$ [15,32-36,12,13]), we have realised that it is the Fukui function [$f(\bar{r})$] term which is altering the expected trend. Although, using the composite function $\lambda[\rho(\bar{r}')] = Nf(\bar{r}')$,[37-44] apparently we could remove the $\frac{1}{N}$ problem from the expression of local hardness $\tilde{\eta}_F^{TFD}(\bar{r})$, N-dependence is re-introduced indirectly through the normalization condition of Fukui function (i.e., $\int f(\bar{r})d\bar{r} = 1$ or $\sum_k^N f(k) = 1$).[4] So, the broader applicability of $\eta(\bar{r})$ as a reliable intermolecular reactivity descriptor

necessitates the removal of its $\frac{1}{N}$ dependence. This is because the comparison of intermolecular reactivity trends is based mainly on the local hardness values of those particular sites (or atoms in the condensed form), with the electronic or any other effects exerted by the rest of the system already incorporated. But the $\frac{1}{N}$ factor creates an impression as if the whole system does equally contribute to the reactivity of that particular site or atom. In reality, parts (or moieties) of the system, far from the site of interest, may have little or no effect on the reactivity of that particular site. Geerlings and collaborators also raised a similar argument in some of their earlier studies.[45] Therefore, the best way to incorporate the electronic (or any other) effects of the rest of the system, without overemphasizing $\frac{1}{N}$ factor, is to consider only the active site (or atoms or group) for which the number of electrons is same. For example, we have shown[20] that if only $C=O$ moiety is considered (N will be same, i.e., $N=14$, for all the chosen carbonyl systems) one can use the modified 'condensed-to-atom' form of Eqs (3.20) and (3.21) as,

$$\tilde{\eta}_D^{TFD}(k)=\frac{10}{9}C_F P(k)^{2/3}-V^{el}(k)-\frac{4}{9}C_X P(k)^{1/3} \tag{3.20}$$

$$\tilde{\eta}_D^{TFD'}(k)=-V^{el}(k) \tag{3.21}$$

The N-dependence problem of local hardness can be eliminated by practical implementation[22] of variants of the hardness potential[21] $h(\bar{r})$. This is another important contribution of this thesis. Starting from Parr and Gázquez's hardness potential[21] $h(\bar{r})$, we have proposed a working equation for evaluating $h(\bar{r})$. As $h(\bar{r})$ does not contain $\frac{1}{N}$ factor, using it the N-dependence problem of $\eta(\bar{r})$ can be done away with. We have shown that the trends of electrophilicity and nucleophilicity generated by $h(k)$ values of the strongest electrophilic and nucleophilic atoms, respectively, are exactly as expected when the systems belong to the same homologous series. However, irregularities are observed when systems belonging to different series are compared. To aid in this pursuit, we have explained[22] why $h(\bar{r})$ (or $h(k)$) sometimes gives results contrary to chemical intuition. It seems that $h(\bar{r})$ does not take care of the response of the system in the event

of an approaching electrophile or nucleophile. This lead us to define[22] the corresponding electrophilic ($\Delta^{+}h(k)$) and nucleophilic ($\Delta^{-}h(k)$) variants of the hardness potential, which measure the reactivity toward a nucleophilic (i.e., Nu^{-}) and an electrophilic (i.e., El^{+}) reagent, respectively. It is shown that our proposed reactivity descriptors correctly predict the expected trend in each case.[22] With an eye on future applications, $\Delta^{+}h(k)$ and $\Delta^{-}h(k)$ have the potential to be used as good tools for studying the regio-chemistry of biological systems.

5.3. Future Scope of Work

The novel approaches adopted in this thesis are in the direction of the application of DFT based reactivity descriptors to predict the regioselectivity of large chemical and biological systems. This thesis provides a scope both for calculating and discussing the molecular behavior of large molecular systems.

We have shown that by combining Conceptual DFT[1-11] with fragmentation approach, one can obtain detailed physical descriptions particularly the regiochemistry of large and complicated biological molecules. However, further studies are required to judge the effectiveness of the proposed method by applying it to other large systems. Finding out the limit of 'buffer zone' in the proposed 'One-into-Many'[12,13] model might be a worthwhile attempt in future.

A combined search by using 'One-into-Many'[12,13] model and electrophilic ($\Delta^{+}h(k)$) and nucleophilic ($\Delta^{-}h(k)$) hardness potential[21,22] seems to be an interesting way to explore the biological functions. In our present study, we have used Thomas-Fermi-Dirac (TFD)[17-19] approach for evaluating hardness potential. One can extend the study of hardness potential to other approaches (by adding $\frac{1}{9}$ of the Weizsäcker functional[46] and a Wigner-type[47] local correlation functional) which takes care of the loopholes in TFD. Moreover, the preferable site of attack depends upon the mode of attack of that approaching species, pH and nature of the solvent. Further work on involving these criteria in the evaluation of $\Delta^{+}h(k)$ and $\Delta^{-}h(k)$ would be of considerable interest.

Like most conceptual DFT based reactivity descriptors, hardness potential and its electrophilic ($\Delta^{+}h(k)$) and nucleophilic ($\Delta^{-}h(k)$) variants possess strong interpretive power, which is important as it leads to a deep understanding of diverse class of chemical and biochemical processes. However, their predictive capacity has yet to be assessed fully.

In our opinion this study paves the way for further studies which may involve more complex molecules. With increased computational power - both software and hardware - and DFT based methodology, one may finally achieve better results. This will be of great importance in areas such as drug design and related applications.

References

1. R. G. Parr, R. A. Donnelly, M. Levy and W. E. Palke, *J. Chem. Phys.*, 1978, **68**, 3801.
2. R. G. Parr, *Annu. Rev. Phys. Chem.*, 1983, **34**, 631.
3. R. F. Nalewajski, *J. Phys. Chem.*, 1985, **89**, 2831.
4. R. G. Parr and W. Yang, *Density–Functional Theory of Atoms and Molecules*, Oxford University Press, New York, 1989.
5. R. G. Parr and W. Yang, *Annu. Rev. Phys. Chem.*, 1995, **46**, 701.
6. H. Chermette, *J. Comput. Chem.*, 1999, **20**, 129.
7. (a) P. Geerlings and F. De Proft, *Int. J. Mol. Sci.*, 2002, **3**, 276; (b) P. Geerlings, F. De Proft and W. Langenaeker, *Chem. Rev.*, 2003, **103**, 1793.
8. R. K. Roy and S. Saha; *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.*, 2010, DOI: 10.1039/b811052m (In Press).
9. M. H. Cohen and A. Wasserman, *J. Phys. Chem. A*, 2007, **111**, 2229.
10. J. L. Gázquez, *J. Mex. Chem. Soc.*, 2008, **52**, 3.
11. in *Chemical reactivity theory: a density functional view*, ed. P. K. Chattaraj, CRC, 2009.
12. S. Saha and R. K. Roy, *J. Phys. Chem. B*, 2007, **111**, 9664.
13. S. Saha and R. K. Roy, *J. Phys. Chem. B*, 2008, **112**, 1884.
14. M. Berkowitz, S. K. Ghosh and R. G. Parr, *J. Am. Chem. Soc.*, 1985, **107**, 6811.
15. S. K. Ghosh and M. Berkowitz, *J. Chem. Phys.*, 1985, **83**, 2976.
16. P. K. Chattaraj, D. R. Roy, P. Geerlings and M. Torrent-Sucarrat, *Theor. Chem. Acc.*, 2007, **118**, 923.
17. L. H. Thomas, *Proc. Cambridge Philos. Soc.*, 1927, **23**, 542.
18. E. Fermi, *Z. Phys.*, 1928, **48**, 73.
19. P. A. M. Dirac, *Proc. Cambridge Philos. Soc.*, 1930, **26**, 376.
20. S. Saha and R. Roy, *Phys. Chem. Chem. Phys.*, 2008, **10**, 5591.
21. R. G. Parr and J. L. Gazquez, *J. Phys. Chem.*, 1993, **97**, 3939.
22. S. Saha and R. K. Roy (Communicated).
23. H. R. Drew, R. M. Wing, T. Takano, C. Broka, S. Tanaka, K. Itakura and R. E. Dickerson, *Proc. Natl. Acad. Sci. U.S.A.*, 1981, **78**, 2179.

24. P. D. Lawley and P. Brookes, *Nature (London)*, 1961, **192**, 1081.
25. C. Nagata, A. Imamura, H. Salto and K. Fukui, *Gann*, 1963, **54**, 109.
26. A. Dipple, P. Brookes, D. S. Mackintosh and M. P. Rayman, *Biochemistry*, 1971, **10**, 4323.
27. J. J. Danneberg and M. Tomasz, *J. Am. Chem. Soc.*, 2000, **122**, 2062.
28. K. S. Gates, T. Nooner and S. Dutta, *Chem. Res. Toxicol.*, 2004, **17**, 839.
29. A. Loveless, *Nature (London)*, 1969, **223**, 206.
30. B. Singer, *Nature (London)*, 1976, **264**, 333.
31. B. Singer and J. M. Essigmann, *Carcinogenesis*, 1991, **12**, 949.
32. W. Langenaeker, F. de Proft and P. Geerlings, *J. Phys. Chem.*, 1995, **99**, 6424.
33. P. Mignon, S. Loverix and P. Geerlings, *Chem. Phys. Lett.*, 2005, **401**, 40.
34. A. S. Ozen, F. De Proft, V. Aviyente and P. Geerlings, *J. Phys. Chem. A*, 2006, **110**, 5860.
35. A. Borgoo, M. Torrent-Sucarrat, F. D. Proft and P. Geerlings, *J. Chem. Phys.*, 2007, **126**, 234104.
36. P. Mignon, P. Geerlings and R. Schoonheydt, *J. Phys. Chem. C*, 2007, **111**, 12376.
37. S. K. Ghosh, *Chem. Phys. Lett.*, 1990, **172**, 77.
38. S. Liu, F. De Proft and R. G. Parr, *J. Phys. Chem.*, 1997, **101**, 6991.
39. M. Torrent-Sucarrat, M. Duran and M. Sola, *J. Phys. Chem. A*, 2002, **106**, 4632.
40. M. Torrent-Sucarrat and P. Geerlings, *J. Chem. Phys.*, 2006, **125**, 244101.
41. M. Torrent-Sucarrat, P. Salvador, P. Geerlings and M. Solà, *J. Comput. Chem.*, 2007, **28**, 574.
42. M. Torrent-Sucarrat, F. De Proft, P. Geerlings and Paul W. Ayers, *Chem. - Eur. J.*, 2008, **14**, 8652.
43. M. Torrent-Sucarrat, P. Salvador, M. Solà and P. Geerlings, *J. Comput. Chem.*, 2008, **29**, 1064.
44. M. Torrent-Sucarrat, F. D. Proft, P. W. Ayers and P. Geerlings, *Phys. Chem. Chem. Phys.* 2010, 1**2**, 1072.
45. P. Mignon, S. Loverix, F. De Proft and P. Geerlings, *J. Phys. Chem. A*, 2004, **108**, 6038.

46. C. Von Weizsäcker, *Z. Phys*, 1935, **96**, 431.

47. E. Wigner, *Phys. Rev.*, 1934, **46**, 1002.

$$E_E^{TFD}[\rho(\bar{r})] = C_F \int \rho(\bar{r})^{5/3} d\bar{r} + \frac{1}{2}\iint \frac{\rho(\bar{r})\rho(\bar{r}')}{|\bar{r}-\bar{r}'|} d\bar{r}' d\bar{r} - C_X \int \rho(\bar{r})^{4/3} d\bar{r}$$

$$\eta_\lambda(\bar{r}) = \left(\frac{\delta\mu}{\delta\rho(\bar{r})}\right)_{v(\bar{r})} = \frac{1}{N}\int \eta(\bar{r},\bar{r}')\lambda|\rho(\bar{r}')|d\bar{r}' = \frac{1}{N}\int \frac{\delta^2 E_E}{\delta\rho(\bar{r}')\delta\rho(\bar{r})}\lambda|\rho(\bar{r}')|d\bar{r}'$$

Appendices

$$\tilde{\eta}_{fv}(\bar{r}) = \int \eta(\bar{r},\bar{r}') f(\bar{r}') d\bar{r}'$$

$$h(r) = \int \frac{\delta^2 \Gamma[\rho]}{\delta\rho(\bar{r})\delta\rho(\bar{r}')} \rho(\bar{r}') d\bar{r}' = N\eta(\bar{r})$$

Appendix A

A detail derivation of $\tilde{\eta}_D^{TFD}(\bar{r})$

The Thomas-Fermi-Dirac energy functional ($F_E^{TFD}[\rho(\bar{r})]$) comprises three elements, the kinetic energy functional $T_K[\rho]$, the classical Coulomb repulsion functional $J[\rho]$ and the exchange energy functional $E_X[\rho]$. That is,

$$F_E^{TFD}[\rho(\bar{r})] = T_K[\rho] + J[\rho] + E_X[\rho] \tag{A1}$$

where $T_K[\rho]$ is the Thomas-Fermi kinetic energy functional,

$$T_K[\rho] = C_F \int \rho(\bar{r})^{5/3} d\bar{r} \tag{A2}$$

$J[\rho]$ is the classical Coulomb repulsion energy,

$$J[\rho] = \frac{1}{2} \iint \frac{\rho(\bar{r})\rho(\bar{r}')}{|\bar{r} - \bar{r}'|} d\bar{r} d\bar{r}' \tag{A3}$$

and $E_X[\rho]$ is the Dirac exchange energy functional

$$E_X[\rho] = -C_X \int \rho(\bar{r})^{4/3} d\bar{r} \tag{A4}$$

Insertion of (A2)-(A4) into (A1) gives, finally,

$$F_E^{TFD}[\rho(\bar{r})] = C_F \int \rho(\bar{r})^{5/3} d\bar{r} + \frac{1}{2} \iint \frac{\rho(\bar{r})\rho(\bar{r}')}{|\bar{r} - \bar{r}'|} d\bar{r}' d\bar{r} - C_X \int \rho(\bar{r})^{4/3} d\bar{r} \tag{2.1}$$

To evaluate the contributions of these terms in hardness kernel $\eta(\bar{r},\bar{r}')$ one may proceed in the following way.

Consider first the kinetic energy functional

$$T_K[\rho] = C_F \int \rho(\bar{r})^{5/3} d\bar{r} \tag{A2}$$

then the contribution of kinetic energy functional to $\eta_K(\bar{r},\bar{r}')$ becomes

$$\eta_K(\bar{r},\bar{r}') = \frac{\delta^2 T_K}{\delta\rho(\bar{r}')\delta\rho(\bar{r})} = \frac{10}{9} C_F \rho(\bar{r})^{-1/3} \delta(\bar{r} - \bar{r}') \tag{A5}$$

because,

$$\frac{\delta T_K}{\delta\rho(\bar{r})} = \frac{5}{3} C_F \rho(\bar{r})^{2/3} \tag{A6}$$

then, $$\frac{\delta^2 T_K}{\delta\rho(\bar{r}')\delta\rho(\bar{r})} = \frac{10}{9} C_F \rho(\bar{r})^{-1/3} \frac{\delta\rho(\bar{r})}{\delta\rho(\bar{r}')} = \frac{10}{9} C_F \rho(\bar{r})^{-1/3} \delta(\bar{r} - \bar{r}') \quad \text{(A7)}$$

As total local hardness (when evaluating the local hardness $\eta(\bar{r})$ using $\lambda[\rho(\bar{r}')] = \rho(\bar{r}')$) is

$$\eta_D^{TFD}(\bar{r}) = \frac{1}{N} \int \eta(\bar{r}, \bar{r}') \rho(\bar{r}') d\bar{r}' \quad (1.67)$$

then, the contribution of kinetic energy functional to $\eta_K(\bar{r})$ becomes

$$\eta_K(\bar{r}) = \frac{10}{9N} C_F \int \rho(\bar{r})^{-1/3} \delta(\bar{r} - \bar{r}') \rho(\bar{r}') d\bar{r}' \quad \text{(A8)}$$

Finally, $$\eta_K(\bar{r}) = \frac{10}{9N} C_F \rho_N(\bar{r})^{2/3} \quad \text{(A9)}$$

as $$\int \delta(\bar{r} - \bar{r}') \rho(\bar{r}') d\bar{r}' = \rho(\bar{r}) \quad \text{(A10)}$$

Similarly, for exchange term one can write

$$E_X[\rho] = -C_X \int \rho(\bar{r})^{4/3} d\bar{r} \quad \text{(A4)}$$

the $\eta_X(\bar{r})$ must be

$$\eta_X(\bar{r}) = \frac{1}{N} \int \eta_X(\bar{r}, \bar{r}') \rho(\bar{r}') d\bar{r}' \quad \text{(A11)}$$

or $$\eta_X(\bar{r}) = -\frac{4}{9N} C_X \int \rho(\bar{r})^{-2/3} \delta(\bar{r} - \bar{r}') \rho(\bar{r}') d\bar{r}' \quad \text{(A12)}$$

in order to $$\eta_X(\bar{r}, \bar{r}') = \frac{\delta^2 E_X}{\delta\rho(\bar{r}')\delta\rho(\bar{r})} = -\frac{4}{9} C_X \rho(\bar{r})^{-2/3} \delta(\bar{r} - \bar{r}') \quad \text{(A13)}$$

(as $$\frac{\delta^2 E_X}{\delta\rho(\bar{r}')\delta\rho(\bar{r})} = \frac{\delta}{\delta\rho(\bar{r}')}\left[\frac{\delta E_X}{\delta\rho(\bar{r})}\right] = \frac{\delta}{\delta\rho(\bar{r}')}\left[-\frac{4}{3} C_X \rho(\bar{r})^{1/3}\right]$$

$$= -\frac{4}{9} C_X \rho(\bar{r})^{-2/3} \frac{\delta\rho(\bar{r})}{\delta\rho(\bar{r}')} = -\frac{4}{9} C_X \rho(\bar{r})^{-2/3} \delta(\bar{r} - \bar{r}'))$$

So, applying Eq (A10) in Eq (A12), one finds that

$$\eta_X(\bar{r}) = -\frac{4}{9N} C_X \rho(\bar{r})^{1/3} \quad \text{(A14)}$$

Finally consider the Coulomb term,

$$J[\rho] = \frac{1}{2}\iint \frac{\rho(\bar{r})\rho(\bar{r}')}{|\bar{r}-\bar{r}'|} d\bar{r}d\bar{r}' \tag{A3}$$

For this we find,

$$\eta_J(\bar{r},\bar{r}') = \frac{\delta^2 J}{\delta\rho(\bar{r})\delta\rho(\bar{r}')} = \frac{1}{|\bar{r}-\bar{r}'|} \tag{A15}$$

because,

$$\begin{aligned}
\frac{\delta J[\rho]}{\delta\rho(\bar{r}_1)} &= \frac{1}{2}\left[\iint \frac{\delta\rho(\bar{r})}{\delta\rho(\bar{r}_1)}\frac{\rho(\bar{r}')}{|\bar{r}-\bar{r}'|} d\bar{r}d\bar{r}' + \iint \frac{\rho(\bar{r})}{|\bar{r}-\bar{r}'|}\frac{\delta\rho(\bar{r}')}{\delta\rho(\bar{r}_1)} d\bar{r}d\bar{r}'\right] \\
&= \frac{1}{2}\left[\iint \delta(\bar{r}-\bar{r}_1)\frac{\rho(\bar{r}')}{|\bar{r}-\bar{r}'|} d\bar{r}d\bar{r}' + \iint \frac{\rho(\bar{r})}{|\bar{r}-\bar{r}'|}\delta(\bar{r}'-\bar{r}_1) d\bar{r}d\bar{r}'\right] \\
&= \frac{1}{2}\left[\int \frac{\rho(\bar{r}')}{|\bar{r}_1-\bar{r}'|} d\bar{r}' + \int \frac{\rho(\bar{r})}{|\bar{r}-\bar{r}_1|} d\bar{r}\right]
\end{aligned} \tag{A16}$$

$$\begin{aligned}
\frac{\delta J[\rho]}{\delta\rho(\bar{r}_2)\delta\rho(\bar{r}_1)} &= \frac{1}{2}\left[\int \frac{\delta\rho(\bar{r}')}{\delta\rho(\bar{r}_2)}\frac{1}{|\bar{r}_1-\bar{r}'|} d\bar{r}' + \int \frac{\delta\rho(\bar{r})}{\delta\rho(\bar{r}_2)}\frac{1}{|\bar{r}-\bar{r}_1|} d\bar{r}\right] \\
&= \frac{1}{2}\left[\int \delta(\bar{r}'-\bar{r}_2)\frac{1}{|\bar{r}_1-\bar{r}'|} d\bar{r}' + \int \delta(\bar{r}-\bar{r}_2)\frac{1}{|\bar{r}-\bar{r}_1|} d\bar{r}\right] \\
&= \frac{1}{2}\left[\frac{1}{|\bar{r}_1-\bar{r}_2|} + \frac{1}{|\bar{r}_2-\bar{r}_1|}\right] = \frac{1}{|\bar{r}_1-\bar{r}_2|}
\end{aligned}$$

or,
$$\frac{\delta J[\rho]}{\delta\rho(\bar{r}')\delta\rho(\bar{r})} = \frac{1}{|\bar{r}-\bar{r}'|} \tag{A17}$$

So using Eq (A15) one can obtain

$$\eta_J(\bar{r}) = \frac{1}{N}\int \eta_J(\bar{r},\bar{r}')\rho(\bar{r}')d\bar{r}' = \frac{1}{N}\int \frac{1}{|\bar{r}-\bar{r}'|}\rho(\bar{r}')d\bar{r}' \tag{A18}$$

Now, as
$$\int \frac{\rho(\bar{r})}{|\bar{r}-\bar{r}'|} d\bar{r} = -V^{el}(\bar{r}) \tag{A19}$$

so,
$$\eta_J(\bar{r}) = -\frac{1}{N}V^{el}(r) \tag{A20}$$

Putting (A9), (A20) and (A14) together, we arrive at Eq (2.2).

Appendix B

A detail derivation of $\widetilde{\eta}_F^{TFD\ -}(\bar{r})$ and $\widetilde{\eta}_F^{TFD\ +}(\bar{r})$

The frontier local hardness (i.e., when evaluating the local hardness $\eta(\bar{r})$ using $\lambda[\rho(\bar{r}')] = Nf(\bar{r}')$) is

$$\eta_F^{TFD}(\bar{r}) = \int \eta(\bar{r},\bar{r}')f(\bar{r}')d\bar{r}' \tag{3.3}$$

So, the contribution of kinetic energy functional to $\eta(\bar{r})$ becomes

$$\eta_K(\bar{r}) = \int \eta_K(\bar{r},\bar{r}')f(\bar{r}')d\bar{r}' \tag{B1}$$

The contribution of kinetic energy functional to $\eta_K(\bar{r},\bar{r}')$ becomes

$$\eta_K(\bar{r},\bar{r}') = \frac{\delta^2 T_K}{\delta\rho(\bar{r}')\delta\rho(\bar{r})} = \frac{10}{9}C_F\rho(\bar{r})^{-1/3}\delta(\bar{r}-\bar{r}') \tag{A5}$$

hence, $\eta_K(\bar{r})$ becomes

$$\eta_K(\bar{r}) = \frac{10}{9}C_F\int \rho(\bar{r})^{-1/3}\delta(\bar{r}-\bar{r}')f(\bar{r}')d\bar{r}' \tag{B2}$$

As $f^-(\bar{r}')$ measures reactivity toward an electrophilic (El^+), so inserting Eq (1.40) in Eq (B2) one can obtain the corresponding contribution of kinetic energy functional for the local hardness (i.e., $\eta_K^-(\bar{r})$) for studies of electrophilic attack

$$\eta_K^-(\bar{r}) = \frac{10}{9}C_F\int \rho_N(\bar{r})^{-1/3}\delta(\bar{r}-\bar{r}')[\rho_N(\bar{r}')-\rho_{N-1}(\bar{r}')]d\bar{r}' \tag{B3}$$

or,
$$\eta_K^-(\bar{r}) = \frac{10}{9}C_F[\int \rho_N(\bar{r})^{-1/3}\delta(\bar{r}-\bar{r}')\rho_N(\bar{r}')d\bar{r}' - \int \rho_N(\bar{r})^{-1/3}\delta(\bar{r}-\bar{r}')\rho_{N-1}(\bar{r}')d\bar{r}']$$

or,
$$\eta_K^-(\bar{r}) = \frac{10}{9}C_F\left[\rho_N(\bar{r})^{2/3} - \rho_N(\bar{r})^{-1/3}\rho_{N-1}(\bar{r})\right] \tag{B4}$$

(as $\int \delta(\bar{r}-\bar{r}')\rho(\bar{r}')d\bar{r}' = \rho(\bar{r})$)

For studies of nucleophilic (Nu^-) attack the corresponding contribution of kinetic energy functional for the local hardness (i.e., $\eta_K^+(\bar{r})$) can obtain by using $f^+(\bar{r}')$

$$\eta_K^+(\bar{r}) = \frac{10}{9} C_F \int \rho_N(\bar{r})^{-1/3} \delta(\bar{r} - \bar{r}')[\rho_{N+1}(\bar{r}') - \rho_N(\bar{r}')]d\bar{r}' \qquad \text{(B5)}$$

because $f^+(\bar{r}') \equiv \rho_{N+1}(\bar{r}') - \rho_N(\bar{r}')$ (i.e., Eq (1.41))

So,
$$\eta_K^+(\bar{r}) = \frac{10}{9} C_F [\int \rho_N(\bar{r})^{-1/3} \delta(\bar{r} - \bar{r}')\rho_{N+1}(\bar{r}')d\bar{r}' - \int \rho_N(\bar{r})^{-1/3} \delta(\bar{r} - \bar{r}')\rho_N(\bar{r}')d\bar{r}']$$

or,
$$\eta_K^+(\bar{r}) = \frac{10}{9} C_F \left[\rho_N(\bar{r})^{-1/3} \rho_{N+1}(\bar{r}) - \rho_N(\bar{r})^{2/3}\right] \qquad \text{(B6)}$$

$$(\text{as} \int \delta(\bar{r} - \bar{r}')\rho(\bar{r}')d\bar{r}' = \rho(\bar{r}))$$

Similarly, one can write, for exchange energy functional $E_X[\rho]$ term,

$$E_X[\rho] = -C_X \int \rho(\bar{r})^{4/3} d\bar{r} \qquad \text{(A4)}$$

The corresponding contribution of exchange energy functional for the local hardness (i.e., $\eta_X^-(\bar{r})$) for studies of electrophilic attack must be

$$\eta_X^-(\bar{r}) = -\frac{4}{9} C_X \int \rho(\bar{r})^{-2/3} \delta(\bar{r} - \bar{r}') f^-(\bar{r}')d\bar{r}' \qquad \text{(B7)}$$

in order to
$$\eta_X(\bar{r}, \bar{r}') = \frac{\delta^2 E_X}{\delta\rho(\bar{r}')\delta\rho(\bar{r})} = -\frac{4}{9} C_X \rho(\bar{r})^{-2/3} \delta(\bar{r} - \bar{r}') \qquad \text{(A13)}$$

So, applying Eq (1.40) [$f^-(\bar{r}')$] in Eq (B7), one finds that

$$\eta_X^-(\bar{r}) = -\frac{4}{9} C_X \int \rho_N(\bar{r})^{-2/3} \delta(\bar{r} - \bar{r}')[\rho_N(\bar{r}') - \rho_{N-1}(\bar{r}')]d\bar{r}' \qquad \text{(B8)}$$

or,
$$\eta_X^-(\bar{r}) = -\frac{4}{9} C_X [\int \rho_N(\bar{r})^{-2/3} \delta(\bar{r} - \bar{r}')\rho_N(\bar{r}')d\bar{r}' - \int \rho_N(\bar{r})^{-2/3} \delta(\bar{r} - \bar{r}')\rho_{N-1}(\bar{r}')d\bar{r}']$$

or,
$$\eta_X^-(\bar{r}) = -\frac{4}{9} C_X \left[\rho_N(\bar{r})^{1/3} - \rho_N(\bar{r})^{-2/3} \rho_{N-1}(\bar{r})\right] \qquad \text{(B9)}$$

And inserting Eq (1.41) in Eq (B7) one can obtain the corresponding contribution of exchange energy functional for the local hardness (i.e., $\eta_X^+(\bar{r})$) for studies of nucleophilic attack

$$\eta_X^+(\bar{r}) = -\frac{4}{9} C_X \int \rho_N(\bar{r})^{-2/3} \delta(\bar{r} - \bar{r}')[\rho_{N+1}(\bar{r}') - \rho_N(\bar{r}')]d\bar{r}' \tag{B10}$$

or,
$$\eta_X^+(\bar{r}) = -\frac{4}{9} C_X [\int \rho_N(\bar{r})^{-2/3} \delta(\bar{r} - \bar{r}')\rho_{N+1}(\bar{r}')d\bar{r}' - \int \rho_N(\bar{r})^{-2/3} \delta(\bar{r} - \bar{r}')\rho_N(\bar{r}')d\bar{r}']$$

or,
$$\eta_X^+(\bar{r}) = -\frac{4}{9} C_X \left[\rho_N(\bar{r})^{-2/3} \rho_{N+1}(\bar{r}) - \rho_N(\bar{r})^{1/3}\right] \tag{B11}$$

Finally consider the Coulomb term,

$$J[\rho] = \frac{1}{2} \iint \frac{\rho(\bar{r})\rho(\bar{r}')}{|\bar{r} - \bar{r}'|} d\bar{r} d\bar{r}' \tag{A3}$$

For this we find,

$$\eta_J(\bar{r}, \bar{r}') = \frac{\delta^2 J}{\delta\rho(\bar{r})\delta\rho(\bar{r}')} = \frac{1}{|\bar{r} - \bar{r}'|} \tag{A15}$$

So,
$$\eta_J(\bar{r}) = \int \frac{1}{|\bar{r} - \bar{r}'|} f(\bar{r}')d\bar{r}' \tag{B12}$$

Now for studies of electrophilic attack, inserting Eq (1.40) ($f^-(\bar{r}')$) in Eq (B12) we obtain,

$$\eta_J^-(\bar{r}) = \int \frac{1}{|\bar{r} - \bar{r}'|} [\rho_N(\bar{r}') - \rho_{N-1}(\bar{r}')]d\bar{r}' \tag{B13}$$

or,
$$\eta_J^-(\bar{r}) = -V_N^{el}(\bar{r}) + V_{N-1}^{el}(\bar{r}) \tag{B14}$$

as
$$\int \frac{\rho(\bar{r})}{|\bar{r} - \bar{r}'|} d\bar{r} = -V^{el}(\bar{r})$$

Applying Eq (1.41) ($f^+(\bar{r}')$) in Eq (B12), one finds that $\eta_J^+(\bar{r})$ which measures reactivity toward an nucleophilic attack

$$\eta_J^+(\bar{r}) = \int \frac{1}{|\bar{r} - \bar{r}'|} [\rho_{N+1}(\bar{r}') - \rho_N(\bar{r}')]d\bar{r}' \tag{B15}$$

or,
$$\eta_J^+(\bar{r}) = -V_{N+1}^{el}(\bar{r}) + V_N^{el}(\bar{r}) \tag{B16}$$

becauses $$\int \frac{\rho(\bar{r})}{|\bar{r}-\bar{r}'|} d\bar{r} = -V^{el}(\bar{r})$$

Combining Eqs (B4), (B9) & (B14) and Eqs (B6), (B11) & (B16) we arrive Eq (3.13) and Eq (3.17) respectively.

Appendix C

LIST OF PUBLICATIONS

- From Thesis Work

1. 'One-into-Many' Model: An Approach on DFT Based Reactivity Descriptor to Predict the Regioselectivity of Large System.
   S. Saha and R. K. Roy; *J. Phys. Chem. B*, 2007, **111**, 9664.
   (ADDITIONS AND CORRECTIONS: *J. Phys. Chem. B*, 2008, **112**, 1884).
2. *N*-dependence Problem of Local Hardness Parameter.
   S. Saha and R. K. Roy; *Phys. Chem. Chem. Phys.*, 2008, **10**, 5591.
3. Can Hardness Potential Remove *N*-dependence Problem of Local Hardness?
   S. Saha and R. K. Roy (Communicated).

- Review Article

1. Studies of Regioselectivity of Large Molecular Systems using DFT Based Reactivity Descriptors.
   R. K. Roy and S. Saha; *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.*, 2010,
   DOI: 10.1039/b811052m (In Press).

- Other Publications

1. Are the Hirshfeld and Mulliken Population Analysis Schemes Consistent with Chemical Intuition?
   S. Saha, R. K. Roy and P. W. Ayers; *Int. J. Quantum Chem.*, 2009, **109**, 1790.
2. A Comprehensive Decomposition Analysis of Stabilization Energy (CDASE) and its Application in Locating the Rate Determining Step of Multi-step Reactions.
   P. Bagaria, S. Saha, S. Murru, V. Kavala, B. K. Patel and R. K. Roy; *Phys. Chem. Chem. Phys.*, 2009, **11**, 8306.
3. CDASE-A Reliable Scheme to explain the Reactivity Sequence between Diels-Alder Pairs.
   S. Saha, R. K. Roy and S. Pal, *Phys. Chem. Chem. Phys.*, 2010, DOI: 10.1039/B925441B (Accepted).

BRIEF BIOGRAPHY OF Dr. RAM KINKAR ROY

Dr. Ram Kinkar Roy is an Assistant Professor of Chemistry at Birla Institute of Technology and Science (BITS), Pilani. He received his Ph. D. degree from National Chemical Laboratory, Pune in 1996. For his doctoral degree, he worked with Professor Saurav Pal in the research area of theoretical quantum chemistry. After his doctorate, he spent a year at National Chemical Laboratory, Pune, working with Prof. Saurav Pal as a Research Associate (CSIR). As a visiting post-doctoral fellow, he worked with Prof. Paul Geerlings for a year (1997-1998) at Vrije Universiteit Brussel, Belgium and also with Prof. Kimihiko Hirao in 1998-2001 at the University of Tokyo, Tokyo. He joined BITS, Pilani in 2001 and since then he has been here. He has been involved in Research for the last 19 years and teaching for 9 years. As a result of his research accomplishment, he has 31 international publications. He has handled one DST project, and is presently handling another/second DST project as well. Dr. Roy has successfully guided one Ph.D. student and is currently guiding another one.

BRIEF BIOGRAPHY OF Mr. SOUMEN SAHA

Mr. Soumen Saha received his B.Sc. degree (Chemistry Honors) in 2001 from the University of Burdwan and his M.Sc. degree (Chemistry) from the Bengal Engineering College (A Deemed University), Shibpur in 2003 (Result published on March, 2004). In 2005, he joined the research group of Dr. Ram Kinkar Roy for his Ph.D. degree in the Chemistry Group, Birla Institute of Technology and Science (BITS), Pilani. At present he is a Senior Research Fellow in CSIR-SRF scheme.