Studies of regioselectivity of large molecular systems using DFT based reactivity descriptors

Ram Kinkar Roy* and Soumen Saha

DOI: 10.1039/b811052m

This report describes the recent works on Conceptual Density Functional Theory (DFT) based reactivity descriptors used to predict the regioselectivity of large systems, biomolecular systems, in particular. The challenges of bio-systems, 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, fragmentation based method is developed recently (given the name 'One-into-Many' model) for a reliable determination of the regioselectivity of biomolecular systems. Thus, our main motivation to embark on the endeavor of this report is to provide a brief introduction of Conceptual DFT and fragmentation approaches based on these reactivity descriptors for predicting the regioselectivity of large biomolecular systems.

1. Introduction

An important chemical concept prevalent in chemistry (organic chemistry, in particular) is regioselectivity. Regioselectivity[1,2] is defined as the preference of a chemical bond making or breaking in one direction over all other possible ones. Understanding the regioselectivity of a reaction between two chemical species is crucial not only for predicting the corresponding reaction mechanism but also for designing desired products. 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*., were proposed and extensively used 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 behaviours. However, most of these principles remained empirical until a branch of density functional theory (DFT),[29–41] called "Conceptual DFT" or "Chemical Reactivity Theory", has been initiated by its protagonist, 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 to formal definitions of empirical concepts.[42–52] Conceptual DFT was even successful 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

Department of Chemistry, Birla Institute of Technology and Science, Pilani-333031, Rajasthan, India. E-mail: rkroy2@rediffmail.com

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 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 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 shared by the fragment-based approaches is to divide a large molecule into a series of fragments (rather than treating the whole system at once), 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. Furthermore, 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, and then this density matrix is 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 can be approximated as a sum of bond energies or enthalpies. In comparison with those 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]

However, for predicting the regioselectivity of large molecular systems, a simple fragment-based approach, named as "One-into-Many" model, was proposed by us.[114,115] In this model a large system is proposed to be broken into different smaller 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, thus helping the prediction of regioselectivity of the large system.

In this report, we will present some advancement achieved in last few years to predict the regioselectivity of the large bio-molecular systems using Conceptual DFT

(or chemical reactivity theory) based reactivity descriptors with some review of its basic aspects. 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. The limitations of the present approach in predicting the regioselectivity of target systems (*e.g.*, large chemical and biological molecules) as well as likely developments over the next few years will also be discussed. In the first part of our presentation, the Conceptual DFT is described briefly, which provides the theoretical foundations of different reactivity descriptors. The second part takes care of more recent developments enabling evaluations of the regioselectivity for a number of large biological systems using these reactivity descriptors. Finally, in the last part of our presentation we have summarized the whole report.

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

$$E = E[\rho] \tag{1}$$

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[116]

$$E[\rho] = F[\rho] + \int v(\bar{r})\rho(\bar{r})\mathrm{d}\bar{r} \tag{2}$$

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 it's evaluation was developed by Kohn and Sham.[30] In this second theorem, a variational principle is formulated, stating that the ground state density is that density that minimizes the energy of the system for a fixed number of electrons

$$\delta(E[\rho] - \mu\int\rho(\bar{r})\mathrm{d}\bar{r}) = 0 \tag{3}$$

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

$$\mu = v(\bar{r}) + \frac{\delta F}{\delta\rho(\bar{r})} = \text{constant} \tag{4}$$

Kohn and Sham rewrote eqn (4) as an orbital equation having the form[30]

$$\left[-\frac{1}{2}\nabla^2 + v(\bar{r}) + v_{xc}(\bar{r}) + \int\frac{\rho(\bar{r}')}{|\bar{r}-\bar{r}'|}\mathrm{d}\bar{r}'\right]\psi_i = \varepsilon_i\psi_i \tag{5}$$

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

$$v_{xc}(\bar{r}) = \frac{\delta E_{xc}}{\delta \rho(\bar{r})} \tag{6}$$

In eqn (5), ψ_i's are the Kohn-Sham orbitals, the squares of which must sum up to the total electron density of the system

$$\rho(\bar{r}) = \sum_i |\psi_i|^2 \tag{7}$$

In recent years, many accurate forms of the exchange correlation functional were derived.[117–120] However, 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 reactivity indices. The associated reactivity descriptors that arise from differentiation with respect to N (but not the external potential $v(\bar{r})$) are called *global reactivity descriptors*: they are associated with the overall reactivity of the molecule and do not contain any information about regioselectivity. Coefficients that contain exactly one differentiation with respect to the external potential $v(\bar{r})$ are said to be *local reactivity descriptors* because they vary locally from one position to another in a molecule. Local reactivity descriptors provide key information about the relative reactivity of different sites in a molecule. So the local reactivity descriptors are key in making predictions about regioselectivity. Coefficients that contain exactly two or more differentiations with respect to the external potential are called *nonlocal reactivity descriptors* or *reactivity kernels*. 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. The important aspect of this presentation 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 eqn (4). It has been Parr's impressive contribution[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)

$$\mu = \left(\frac{\partial E}{\partial N}\right)_{v(\bar{r})} \tag{8}$$

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.[121,122] Interestingly, Iczkowski and Margrave,[123] in an important contribution to the literature of electronegativity have defined the electronegativity (χ) of a system by the following,

$$\chi = -\left(\frac{\partial E}{\partial N}\right) \tag{9}$$

Mulliken's[124] definition of electronegativity is given as the arithmetic average of two experimentally measurable quantities, *i.e.*, ionization potential (IP) and electron affinity (EA):

$$\chi = \frac{IP + EA}{2} \tag{10}$$

The expression is just the finite difference approximation to the term, $-\left(\frac{\mathrm{d}E}{\mathrm{d}N}\right)$. 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:

$$\chi = -\mu = -\left(\frac{\partial E}{\partial N}\right)_{v(\bar{r})} \tag{11}$$

The idea that electronegativity is a chemical potential originates with Gyftopoulos and Hatsopoulos.[125]

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.

$$\mu^{-} = \frac{E(N-1) - E(N)}{-1} = -IP \tag{12}$$

$$\mu^{+} = E(N+1) - E(N) = -EA \tag{13}$$

$$\left(\frac{\partial E}{\partial N}\right)_{v(\bar{r})} = \mu = \frac{1}{2}(\mu^{+} + \mu^{-}) = -\frac{1}{2}(IP + EA) \tag{14}$$

Thus, from eqn (11) electronegativity (χ) can be written as

$$\chi = -\mu = \tfrac{1}{2}(IP + EA) \tag{15}$$

The expression of χ originated from here is similar to that of Mulliken (*i.e.*, eqn (11)).[124] As an approximation to eqn (15), one can relate chemical potential (μ) to the frontier orbital energies. This can be obtained through the Koopmanns' approximation[126,127] 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,[53,128–133]

$$-E_{\mathrm{HOMO}} = IP \tag{16}$$

$$-E_{\mathrm{LUMO}} = EA \tag{17}$$

Therefore, using Koopmanns' theorem,[126] we can write

$$\mu = -\chi = \frac{E_{\mathrm{LUMO}} + E_{\mathrm{HOMO}}}{2} \tag{18}$$

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 eqn (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* eqn (18).[134–136]

Moreover, theoretical justification was provided for Sanderson's principle of electronegativity equalization[23,26,137,138] 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]

$$\eta = \left(\frac{\partial^2 E}{\partial N^2}\right)_{v(\bar{r})} = \left(\frac{\partial \mu}{\partial N}\right)_{v(\bar{r})} \tag{19}$$

The chemical hardness is 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).[18,19,54,139–142] 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.[143,144] Again, using a finite difference approximation and a quadratic $E = E(N)$ curve, this equation reduces to

$$\eta = IP - EA \tag{20}$$

which, after using Koopmans approximation,[126] becomes

$$\eta = \varepsilon_{\mathrm{LUMO}} - \varepsilon_{\mathrm{HOMO}} \tag{21}$$

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.[145] 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 a combination of statistical mechanics and the fluctuation-dissipation theorem.[55,146–149]

The inverse of the global hardness is called the global softness[143,140]

$$S = \frac{1}{\eta} = \left(\frac{\partial N}{\partial \mu}\right)_{v(\vec{r})} \tag{22}$$

which was empirically shown to be proportional to the polarizability of the system.[149–155] 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,

$$E_A = E_A^o + \mu_A^o(N_A - N_A^o) + \tfrac{1}{2}\eta_A(N_A - N_A^o)^2 + \cdots \tag{23a}$$

$$E_B = E_B^o + \mu_B^o(N_B - N_B^o) + \tfrac{1}{2}\eta_B(N_B - N_B^o)^2 + \cdots \tag{23b}$$

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

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

or

$$(E_A + E_B) - (E_A^o + E_B^o) = \Delta E_A + \Delta E_B = \Delta(E_A + E_B) = (\mu_A^o - \mu_B^o)\Delta N + \tfrac{1}{2}(\eta_A + \eta_B)(\Delta N)^2 + \cdots \tag{24}$$

where,

$$\Delta N = N_B^o - N_B = N_A - N_A^o \tag{25}$$

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 eqn (24) one can write,

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

or

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

or

$$\mu_A = \mu_B \tag{26}$$

where,

$$\mu_A = \left(\frac{\partial E_A}{\partial N_A}\right)_{v(\vec{r})} = \mu_A^o + \eta_A \Delta N \tag{27a}$$

$$\mu_B = \left(\frac{\partial E_B}{\partial N_B}\right)_{v(\vec{r})} = \mu_B^o + \eta_B \Delta N \tag{27b}$$

Hence, from eqn (26), (27a) and (27b), we can write,

$$\Delta N = \frac{\mu_B^o - \mu_A^o}{(\eta_A + \eta_B)} \tag{28}$$

Substituting the values of ΔN from eqn (28) in eqn (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

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

$$\Delta E_{SE} = \Delta E_A + \Delta E_B = -\frac{(\mu_B^o - \mu_A^o)^2}{2(\eta_A + \eta_B)} \tag{29}$$

Here, eqn (29) represents the stabilization energy due to transfer of ΔN amount of electron from B to A (from eqn (29) it is obvious that ΔE_{SE} is negative *i.e.*, energy is lowered due to charge transfer).[20,156]

Another global reactivity descriptor is global electrophilicity (w), also proposed by Parr *et al.*[157] while tried to validate the experimental findings of Maynard *et al.*[158] 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 (*e.g.*, 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 $v(\vec{r})$ is

$$\Delta E = \mu \Delta N + \eta \frac{\Delta N^2}{2} \tag{30}$$

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

$$\frac{\Delta E}{\Delta N} = 0 \tag{31}$$

implying

$$\Delta N_{\max} = -\frac{\mu}{\eta} \tag{32}$$

which yields stabilization energy,

$$\Delta E = -\frac{\mu^2}{2\eta} \tag{33}$$

In eqn (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 $\left(\text{power} \equiv W = -\frac{V^2}{R}\right)$, Parr *et al.*[157] defined $w = -\Delta E$ as a measure of electrophilicity of the system (atom, molecule, or ion). The resulting equation is

$$w = \frac{\mu^2}{2\eta} \tag{34}$$

This quantity w is called the "electrophilicity index". Kinetic and thermodynamic aspects of w was investigated by Chattaraj and collaborators[159] 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.[160] 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 $\left(\frac{\Delta E}{\Delta N}\right)_v = 0$, $\Delta E \approx -w = -\frac{\mu^2}{2\eta}$. 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 eqn (34)) in terms of first vertical IP and first vertical EA as (by using eqn (14) and (20)),

$$w = \frac{\mu^2}{2\eta} = \frac{[-(IP + EA)/2]^2}{2(IP - EA)} = \frac{(IP + EA)^2}{8(IP - EA)} \tag{35}$$

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[160] 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, it is 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.*[161] extended the use of global electrophilicity descriptor, as proposed by Parr *et al.*,[157] 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.*, $|\Delta E_{A(B)}|$ and $|\Delta E_{B(A)}|$) together with the global electrophilicity descriptor of the acceptor (w_A), could explain the rate determining step of a multistep chemical reaction.[161] They also showed that eqn (33) is a special case of eqn (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,[162–166] electrofugality and nucleofugality,[167–169] potentialphilicity and potentialphobicity,[170] chargephilicity and chargephobicity[171] are also proposed recently, which are all conceptually related to w.

(ii) Local reactivity descriptor. 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. So, local reactivity descriptors are the key to determining which places 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,

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

or,

$$\mathrm{d}E = \mu \mathrm{d}N + \int \left(\frac{\delta E}{\delta v(\bar{r})}\right)_N \delta v(\bar{r}) \mathrm{d}\bar{r} \tag{36}$$

thus, one has the most fundamental local reactivity descriptor, the ground state electron density $\rho(\bar{r})$[31,172–178]

$$\rho(\bar{r}) = \left(\frac{\delta E}{\delta v(\bar{r})}\right)_N \tag{37}$$

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

$$\mathrm{d}\mu = \left(\frac{\partial \mu}{\partial N}\right)_{v(\bar{r})} \mathrm{d}N + \int \left(\frac{\delta \mu}{\delta v(\bar{r})}\right)_N \delta v(\bar{r}) \mathrm{d}\bar{r} \tag{38a}$$

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

$$\mathrm{d}\mu = \eta \mathrm{d}N + \int \left(\frac{\delta \mu}{\delta v(\bar{r})}\right)_N \delta v(\bar{r}) \mathrm{d}\bar{r} \tag{38b}$$

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

$$f(\bar{r}) = \left(\frac{\delta \mu}{\delta v(\bar{r})}\right)_N = \left(\frac{\partial \rho(\bar{r})}{\partial N}\right)_{v(\bar{r})} \tag{39}$$

which is known as Fukui Function (FF).[143,179–181] This quantity integrates to unity, $\int f(\bar{r})\mathrm{d}\bar{r} = 1$. The second formula for $f(\bar{r})$ in eqn (39) is a "Maxwell relation"[182] following from the fact that $\mathrm{d}E$ is an exact differential. There is a discontinuity[179,183,184] in the derivative of the Fukui function just as there is for chemical potential.[185] 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[179] have defined the left $[f^-(\bar{r})]$, right $[f^+(\bar{r})]$ and central $[f^0(\bar{r})]$ derivatives of eqn (39).

These three Fukui functions can be written by applying a finite difference approximation and the frontier-orbital theory[3–6] of reactivity as,

$$f^{-}(\bar{r}) \cong \rho_N(\bar{r}) - \rho_{N-1}(\bar{r}) \approx \rho_{\mathrm{HOMO}}(\bar{r})$$ measures reactivity toward an electrophilic (El^{+}) reagent (derivative as N increases from $N - \delta \rightarrow N$), (40)

$$f^{+}(\bar{r}) \cong \rho_{N+1}(\bar{r}) - \rho_N(\bar{r}) \approx \rho_{\mathrm{LUMO}}(\bar{r})$$ measures reactivity toward a nucleophilic (Nu^{-}) reagent (derivative as N increases from $N \rightarrow N + \delta$), (41)

and

$$f^{0}(\bar{r}) = \tfrac{1}{2}[f^{+}(\bar{r}) + f^{-}(\bar{r})] \cong \tfrac{1}{2}[\rho_{N+1}(\bar{r}) - \rho_{N-1}(\bar{r})] \approx \tfrac{1}{2}[\rho_{\mathrm{HOMO}}(\bar{r}) + \rho_{\mathrm{LUMO}}(\bar{r})]$$ measures reactivity toward an innocuous (radical) reagent (mean of left and right derivatives) (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[186] 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

$$f^{-}(k) \cong P_N(k) - P_{N-1}(k) \quad (43)$$

$$f^{+}(k) \cong P_{N+1}(k) - P_N(k) \quad (44)$$

and

$$f^{0}(k) \cong \tfrac{1}{2}[P_{N+1}(k) - P_{N-1}(k)] \quad (45)$$

where $P(k)$ denotes the electronic population on atom k. Parr and Yang[179] 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.[187–202] 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)[203] (also known as stockholders charge-partitioning technique), as proposed by Hirshfeld is used and shown that HPA yields only positive Fukui functions.[59,162,187,191,192,194,204] 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),[205] natural bond orbital (NBO) analysis,[206–209] 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.*[187,191,192] clearly demonstrated that HPA is superior to other charge-partitioning schemes. Subsequently, there are quite a number of studies in this area,[189,194,210] 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.[188–190,193–196,198–202,204,211,212] 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.*,[187] 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 one and not a generalized method.

In order to mitigate the problems associated with the above Hirshfeld scheme, recently in 2007, Bultinck *et al.*[213] have proposed an alternative, iterative version of the Hirshfeld partitioning procedure, known as "Hirshfeld-I" method. They have verified this method on the 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 eliminates arbitrariness in the choice of the promolecule, so the atomic populations are determined solely by the molecular electron density, increases the magnitudes of the charges, and 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[214] 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.[215–217] This is important in view of the fact that although the problem of negative Fukui function indices has been looked upon in detail, no definitive answers has been given yet to the question whether negative values are physically acceptable or are artifacts. Thus, the occurrence of negative Fukui function has remained a puzzle for a long time. And according to some computational studies, it is truly impossible to exclude negative Fukui function.[189,218–221] It has been pointed out that the possibility of negative atom condensed Fukui function values depend critically on the properties of the hardness matrix.[195,204,215,222]

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.[223] 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, $s(\bar{r})$[143]

$$s(\bar{r}) = \left(\frac{\partial \rho(\bar{r})}{\partial \mu}\right)_{v(\bar{r})} = \left(\frac{\partial \rho(\bar{r})}{\partial N}\right)_{v(\bar{r})} \left(\frac{\partial N}{\partial \mu}\right)_{v(\bar{r})} = f(\bar{r})S \tag{46}$$

where S is global softness (vide eqn (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,

$$\int s(\bar{r})\mathrm{d}\bar{r} = S \tag{47}$$

The Fukui function in eqn (46) can be identified with the Fukui function from above (eqn (40)), the Fukui function from below (eqn (41)), or from the average of the two (eqn (42)). Similarly, the three approximate atomic $f(\bar{r})$ indices (from eqn (43)–(45)), when multiplied by S, provide three different local softnesses for any particular atom (k). These can be written as

$$s^{-}(k) \cong [P_N(k) - P_{N-1}(k)]S \text{ (suited for studies of electrophilic attack)} \tag{48}$$

$$s^{+}(k) \cong [P_{N+1}(k) - P_N(k)]S \text{ (suited for studies of nucleophilic attack)} \tag{49}$$

and

$$s^{0}(k) \cong \tfrac{1}{2}[P_{N+1}(k) - P_{N-1}(k)]S \text{ (suited for studies of radical attack)} \tag{50}$$

From eqn (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.[143] 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).[224] 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.[143] This formula and other similar DFT fluctuation formulae[225,226] 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.[143,227]

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.*[228] 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[229] have given a derivation of local softness that reveals its relation to its reciprocal property, local hardness.[230–232]

Substitution of eqn (4) into eqn (2) follows, for a ground state

$$\begin{aligned} E[\rho] &= N\mu - \left[\int \frac{\delta F[\rho]}{\delta\rho(\bar{r})}\rho(\bar{r})\mathrm{d}\bar{r} - F[\rho]\right] \\ &= N\mu - H[\rho] \end{aligned} \tag{51}$$

where the *hardness functional* $H[\rho]$ is defined by the formula[233]

$$H[\rho] = \int \frac{\delta F[\rho]}{\delta\rho(\bar{r})}\rho(\bar{r})\mathrm{d}\bar{r} - F[\rho] \tag{52}$$

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

The total differential of eqn (51), associated with a change of the system of interest, is simply

$$\mathrm{d}E[\rho] = N\mathrm{d}\mu + \mu\mathrm{d}N - \mathrm{d}H[\rho] \tag{53}$$

Comparing this with eqn (36), one can obtain the equation

$$N\mathrm{d}\mu = \mathrm{d}H[\rho] + \int\rho(\bar{r})\delta v(\bar{r})\mathrm{d}\bar{r} \tag{54}$$

The differential of $H[\rho]$ has a surprisingly simple form. From eqn (52) one has[42]

$$\begin{aligned} \mathrm{d}H[\rho] &= -\mathrm{d}F[\rho] + \mathrm{d}\left[\int \frac{\delta F[\rho]}{\delta\rho(\bar{r})}\rho(\bar{r})\mathrm{d}\bar{r}\right] = -\int \frac{\delta F[\rho]}{\delta\rho(\bar{r})}\mathrm{d}\rho(\bar{r})\mathrm{d}\bar{r} + \int \frac{\delta F[\rho]}{\delta\rho(\bar{r})}\mathrm{d}\rho(\bar{r})\mathrm{d}\bar{r} \\ &+ \int\left[\mathrm{d}\left(\frac{\delta F[\rho]}{\delta\rho(\bar{r})}\right)\right]\rho(\bar{r})\mathrm{d}\bar{r} = \int\int \frac{\delta^2 F[\rho]}{\delta\rho(\bar{r})\delta\rho(\bar{r}')}\mathrm{d}\rho(\bar{r}')\rho(\bar{r})\mathrm{d}\bar{r}\mathrm{d}\bar{r}' \end{aligned} \tag{55}$$

So, from eqn (53),

$$\mathrm{d}(E[\rho] - N\mu) = -\int\int \frac{\delta^2 F[\rho]}{\delta\rho(\bar{r})\delta\rho(\bar{r}')}\mathrm{d}\rho(\bar{r}')\rho(\bar{r})\mathrm{d}\bar{r}\mathrm{d}\bar{r}' \tag{56}$$

This is where the local hardness comes in. Following Ghosh and Berkowitz,[231] local hardness $\eta(\bar{r})$ is defined as:[230]

$$\eta(\bar{r}) = \frac{1}{N}\int \frac{\delta^2 F}{\delta\rho(\bar{r}')\delta\rho(\bar{r})}\rho(\bar{r}')\mathrm{d}\bar{r}' \tag{57}$$

One can find accurate to all orders,

$$\frac{\delta H[\rho]}{\delta\rho(\bar{r})} = N\eta(\bar{r}) = h(\bar{r}) \tag{58}$$

Parr and Gázquez called $h(\bar{r})$ the hardness potential for the system.[233]

Introducing the symbol of $\eta(\bar{r})$, one can rewrite eqn (56) as

$$\mathrm{d}(E[\rho] - N\mu) = -N\int\eta(\bar{r})\mathrm{d}\rho(\bar{r})\mathrm{d}\bar{r} \tag{59}$$

Eliminating $\mathrm{d}E$ from eqn (36) and (59), was obtained

$$\mathrm{d}\mu = \int \eta(\bar{r})\mathrm{d}\rho(\bar{r})\mathrm{d}\bar{r} + \frac{1}{N}\int \rho(\bar{r})\delta v(\bar{r})\mathrm{d}\bar{r} \tag{60}$$

This is the local counterpart of eqn (38) in the sense of Nalewajski,[44] in which it now appears the local hardness in place of the global hardness.

From eqn (60) one can find another formula for $\eta(\vec{r})$,[230]

$$\eta(\vec{r}) = \left(\frac{\delta\mu}{\delta\rho(\vec{r})}\right)_{v(\vec{r})} \tag{61}$$

Eqn (61) is an example of an ambiguous "constrained functional derivative".[225,230,234–251] The functional derivative is ambiguous because of the interdependence of $\rho(\vec{r})$ and $v(\vec{r})$.[225] 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:

$$\begin{aligned}\eta &= \left(\frac{\partial^2 E}{\partial N^2}\right)_{v(\vec{r})} = \left(\frac{\partial\mu}{\partial N}\right)_{v(\vec{r})} \\ &= \int\left(\frac{\delta\mu}{\delta\rho(\vec{r})}\right)_{v(\vec{r})}\left(\frac{\delta\rho(\vec{r})}{\delta N}\right)_{v(\vec{r})}\mathrm{d}\vec{r} = \int \eta(\vec{r})f(\vec{r})\mathrm{d}\vec{r}\end{aligned} \tag{62}$$

An explicit expression for $\eta(\vec{r})$ can be deduced from eqn (57) and the definition of Fukui function (eqn (39))[234]

$$\eta(\vec{r}) = \int \frac{\delta^2 F}{\delta\rho(\vec{r}')\delta\rho(\vec{r})} f(\vec{r}')\mathrm{d}\vec{r}' \tag{63}$$

Local hardness and local softness are reciprocals in the sense that

$$\int \eta(\vec{r})s(\vec{r})\mathrm{d}\vec{r} = 1 \tag{64}$$

One can simplify the definitions of local hardness by writing[236]

$$\eta_\lambda(\vec{r}) = \left(\frac{\delta\mu}{\delta\rho(\vec{r})}\right)_{v(\vec{r})} = \frac{1}{N}\int \frac{\delta^2 F}{\delta\rho(\vec{r}')\delta\rho(\vec{r})}\lambda[\rho(\vec{r}')]\mathrm{d}\vec{r}' \tag{65}$$

where, $\lambda[\rho(\vec{r}')]$[225] is a composite function that integrates to N (*i.e.*, total number of electrons of the system),

$$\int \lambda[\rho(\vec{r}')]\mathrm{d}\vec{r} = N \tag{66}$$

Two important choices of the composite function $\lambda[\rho(\vec{r}')]$ are $\rho(\vec{r}')$[114,115,231,236,245,246,248,249,252] and $Nf(\vec{r}')$,[234,237,239,241,243,247,250,251] when the following possibilities emerge:

$$\lambda[\rho(\vec{r}')] = \rho(\vec{r}') \text{ yielding } \tilde{\eta}_D(\vec{r}) = \frac{1}{N}\int \frac{\delta^2 F}{\delta\rho(\vec{r}')\delta\rho(\vec{r})}\rho(\vec{r}')\mathrm{d}\vec{r}' \tag{67}$$

and

$$\lambda[\rho(\vec{r}')] = Nf(\vec{r}') \text{ yielding } \tilde{\eta}_F(\vec{r}) = \int \frac{\delta^2 F}{\delta\rho(\vec{r}')\delta\rho(\vec{r})} f(\vec{r}')\mathrm{d}\vec{r}' \tag{68}$$

But $\tilde{\eta}_F(\vec{r})$ is shown to be equal to the global hardness η at every point of space,[225] when the exact functional $F[\rho]$ is used[253] in eqn (68)

$$\begin{aligned}\eta = \left(\frac{\partial\mu}{\partial N}\right)_{v(\vec{r})} &= \frac{\partial}{\partial N}\left(\frac{\delta F[\rho]}{\delta\rho(\vec{r})}\right) = \int \frac{\delta^2 F[\rho]}{\delta\rho(\vec{r})\delta\rho(\vec{r}')}\left(\frac{\partial\rho(\vec{r}')}{\partial N}\right)\mathrm{d}\vec{r}' \\ &= \int \frac{\delta^2 F[\rho]}{\delta\rho(r)\delta\rho(\vec{r}')} f(\vec{r}')\mathrm{d}\vec{r}' = \tilde{\eta}_F(\vec{r})\end{aligned} \tag{69}$$

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".[140]

Based on the global electrophilicity index w (eqn (34)) as defined by Parr *et al.*,[157] 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 eqn (34) and using the inverse relationship between global hardness and global softness[143] (eqn (22)) one may obtain

$$w = \frac{\mu^2}{2\eta} = \frac{\mu^2}{2} S = \frac{\mu^2}{2} \sum_k s^+(k) = \sum_k w(k) \tag{70}$$

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

$$w = \int w(\bar{r}) \mathrm{d}\bar{r} \tag{71}$$

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

$$w = w \int f(\bar{r}) \mathrm{d}\bar{r} = \int w f(\bar{r}) \mathrm{d}\bar{r} = \int w(\bar{r}) \mathrm{d}\bar{r} \tag{72}$$

where

$$w(\bar{r}) = w f(\bar{r}) \tag{73}$$

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

$$w^\alpha(\bar{r}) = w f^\alpha(\bar{r}) \tag{74}$$

where $\alpha = +, -$, and 0 for attacks by a nucleophile, electrophile, and radical, respectively. It is obvious that eqn (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

$$w^\alpha(k) = w f^\alpha(k) \tag{75}$$

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:

$$w(\bar{r}) = w f(\bar{r}) = \frac{\mu^2}{2\eta} f(\bar{r}) = \mu^2 S f(\bar{r}) = \mu^2 s(\bar{r}) \tag{76}$$

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.*,[223] have discussed the 'extensive', 'intensive' and 'subintensive' nature of DFT based reactivity descriptors.

(iii) Nonlocal reactivity descriptors. These are reactivity descriptors which depend on two or more spatial positions, $\bar{r},\bar{r}'$, *etc*. Interest on 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.[229]

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 (eqn (4)),

$$u(\bar{r}) = v(\bar{r}) - \mu = -\frac{\delta F[\rho]}{\delta \rho(\bar{r})} \tag{77}$$

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, $\eta(\bar{r},\bar{r}')$[44,143,227,229,263]

$$\eta(\bar{r},\bar{r}') \equiv -\frac{\delta u(\bar{r}')}{\delta \rho(\bar{r})} = -\frac{\delta u(\bar{r})}{\delta \rho(\bar{r}')} = \frac{\delta^2 F_E}{\delta \rho(\bar{r}')\delta \rho(\bar{r})} \tag{78}$$

the last equality coming from eqn (77). Recall the first definition of local hardness, $\eta(\bar{r})$ (eqn (57)), by introducing the symbol of $\eta(\bar{r},\bar{r}')$, one can find

$$\eta(\bar{r}) \equiv \frac{1}{N}\int \eta(\bar{r},\bar{r}')\rho(\bar{r}')\mathrm{d}\bar{r}' \tag{79}$$

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 $s(\bar{r},\bar{r}')$[44,143,227,229,263]

$$s(\bar{r},\bar{r}') \equiv -\frac{\delta \rho(\bar{r}')}{\delta u(\bar{r})} = -\frac{\delta \rho(\bar{r})}{\delta u(\bar{r}')} \tag{80}$$

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

$$s(\bar{r}) \equiv \int s(\bar{r},\bar{r}')\mathrm{d}\bar{r}' \tag{81}$$

Moreover, since both the functional derivatives exist,

$$\int \frac{\delta\rho(\bar{r})}{\delta u(\bar{r}')}\frac{\delta u(\bar{r}')}{\delta\rho(\bar{r}'')}\mathrm{d}\bar{r}' = \delta(\bar{r}'' - \bar{r}) \tag{82}$$

so that

$$\int s(\bar{r},\bar{r}')\eta(\bar{r}',\bar{r}'')\mathrm{d}\bar{r}' = \delta(\bar{r}'' - \bar{r}) \tag{83}$$

The hardness and softness kernels are true inverses.

Multiplying eqn (83) by $\rho(\bar{r}'')$ then integrating over $\bar{r}''$ and making use of eqn (79) one can write,

$$\int s(\bar{r},\bar{r}')\eta(\bar{r}')\mathrm{d}\bar{r}' = \frac{\rho(\bar{r})}{N} \tag{84}$$

Integrate this over $\bar{r}$, and employing eqn (81), gives previous eqn (64)

$$\int \eta(\bar{r})s(\bar{r})\mathrm{d}\bar{r} = 1 \tag{64}$$

To achieve eqn (46), writing

$$\begin{aligned}
\mathrm{d}\rho(\bar{r}) &= \int \frac{\delta\rho(\bar{r})}{\delta u(\bar{r}')}\mathrm{d}u(\bar{r}')\mathrm{d}\bar{r}' \\
&= -\int s(\bar{r},\bar{r}')\mathrm{d}u(\bar{r}')\mathrm{d}\bar{r}' \text{ (by using eqn (80))} \\
&= -\int s(\bar{r},\bar{r}')[\delta v(\bar{r}') - \mathrm{d}\mu]\mathrm{d}\bar{r}' \text{ (applying eqn (77))} \\
&= \left[\int s(\bar{r},\bar{r}')\mathrm{d}\bar{r}'\right]\mathrm{d}\mu - \int s(\bar{r},\bar{r}')\delta v(\bar{r}')\mathrm{d}\bar{r}'
\end{aligned} \tag{85}$$

utilizing $s(\bar{r}) \equiv \int s(\bar{r},\bar{r}')\mathrm{d}\bar{r}'$ from eqn (38b) and (39) we get,

$$\mathrm{d}\mu = \eta\mathrm{d}N + \int f(\bar{r}')\delta v(\bar{r}')\mathrm{d}\bar{r}' \tag{38c}$$

generating

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

or,

$$\mathrm{d}\rho(\bar{r}) = s(\bar{r})\eta\mathrm{d}N + \int[-s(\bar{r},\bar{r}') + s(\bar{r})f(\bar{r}')]\delta v(\bar{r}')\mathrm{d}\bar{r}' \tag{86}$$

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

$$\mathrm{d}\rho(\bar{r}) = f(\bar{r})\mathrm{d}N + \int \left[\frac{\delta\rho(\bar{r})}{\delta v(\bar{r}')}\right]_N \delta v(\bar{r}')\mathrm{d}\bar{r}' \tag{87}$$

N and $v(\bar{r})$ being independent, coefficients of $\mathrm{d}N$ and $\delta v(\bar{r})$ in eqn (86) and (87) must be equal. Consequently, we have

$$s(\bar{r}) = \frac{f(\bar{r})}{\eta} = f(\bar{r})S = \left(\frac{\partial\rho(\bar{r})}{\partial N}\right)_{v(\bar{r})}\left(\frac{\partial N}{\partial\mu}\right)_{v(\bar{r})} = \left(\frac{\partial\rho(\bar{r})}{\partial\mu}\right)_{v(\bar{r})} \tag{46}$$

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[229]

$$\chi(\bar{r},\bar{r}') = \left[\frac{\delta\rho(\bar{r})}{\delta v(\bar{r}')}\right]_N = -s(\bar{r},\bar{r}') + \frac{s(\bar{r})s(\bar{r}')}{S} \tag{88}$$

Eqn (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

$$\rho(\bar{r}) = \rho_\alpha(\bar{r}) + \rho_\beta(\bar{r}) \tag{89}$$

and

$$\rho_S(\bar{r}) = \rho_\alpha(\bar{r}) - \rho_\beta(\bar{r}) \tag{90}$$

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

$$N = N_\alpha + N_\beta = \int\rho(\bar{r})\mathrm{d}\bar{r} \tag{91}$$

$$N_S = N_\alpha - N_\beta = \int\rho_S(\bar{r})\mathrm{d}\bar{r} \tag{92}$$

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 put 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.[226,227,296–305]

In the light of the above discussion on DFT based reactivity descriptors we will try to analyze the regioselectivity criteria of large molecular systems in the next section. Although these reactivity indices have become very useful in predicting the regioselectivity of chemical reactions, for tracing the proper reactivity descriptor to explain the intermolecular reactivity trend the argument still continues. Therefore, in the next section, first we will explore the suitable reactivity descriptor for describing the intermolecular reactivity trend as well as its feasibility for computing large systems.

After carrying out the above task we will contemplate the regioselectivity for a number of large biological systems.

3. Regioselectivity of large system in the context of conceptual DFT

A On the way of detecting suitable intermolecular reactivity index

In 1968, G. Klopman[306] attempted to quantify Pearson's HSAB principle[18] using polyelectronic perturbation theory and for that he defined two types of interaction 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 is realized that orbital-controlled reactivity descriptors include Fukui function index [$f(\bar{r})$ or $f(k)$],[143,179,181,307] local softness [$s(\bar{r})$ or $s(k)$],[143] philicity [$w(k)$],[255] 'relative electrophilicity' (s_k^+/s_k^-) and 'relative nucleophilicity' indices (s_k^-/s_k^+),[228] whereas local hardness, $\eta(\bar{r})$[230–232] (when evaluated through Thomas-Fermi-Dirac (TFD) approach[308–310]), 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.*, $\int f(\bar{r})\mathrm{d}\bar{r} = 1$ or $\sum_k^N f(k) = 1$)[45] 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 $s(k)$]. The reasoning behind this demand is that local softness is 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. Furthermore, as $s(\bar{r}) = f(\bar{r})S$ (eqn (46)) local softness seems to be a much more potential index as it contains local as well as global information. However, local softness (particularly individual values of s_k^+ and s_k^-) is strongly influenced by the basis set or correlation effects and because of the 'intensive' nature[223] 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 $w(k)$][255] and is believed to be a reliable intra as well as intermolecular local reactivity index. However, $w(\bar{r})$ will not provide any extra information than that by $s(\bar{r})$ or $f(\bar{r})$ as far as intramolecular reactivity is concerned.[259] Even though, individually this descriptor is 'extensive' (*i.e.*, does not go to zero in the thermodynamic limit) in nature, here also 'intensive' nature of $f(\bar{r})$ or $f(k)$ makes philicity [$w(\bar{r})$ or $w(k)$] indices applicable to limited cases[160,259,261,262,311,312] of intermolecular reactivities.

In Conceptual DFT the predominantly charge-controlled reactivity descriptors 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. However, before going into details of the

usefulness of local hardness, $[\eta(\bar{r})]$, as an intermolecular descriptor a brief discussion on the derivations of working equations of $\eta(\bar{r})$ seems to be justified.

From eqn (67) and (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 $F[\rho]$[29] is unknown. It can be done by using the approximated $F[\rho]$. These approximations are based on the Thomas-Fermi-Dirac (TFD)[308–310] approach to DFT. If we keep in mind that the nucleus-electron attraction is not contained in $F[\rho(\bar{r})]$, the following equation is obtained from the general form of the energy functional $E^{\mathrm{TFD}}[\rho(\bar{r})]$,[45] without further approximations:

$$F_E^{TFD}[\rho(\bar{r})] = C_F \int \rho(\bar{r})^{5/3} \mathrm{d}\bar{r} + \frac{1}{2} \int \int \frac{\rho(\bar{r})\rho(\bar{r}')}{|\bar{r} - \bar{r}'|} \mathrm{d}\bar{r}' \mathrm{d}\bar{r} - C_X \int \rho(\bar{r})^{4/3} \mathrm{d}\bar{r} \qquad (93)$$

Here, $C_F = \frac{3}{10}(3\pi^2)^{\frac{2}{3}} = 2.8712$ 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.[45]

Inserting eqn (93) in eqn (67) (where $\lambda[\rho(\bar{r}')]$ is replaced by $\rho(\bar{r}')$) Ghosh *et al.*[231] derived the expression of local hardness as,

$$\tilde{\eta}_D^{TFD}(\bar{r}) = \frac{10}{9N} C_F \rho(\bar{r})^{2/3} - \frac{1}{N} V^{\mathrm{el}}(\bar{r}) - \frac{4}{9N} C_X \rho(\bar{r})^{1/3} \qquad (94)$$

Starting from eqn (93) 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,[230]

$$\tilde{\eta}_D^{TFD'}(\bar{r}) = -\frac{1}{N} V^{\mathrm{el}}(\bar{r}) \qquad (95)$$

here, N is the total number of electrons of the system and $V^{\mathrm{el}}(\bar{r})$ is the electronic part of the molecular electrostatic potential. However, eqn (95) can be derived by approximating just the coulombic contribution (*i.e.*, only the middle term of eqn (93)).[230,237] It was shown that this approximated form of local hardness, (*i.e.*, $-V^{\mathrm{el}}(\bar{r})/N$) can be used as a reliable parameter for comparison of intermolecular reactivity sequences of any particular site in a series of molecules.[114,115,236,247,252,313] There are also some recent studies by Geerlings and his collaborators on the relative contributions of different energetic components to the global and local hardness values.[232,247,314]

In a very recent study, Saha and Roy[252] critically illustrated the limitation[250,251] of $\eta(\bar{r})$ (evaluated from two composite functions *i.e.*, $\lambda[\rho(\bar{r}')] = \rho(\bar{r}')$ eqn (67), (*i.e.*, total local hardness)[114,115,231,236,245,246,248,249,252] and $\lambda[\rho(\bar{r}')] = Nf(\bar{r}')$ eqn (68), *i.e.*, frontier local hardness)[234,237,239,241,243,247,250,251]) when used for comparison of intermolecular reactivity trends between systems of different sizes but having common reactive centers. After a careful analysis they revealed that as the number of electrons increases with the size of the system, the $\frac{1}{N}$ factor alters the expected trends of $\tilde{\eta}_D^{\mathrm{TFD}}(\bar{r})$ or $\tilde{\eta}_D^{TFD'}(\bar{r})$ values. 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 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 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.[315] 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. Thus the modified form of eqn (94) and (95) can be written as

$$\tilde{\eta}_D^{TFD}(\bar{r}) = \frac{10}{9} C_F \rho(\bar{r})^{2/3} - V^{\mathrm{el}}(\bar{r}) - \frac{4}{9} C_X \rho(\bar{r})^{1/3} \tag{96}$$

$$\tilde{\eta}_D^{TFD'}(\bar{r}) = -V^{\mathrm{el}}(\bar{r}) \tag{97}$$

For example they have shown 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 eqn (96) and (97) as,

$$\tilde{\eta}_D^{TFD}(k) = \frac{10}{9} C_F P(k)^{2/3} - V^{\mathrm{el}}(k) - \frac{4}{9} C_X P(k)^{1/3} \tag{98}$$

$$\tilde{\eta}_D^{TFD'}(k) = -V^{\mathrm{el}}(k) \tag{99}$$

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.[236,316] 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. Because electrostatic force operates from large distance any hardness based (or charge-controlled) reactivity descriptors will be more suitable for explaining the reactivity at this stage (*i.e.*, intermolecular reactivity sequence).[230] 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 $s(k)$],[143] Fukui function index [$f(\bar{r})$ or $f(k)$],[143,179,181,307] philicity [$w(k)$][255]) will be more suitable to describe the intramolecular reactivity. The findings by Klopman[306] also support this argument.

B On the way of predicting the regioselectivity of large molecular systems within the framework of conceptual DFT

The bottleneck in predicting the intramolecular reactivity trends of large chemical and biological systems lie in the fact that the calculation for the whole system is needed to be performed. This is because of the incorporation of the global softness part in the expression of local softness (*i.e.*, $s^{\alpha}(\bar{r}) = f^{\alpha}(\bar{r})S$ where $\alpha = +, -$ and 0).[143] Thus, larger and larger the system becomes, lower and lower level of calculation we have to opt for. Although the regioselectivity plays an important role to understand a reaction, because of 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 molecular fragmentation approach. It is worthwhile to point out here that for predicting the regioselectivity of large molecular systems in the framework of Conceptual DFT, a simple fragmentation approach was first proposed by Saha and Roy.[114,115] They proposed a model, named as 'One-into-Many' model, 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). As argued in the previous sub-section [3 A], local hardness is mainly a charge-controlled descriptor when evaluated using eqn (96) and (97) (or eqn (98) and (99)). Since it can take care of the long-range reactivity (*i.e.*, intermolecular reactivity), 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. Thus, while the technique to be adopted is similar to the divide-and-conquer (DC) approach formulated by Yang,[86] 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. Because, 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)[308–310] approach of density functional theory) from the distant atoms or moieties are less significant and thus can be neglected. However, contribution from close 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, 'buffer zone' refers to the moiety of the chemical system, which is common to two adjacent reactive sites (local hardness values of which are to be evaluated).

To implement the 'One-into-Many' model,[114,115] Saha and Roy has chosen right-handed B-DNA (PDB ID: 1BNA)[317] as a model system. The structure of the DNA is shown in Fig. 1. In this model system there are 12 base-pairs with the sequence d(CpGpCpGpApApTpTpCpGpCpG).[317] Detail literature study on adduct formation indicates that the majority of known carcinogens react with DNA through N^2 and N7 positions of guanine.[318–322] 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)[323,324] and the exocyclic oxygen of thymine (O^2)[324,325] residues are the reactive sites for electrophilic attack. Saha and Roy calculated $\tilde{\eta}_D^{\mathrm{TFD}}(k)$ (eqn (98)) and $\tilde{\eta}_D^{TFD'}(k)$ (eqn (99)) values of those 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 (Fig. 2). The geometries associated with these two base-pairs are taken from 1BNA[317] without any modification. However, to generate more reliable data a buffer zone around each reactive site is required. So, three base-pairs were chosen at a time (named as '*Triple-Base-Pair Systems*') and evaluated the $\tilde{\eta}_D^{\mathrm{TFD}}(k)$ and also $\tilde{\eta}_D^{TFD'}(k)$ values of the reactive sites in the middle base-pair. Here the base-pairs, which are on either side of the central base-pair, create the buffer zone *i.e.*, try to mimic the environment of the 1BNA.[317] Although, in reality, some contribution from base-pairs which are next to the adjacent base pairs are also expected, for computational limitation the calculation has been restricted to the '*Triple-Base-Pair Systems*' only. All possible combinations of '*Triple-Base-Pair Systems*' (Fig. 3) are also taken from 1BNA[317] and the geometrical parameters are generated[326] without any geometrical changes in the study. Relevant calculations have been performed using Gaussian suite of programs.[327]

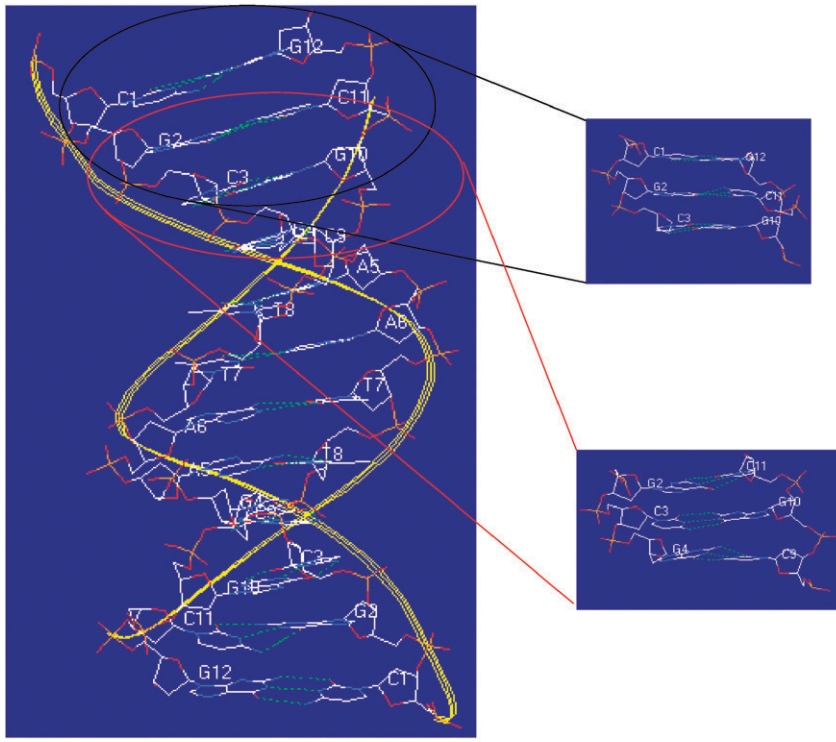

Fig. 1 Fragmentation of Watson-Crick double-stranded B-DNA (PDB ID: 1BNA)[317] into *Triple-Base-Pair Systems* (Reprinted with permission from ref. 114. Copyright 2010 American Chemical Society).

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 equal in all the base-pairs (whether Single or Triple). So, for DNA one can directly apply the modified form of eqn (96) and (97), *i.e.*, eqn (98) and (98).[114,115]

Saha and Roy[114,115] concluded 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. In the case of exocyclic nitrogen, $\tilde{\eta}_D^{\mathrm{TFD}}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$ values of N^2 position of G is found to be the highest among all the exocyclic nitrogen's present in DNA, which agrees with the previous experimental results.[318,320,321] While comparing the N7 position for *Single-Base-Pair Systems* and *Triple-Base-Pair Systems*, highest $\tilde{\eta}_D^{\mathrm{TFD}}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$ values belong to that of G.[115] Only for exocyclic O-atom in DNA base-pairs, the method proposed by Saha and Roy[114,115] fails to generate expected trends of hardness values ($\tilde{\eta}_D^{\mathrm{TFD}}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$). They have reported that (both from $\tilde{\eta}_D^{\mathrm{TFD}}(k)$ and $\tilde{\eta}_D^{TFD'}(k)$ values) O^2 position of C is the most reactive sites even though Singer[324] has observed that O^6 of G and O^2 of T are the two significant reactive sites towards electrophilic attack in the double helical DNA. This failure of Saha and Roy[114,115] is attributed to their 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*. It is worth mentioning here that Fan *et al.*, in one interesting study,[328] suggested a general guideline for the computation of large

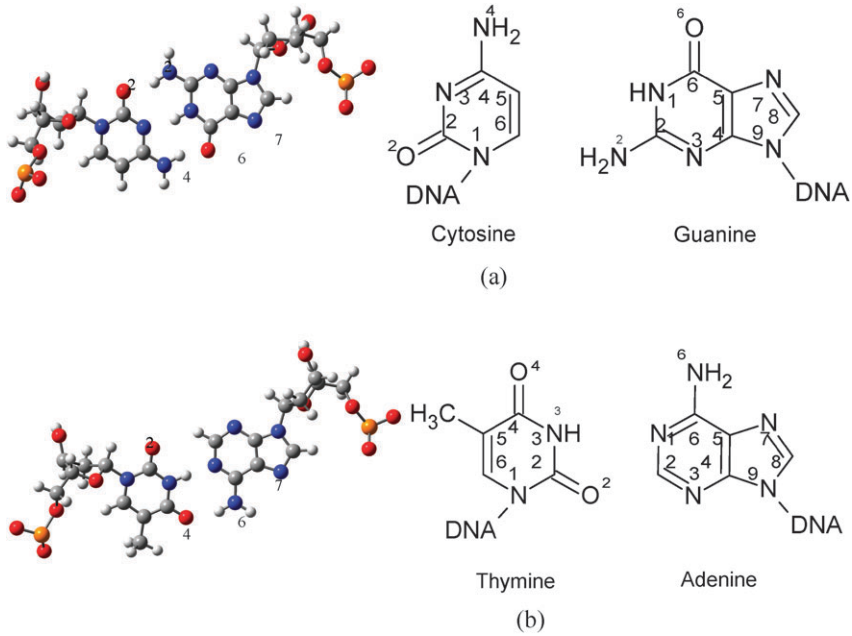

Fig. 2 Two different *Single-Base-Pair Systems* (a) Cytosine-Guanine and (b) Thymine-Adenine (Reprinted with permission from ref. 114. Copyright 2010 American Chemical Society).

biological systems with considerably high accuracy and low computational expense. By using DFT based reactivity indices (namely, chemical potential, hardness, softness and electrophilicity index) of the nucleic acids base pairs they concluded that (i) the A-T base pair has a larger hardness (η) than the G-C pair, indicating that in gas phase the former should be more stable than the latter, and (ii) the G-C base pair possesses a bigger electrophilicity index (w), indicating that it has better capability to accept electrons than the other pair, in agreement with the experimental finding.[328] However, in their study they confined into the *Single-Base-Pair Systems* only.

In a very recent study, Barrientos-Salcedo *et al.*,[329] successfully reproduced experimental findings 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)[330] 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)[330] was taken from the Protein Data Bank. The chemical structures of the three fragments studied in this work are shown in Fig. 4, while the convention of atom-numbering for heavy atoms in this study is shown in Fig. 5. They revealed[329] 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),

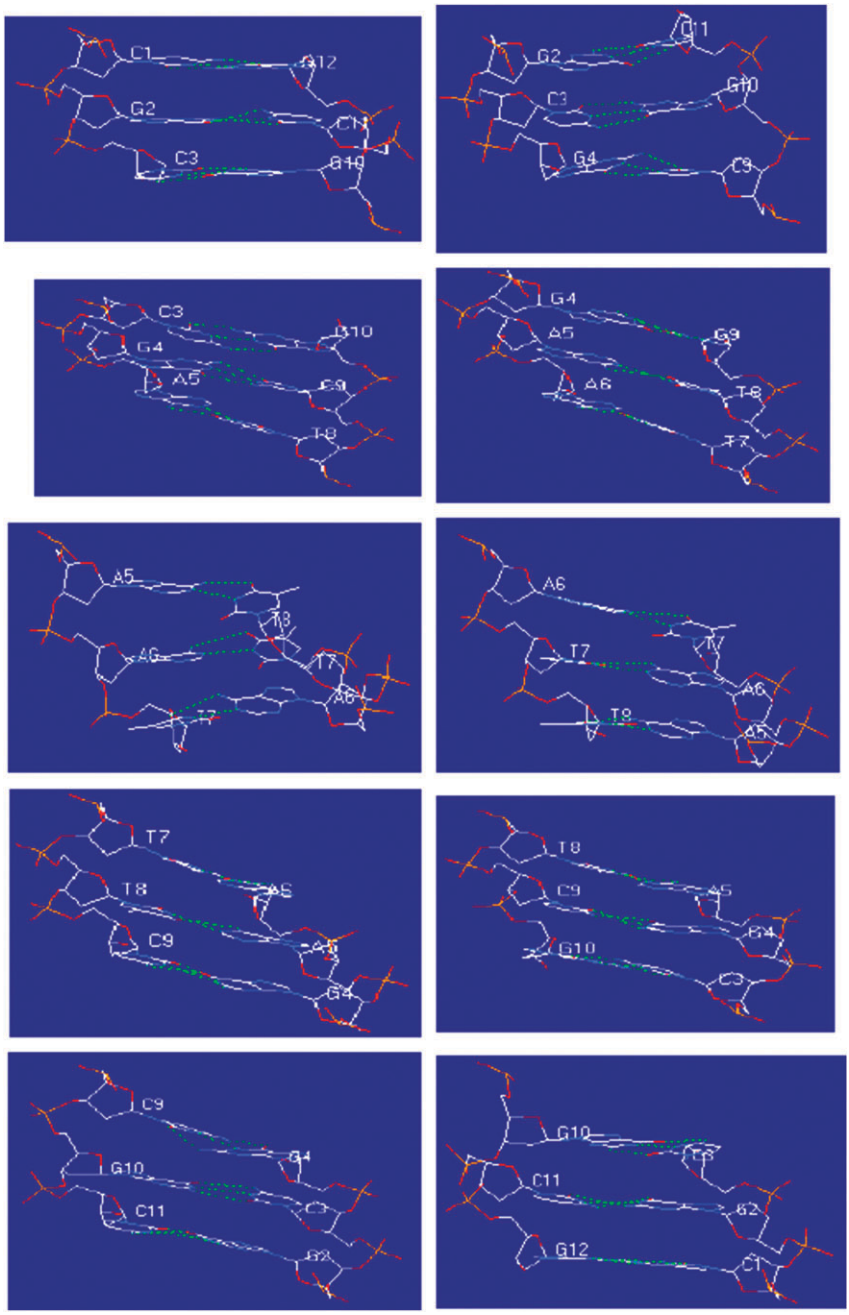

Fig. 3 Ten different *Triple-Base-Pair Systems* of B-DNA (PDB ID: 1BNA)[317] (Reprinted with permission from ref. 114. Copyright 2010 American Chemical Society).

respectively, they observed[329] that these have larger negative charges as compared with 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.*,[331]

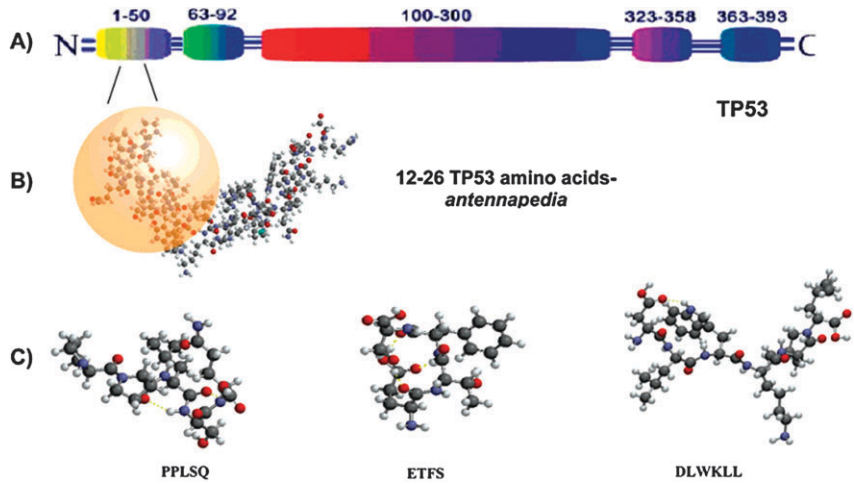

Fig. 4 Images showing (A) TP53 protein, (B) amino acids 12–26 of TP53 protein (shaded circle) and penetratin and (C) fragments analyzed in ref. 329 (ball and stick model) (Reprinted with permission from ref. 329. Copyright 2010 American Chemical Society).

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)(eqn (34)),[157] and the spatial extent measured though the $\langle R^2 \rangle$, Barrientos-Salcedo *et al.* observed[329] that the ETFS fragment exhibits a larger value for ionization potential, which might be related with 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. Moreover from the hydrogen atom charges, they concluded that these atoms might be forming hydrogen bonds (*i.e.*, intramolecular interactions), which might contribute in increasing the stability of these (*i.e.*, PPLSQ and DLWKLL) peptides structures. Furthermore, results of the frontier molecular orbitals (HOMO–LUMO) and electrostatic potential surfaces revealed[329] the possible reactive sites of the amino acid fragment.

Until now, attention was focused only on the fragment-based approach. On the basis of an energy perturbation method, Li and Evans[332,333] presented a slightly different formulation, indicating that, for a hard reaction, the site of minimal Fukui function[143,179,181,307,334,335] 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).[333] PEP is involved in a number of important enzymatic reactions, *i.e.*, 5-enolpyru-vylshikimate 3-phosphate (EPSP) synthase, an enzyme of the shikimate pathway in plants; 3-deoxy-D-manno-2-octu-losonate-8-phosphate (KDO8P) synthase, an enzyme of the glycolysis; 2-dehydro-3-deoxyphosphoheptonate aldolase (DAHP, also referred to as DHAP).[336–342] During the reaction catalyzed by EPSP synthase[339,340] (Fig. 6, path I), C3 of PEP is protonated during initiation of the reaction, which implies, that this carbon is a hard base (because a proton is believed to be a hard acid). In contrast to the EPSP synthase-catalyzed reaction, in the KD08P synthase-catalyzed reaction[341,342] (Fig. 6, path II), this same carbon acts as a soft base and is used to

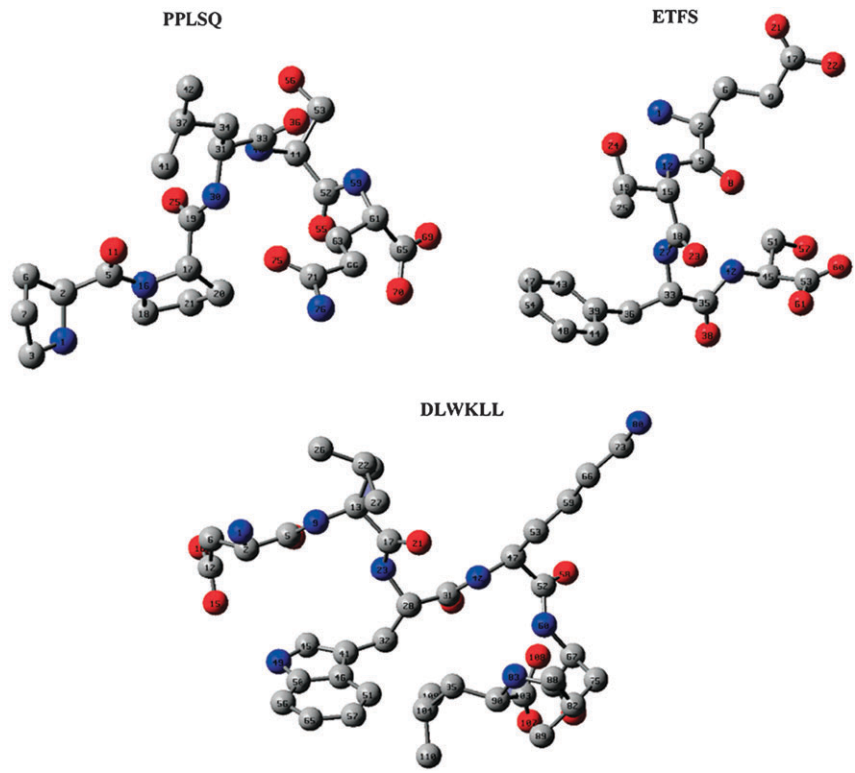

Fig. 5 Convention used in numbering heavy atoms of atomic charges in ref. 329 (ball and stick model) (Reprinted with permission from ref. 329. Copyright 2010 American Chemical Society).

attack an aldehyde carbon, which is believed to be a soft base, of the cosubstrate of the enzyme. As in the KD08P synthase catalyzed reaction, a soft base at C3 is needed for the reaction catalyzed by DAHP synthase to attack the carbonyl carbon[336–338] (Fig. 6, path III). Conceptual DFT leads to hypothesis that this dual nature of C3 of PEP depends on the ionization state of PEP and on the conformation 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[307]) 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.*,[343] where electronegativities and hardnesses were determined for artificially constructed amino acid groups, in both the α-helix and β-sheet conformations. Group hardness[344–351] (by using $\eta = \frac{IP-EA}{2}$) was found to play the dominant role, whereas group electronegativity (by using eqn (10)) only had a minor influence on the sequence.[343]

As an example of how the Conceptual DFT have predicted reactivity of ligands for the NCp7 Zn fingers, Rice and co-workers studied,[158,352] *via* $f^{-}(\bar{r})$ (eqn (43))[143,179,181,307] and $s^{-}(\bar{r})$ (eqn (48)),[143] the regional reactivity of the two retroviral zinc fingers of the HIV-1 nucleocapsid p7 (NCp7) protein, representing antiviral targets. Regional reactivity is displayed spectrally on the solvent-accessible surface of the Zn fingers (Fig. 7). The reactivity spectrum corresponds to the Fukui function which probes the regions of the Zn fingers most able to donate electron density and thus participate in

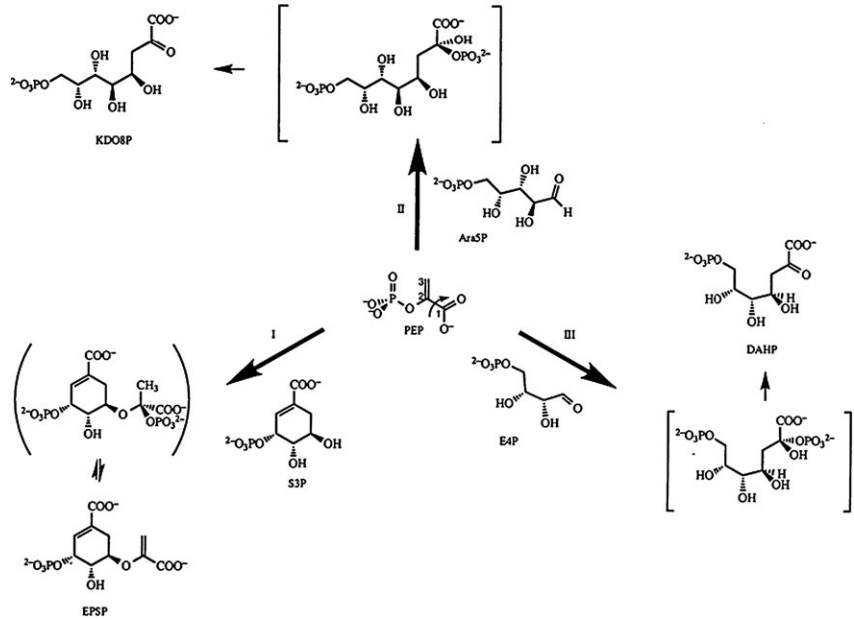

Fig. 6 Schematic representation of three enzymatic reactions that use the enolpyruvyl moiety of PEP. The arrow in the structure of PEP indicates the rotation of the carboxylate group of the molecule around the single bond. Structures in parentheses indicate intermediates observed experimentally; structures in brackets indicate intermediates proposed here that are not yet observed. Path I is the 5-enolpyruvyl-shikimate-3-phosphate (EPSP) synthase catalyzed reaction; path II is that 3-deoxy-D-manno-2-octulosonate-8-phosphate (KD08P) synthase catalyzed reaction; and path III is the 3-deoxy-D-arabinoheptulosonate-7-phosphate (DAHP) synthase catalyzed reaction (Reproduced from ref. 333 with permission. Copyright 2010 National Academy of Sciences, U.S.A.).

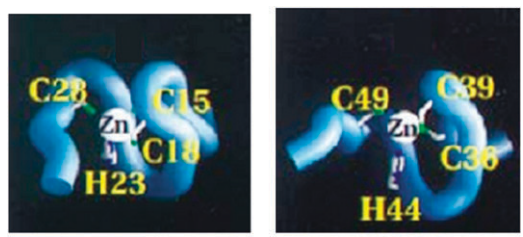

Fig. 7 NCp7-ligand docking arrangements. Ligand atom coloring: (A) Finger 1 (B) Finger 2 orientations (Reproduced from ref. 158 with permission. Copyright 2010 National Academy of Sciences, U.S.A.).

covalent bond formation. The regions of both Zn fingers prove that the Cys thiolates dominate the reactivity profile of NCp7. The reactive sites of finger 2 form a more contiguous reactive surface in comparison to finger 1, where they appear more isolated. 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.*[353] In the proteolysis of peptides by serine proteases, a nucleophilic attack by the hydroxyl oxygen (O_γ) of Ser221 at the carbon atom of the scissile peptide bond occurs when simultaneously the hydroxyl proton of the catalytic Ser221 is transferred to the N_{ε_2} atom of His64 [Fig. 8, X-ray Structure of subtilisin BPN (PDB code: 2ST1)[354–357]]. As a possible reactivity index for the serine nucleophilicity both local softness[143] and local hardness[230–232] were looked at.[353] They found that local softness (eqn (48)) is not suitable for describing the nucleophilicity of Ser221. Local hardness, approximated by the minimum of the MEP, $V(\bar{R})_{\min}$, and the atomic charge $q(k)$ for the serine hydroxyl oxygen (O_γ), however performs very well. The basicity of the histidine was studied using the charge (q) on the basic nitrogen atom (*i.e.*, N_{ε_2}) and the "protonation energy" ΔE (the difference between the energy of the model system with the histidine protonated on the N_{ε_2} atom and the energy of the model system having an unprotonated histidine). The diminished basicity of histidine in the aspartate mutants also was reflected in decreased q of $N_{\varepsilon_2}$and ΔE values.

Mignon *et al.*[358] also used local hardness[230–232] 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)[359] and RNase T1 (EC 3.1.27.3).[360,361] They have shown that the negative charge build up on the

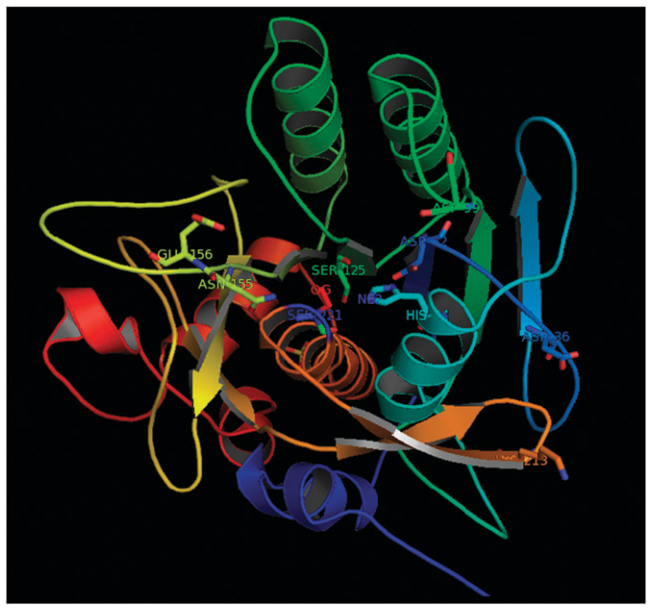

Fig. 8 Schematic drawing of subtilisin, side chains of the residues of importance in ref. 353 are shown (X-ray Structure of subtilisin BPN (PDB code: 2ST1))[354] (Reprinted from ref. 353. Copyright 2010, with permission from Elsevier Publications.).

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[143] of the involved atomic centers),[344,346,349,362] 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.[363,364] Different atomic centers of lumiflavine and nicotinamide were tested, but the smallest difference in group softness was found between the C3, Ht, and C5 atoms of nicotinamide and the C4a, N1, and N5 atoms of lumiflavine, supporting the hydride transfer[365–371] between these regions (Fig. 9)[363] In the lumiflavine molecule, the local electrophilicity[255] (*i.e.*, $w^+(k)$; eqn (75)) of the N5 atom is higher than the local electrophilicity of the N1 and C4a atoms. So, 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 the series of subsequent papers, Roos *et al.*[364,372–375] scrutinizes the experimental findings of the enzymatic reaction mechanism of *Staphylococcus aureus* arsenate reductase (ArsC) within the Conceptual DFT framework. The ArsC gene product from *Staphylococcus aureus* pI258 plasmid, has a low-molecular-weight phosphatase (LMW PTPase) anion-binding motif[376] known as the P-loop. The amino acid sequence of this ligand-binding loop in pI258 ArsC is Cys10–Thr11–Gly12–Asn13–Ser14–Cys15–Arg16–Ser17.[376] The first step in the multistep catalytic mechanism of ArsC consists of a nucleophilic displacement reaction carried out by Cys10 on arsenate, by which a covalent Cys10–arseno adduct is formed (Fig. 10, step 1).[377] In the second step, another nucleophile, Cys82, attacks the covalent Cys10–arseno adduct, thereby leading to the release of arsenite and the formation of a Cys10–Cys82 disulfide bridge (Fig. 10, step 2).[376,378–380] After the second reaction step (Fig. 10, step 2), when the Cys10–Cys82 disulfide intermediate has been formed, the conformation of the redox helix is changed into a transitional conformation

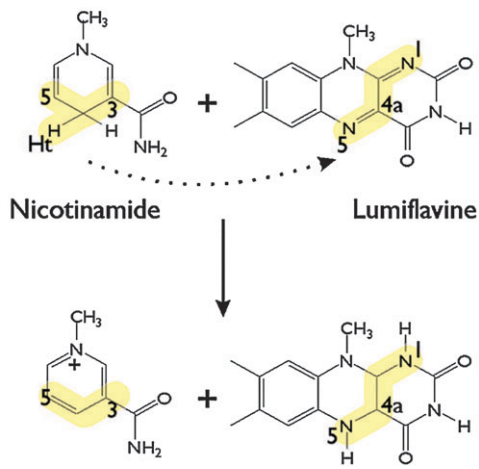

Fig. 9 Hydride transfer reaction between lumiflavine and 1-methylnicotinamide (NH) (Reprinted from ref. 363. Copyright 2010, with permission from Elsevier Publications. Reprinted with permission from ref. 364. Copyright 2010 American Chemical Society).

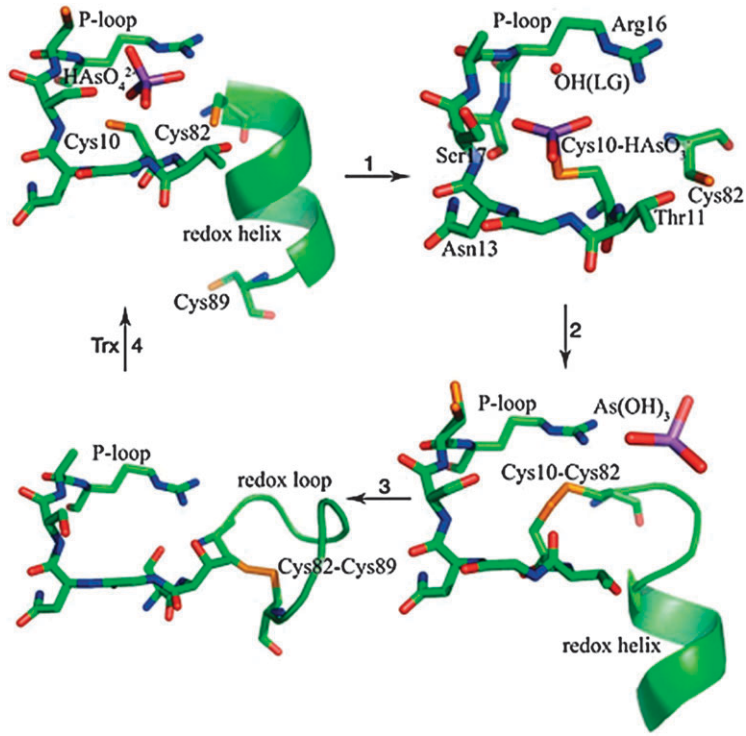

Fig. 10 Scheme of the reaction mechanism of PI258 ArC. (1) The reaction starts with the nucleophilic attack of Cys10 on arsenate leading to a covalent enzyme-arseno intermediate. (2) Arsenite is released after the nucleophilic attack of the thiol of Cys82. A Cys10–Cys82 intermediate is formed, and the redox helix partially unfolds. (3) At the end of the reduction cycle, Cys89 attacks Cys82, forming a Cys82–Cys89 disulde. The redox helix is looped out and presents the disulde bridge at the surface of the enzyme to thioredoxin. (4) Thioredoxin (Trx) regenerates the reduced form of arsenate reductase for a subsequent catalytic cycle. (Reproduced from ref. 386 with permission. Copyright Wiley-VCH Verlag GmbH & Co. KGaA. And reprinted with permission from ref. 364. Copyright 2010 American Chemical Society).

between a helix and a loop (Fig. 10).[379,380] The subsequent third reaction step (Fig. 10, step 3) consists of the nucleophilic attack of Cys89 on the Cys10–Cys82 disulfide, a process resulting in the formation of the Cys82–Cys89 disulfide and the reduction of Cys10.[381] Thioredoxin (Trx) converts oxidized ArsC back to its initial reduced state[382] (Fig. 10, step 4) through a subsequent catalytic cycle (Fig. 11). Thioredoxins are proteins that act as antioxidants by facilitating the reduction of other proteins by cysteine thiol-disulfide exchange.[382] All thioredoxins have a similar three-dimensional fold comprising a central core of four β-strands surrounded by three α-helices.[383] They also feature a conserved active-site loop containing two redoxactive cysteine residues in the sequence Trp–Cys–Gly–Pro–Cys[384] numbered as Trp28 to Cys32 in both *Bacillus subtilis* (Bs_Trx) and *S. aureus* (Sa_Trx) Trx.[364,372–375] $Cys29^{Trx}$ nucleophilically attacks $Cys89^{ArsC}$ of the $Cys82^{ArsC}$–$Cys89^{ArsC}$ disulfide, leading to the reduction of $Cys82^{ArsC}$ and the formation of the Trx–ArsC mixed disulfide intermediate complex between $Cys29^{Trx}$ and $Cys89^{ArsC}$ (Fig. 10 and 11A).[382,385] In this complex, $Cys32^{Trx}$ performs a nucleophilic attack on $Cys29^{Trx}$ of the $Cys29^{Trx}$–$Cys89^{ArsC}$ disulfide

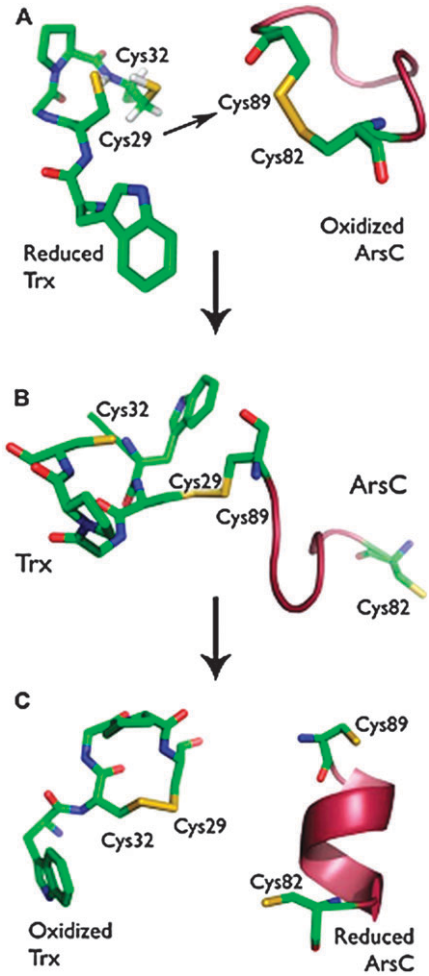

Fig. 11 Bs_Trx reduces Bs_ArsC *via* an intermediate Trx–ArsC complex. (**A**) $Cys29^{Trx}$ of reduced Bs_Trx nucleophilically attacks $Cys89^{ArsC}$ of the $Cys82^{ArsC}$–$Cys89^{ArsC}$ disulfide of oxidized Bs_ArsC, leading to the formation of the mixed $Cys29^{Trx}$–$Cys89^{ArsC}$ disulfide. (**B** and **C**) $Cys32^{Trx}$ performs a nucleophilic attack on $Cys29^{Trx}$, leading to the release of reduced Bs_ArsC and oxidized Bs_Trx (Reproduced from ref. 375).

(Fig. 11B). Accordingly, the Trx–ArsC complex dissociates, releasing reduced ArsC and oxidized Trx (Fig. 11C).

To implement the Conceptual DFT for describing pI258 ArsC enzyme, one needs an adequate model, combining accuracy with computational tractability. Ross *et al.* in their studies applied an interesting modeling technique.[372–375] 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).[379] 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

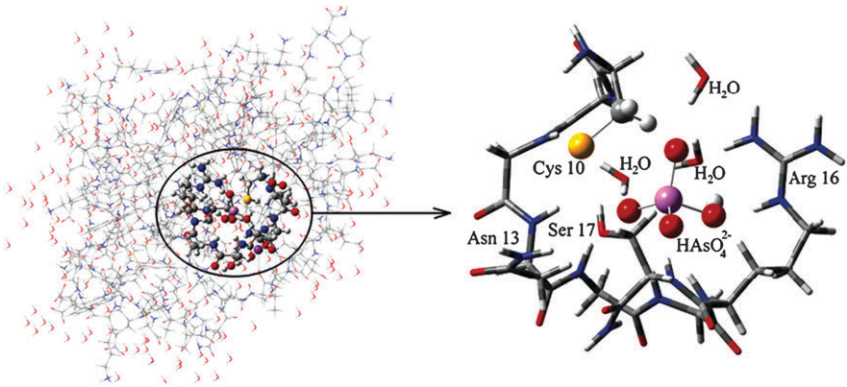

Fig. 12 Reduction of the X-ray structure of ArsC (PDB: 1LJU)[379] to the WT model. Partitioning of the WT model system of ArsC into two layers: high level represented in "Ball & Stick"; low level in "Tube". A similar division is made for the Asn13Ala and the Arg16Ala mutants. Color code: hydrogen, white; nitrogen, blue; carbon, gray; oxygen, red; sulfur, yellow; arsenic, purple (Reprinted with permission from ref. 373. Copyright 2010 American Chemical Society).

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)" [372–375] (Fig. 12). The Ser17Ala Arg16Ala and the Asn13Ala mutants were built "in silico",[158] 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).[379] 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).[379] In both structures, hydrogen atoms were placed and optimized.

The Conceptual DFT is used to assess the nucleophilic attack of Cys10 on arsenate (Fig. 10 step 1). The difference in local softness (*i.e.*, $\Delta s(k) = |s^-(k) - s^+(k)|$) between the attacking nucleophilic sulfur atom of Cys10 and the receiving electrophilic arsenic atom of arsenate is minimal when dianionic arsenate was considered.[372] In addition, calculation of the binding energy showed that the binding of dianionic arsenate in ArsC turned out to be more favorable than that of monoanionic arsenate.[373] 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.[364,386] Conceptual DFT analysis indicates Ser17 to be the major activator of the electrophilic Cys10–arseno adduct. Calculation of the nucleofugality (*i.e.*, $\Delta E_{\text{nucleofuge}} = \frac{(IP-3EA)^2}{8(IP-EA)}$)[167–169] indicates that the enzyme increases the leaving-group capacity of OH^- (first reaction step) and of $HAsO_3^{2-}$ (second reaction step). Further, on the basis of the correlation between the natural population analysis (NPA) charge on the sulfur of the thiolate compound and the pKa of Cys82 and Cys89, Roos *et al.* were able to explain why the presence of arsenate in the active site of ArsC brings Arg16 within hydrogen bonding distance to Cys82, leading to an additional pKa decrease of 0.95 pKa.[374] Thus, the substrate itself contributes to the nucleophilic

character of Cys82. Prior to the third reaction step, Cys89 is kept in the nonactive high pKa form by the presence of the Cys82–Cys89 redox helix. This helix partially unfolds when the Cys10–Cys82 disulfide is formed, thereby favouring the thiolate form of Cys89 and enabling the third reaction step.[386]

In a most recent study Roos *et al.*[364,375] provided fresh insight into the mechanism behind the dissociation of the mixed disulfide complexes between thioredoxin (Trx)[382] and its substrates. As a key model, the complex between Trx and its endogenous substrate, arsenate reductase (ArsC), was used (Fig. 11).[375] In this structure, a Cys29Trx–Cys89ArsC intermediate disulfide is formed by the nucleophilic attack of Cys29Trx on the exposed Cys82ArsC–Cys89ArsC in oxidized ArsC. 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.*[364,375] Information regarding the selectivity of the nucleophilic attack was obtained from a DFT based reactivity analysis.[375] In the Trx–ArsC complex, four possible reactions between the attacking nucleophilic cysteines (Cys32Trx and Cys82Trx) and the accepting electrophilic disulde (Cys29Trx–Cys89ArsC) can be considered. The minimal local softness[143] difference (*i.e.*, $\Delta s(k) = |s^{-}(k) - s^{+}(k)|$) of the interacting sulfur atoms favors the nucleophilic attack of Cys32Trx on Cys29Trx.[375] By studying the s_k^+/s_k^-[228] of the sulfur atoms of the nucleophilic Cysteines in the Bs_Trx–ArsC complex they found that the Cys29Trx –Cys89ArsC disulde is less soft than Cys32Trx and Cys82ArsC, and Cys32Trx is softer than Cys82ArsC.[375] The high reactivity of Cys32Trx toward the Cys29Trx–Cys89ArsC disulde 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.[375]

4. Conclusion

The study of regioselectivity remains a very important topic in the chemical literature. Thousand of papers have appeared in this topic, 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 a description generally in full agreement with a lot of previous work 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). Although the world has witnessed a 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.

Motivated by its potential importance and guided by the insights (as described in section 2) of the present report, the authors have given one simple fragmentation (One-into-Many model[114,115]) 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. The reactivity descriptor, used in this model as a key tool, is local hardness [$\eta(\bar{r})$] because it's predominant component is electronic contribution to the molecular electrostatic potential (MEP). MEP has a long distance 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 it's N-dependence problem.[252] In case of studying the regioselectivity (using One-into-Many model) of DNA systems[114,115] 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 (may be nature has created base pairs like this!!!). To, solve this N-dependence problem the present authors have suggested to consider only those moieties in different systems for which the number of electrons is same (*e.g.*, C═O moiety, when intermolecular reactivity of carbonyl compounds are studied).[252] But, this is not a general solution, applicable to all kinds of systems. So, 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. The present authors are working on it and hope to find such a descriptor soon.

Acknowledgements

The authors are extremely grateful to B. M. Deb, D. Mukherjee, S. Pal, K. Hirao, P. Geerlings and P. W. Ayers. This goal would not have been attainable without the intellectual support and the fruitful exchange of ideas of numerous teachers, colleagues, coworkers, theoreticians as well as experimentalists, from the BITS-Pilani, India, as well as from numerous other countries. R.K.R. wish to express his deepest sense of gratitude and sincere thanks to them: P. Bultinck, P. Bagaria, A. K. Chandra, K. R. S. Chandrakumar, P. K. Chattaraj, R. Ch. Deka, M. Durga Prasad, S. R. Gadre, S. K. Ghosh, M. K. Harbola, E. D. Jemmis, S. Krishnamurty, A. Kumar, D. Kumar, W. Langenaeker, B. K. Patel, J. Paulovič, F. de Proft, P. Purkayastha, P. U. Manohar, G. Narahari Sastry, N. Tajima, V. Usha and N. Vaval. S.S. acknowledges R. Goswami, Md. S. Alam and T. K. Roy for their helping hand. R.K.R. is greatly indebted to Prof. G. Webb for requesting him to write this article. Finally, the senior author gratefully acknowledges research support to DST (Project Ref. No. SR/S1/PC-41/2008) and Departmental Research Support under University Grants Commission (UGC) Special Assistance Programme (SAP) [Project Ref. No. F. 540/14/DRS/2007 (SAP-I)], Government of India. S. S. thanks CSIR, New Delhi, India for granting him Senior Research Fellowships.

References

1 A. Hassner, *J. Org. Chem.*, 1968, **33**, 2684.
2 J. March, *Advanced Organic Chemistry: Reactions, Mechanisms, and Structure*, Wiley and Sons, NY, 4th edn, 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 *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*, ed. R. Maksic, Springer, Berlin, 1990, p. 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*, Academic Press, New York, 2nd edn, 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, p. 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 S. Saha and R. K. Roy, *J. Phys. Chem. B*, 2007, **111**, 9664.
115 S. Saha and R. K. Roy, *J. Phys. Chem. B*, 2008, **112**, 1884.
116 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.
117 E. J. Baerends and P. Ros, *Int. J. Quantum Chem., Quantum Chem. Symp.*, 1978, **12**, 169.
118 B. I. Dunlap, J. W. D. Connolly and J. R. Sabin, *J. Chem. Phys.*, 1979, **71**, 3396.
119 P. M. Boerrigter, G. Velde and E. J. Baerends, *Int. J. Quantum Chem.*, 1988, **33**, 87.
120 J. Andzelm and E. Wimmer, *J. Chem. Phys.*, 1992, **96**, 1280.
121 L. Pauling, *J. Am. Chem. Soc.*, 1932, **54**, 3570.
122 L. Pauling, *The Nature of the Chemical Bond*, Cornell University Press, Ithaca, NY, 1960.
123 R. P. Iczkowski and J. L. Margrave, *J. Am. Chem. Soc.*, 1961, **83**, 3547.
124 R. S. Mulliken, *J. Chem. Phys.*, 1934, **2**, 782.
125 E. P. Gyftopoulos and G. N. Hatsopoulos, *Proc. Natl. Acad. Sci. U.S.A.*, 1968, **60**, 786.
126 T. A. Koopmans, *Physica*, 1933, **1**, 104.
127 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.
128 B. M. Deb, *J. Am. Chem. Soc.*, 1974, **96**, 2030.
129 B. M. Deb, P. N. Sen and S. K. Bose, *J. Am. Chem. Soc.*, 1974, **96**, 2044.
130 B. M. Deb, *J. Chem. Educ.*, 1975, **52**, 314.
131 R. G. Pearson, *Proc. Natl. Acad. Sci. U. S. A.*, 1986, **83**, 8440.
132 R. K. Roy and S. Pal, *J. Phys. Chem.*, 1995, **99**, 17822.
133 R. G. Pearson, *J. Chem. Educ.*, 1999, **76**, 267.
134 J. Simons and K. D. Jordan, *Chem. Rev.*, 1987, **87**, 535.
135 A. Z. Szarka, L. A. Curtiss and J. R. Miller, *J. Chem. Phys.*, 1999, **246**, 147.
136 A. Dreuw and L. S. Cederbaum, *Chem. Rev.*, 2002, **102**, 181.
137 R. T. Sanderson, *Science*, 1955, **121**, 207.
138 R. T. Sanderson, *Polar Covalence*, Academic Press, New York, 1983.
139 R. G. Pearson, *Coord. Chem. Rev.*, 1990, **100**, 403.
140 R. G. Pearson, *Chemical Hardness*, Wiley, New York, 1997.
141 S. Pal and K. R. S. Chandrakumar, *J. Am. Chem. Soc.*, 2000, **122**, 4145.
142 K. R. S. Chandrakumar and S. Pal, *J. Phys. Chem. A*, 2002, **106**, 5737.
143 W. Yang and R. G. Parr, *Proc. Natl. Acad. Sci. U.S.A.*, 1985, **82**, 6723.
144 The Factor of 2 is arbitrary, to create a symmetry between Mulliken's electronegativity and chemical hardness.
145 F. De Proft and P. Geerlings, *Chem. Rev.*, 2001, **101**, 1451.
146 K. L. Sebastian, *Chem. Phys. Lett.*, 1994, **231**, 40.
147 P. K. Chattaraj, G. H. Liu and R. G. Parr, *Chem. Phys. Lett.*, 1995, **237**, 171.
148 K. L. Sebastian, *Chem. Phys. Lett.*, 1995, **236**, 621.
149 P. K. Chattaraj, A. Cedillo and R. G. Parr, *Chem. Phys.*, 1996, **204**, 429.
150 P. Politzer, *J. Chem. Phys.*, 1987, **86**, 1072.
151 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.
152 T. K. Ghanty and S. K. Ghosh, *J. Phys. Chem.*, 1993, **97**, 4951.
153 S. Hati and D. Datta, *J. Phys. Chem.*, 1994, **98**, 10451.
154 R. K. Roy, A. K. Chandra and S. Pal, *J. Phys. Chem.*, 1994, **98**, 10447.
155 Y. Simon-Manso and P. Fuentealba, *J. Phys. Chem. A*, 1998, **102**, 2029.
156 R. F. Nalewajski, *J. Am. Chem. Soc.*, 1984, **106**, 944.
157 R. G. Parr, L. von Szentpaly and S. Liu, *J. Am. Chem. Soc.*, 1999, **121**, 1922.
158 A. T. Maynard, M. Huang, W. G. Rice and D. G. Covell, *Proc. Natl. Acad. Sci. U.S.A.*, 1998, **95**, 11578.

159 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.
160 P. Bagaria and R. K. Roy, *J. Phys. Chem. A*, 2008, **112**, 97.
161 P. Bagaria, S. Saha, S. Murru, V. Kavala, B. Patel and R. Roy, *Phys. Chem. Chem. Phys.*, 2009, **11**, 8306.
162 P. W. Ayers and R. G. Parr, *J. Am. Chem. Soc.*, 2001, **123**, 2007.
163 P. Jaramillo, P. Fuentealba and P. Pérez, *Chem. Phys. Lett.*, 2006, **427**, 421.
164 P. Jaramillo, P. Perez, R. Contreras, W. Tiznado and P. Fuentealba, *J. Phys. Chem. A*, 2006, **110**, 8181.
165 A. Cedillo, R. Contreras, M. Galvan, A. Aizman, J. Andres and V. S. Safont, *J. Phys. Chem. A*, 2007, **111**, 2442.
166 F. D. Vleeschouwer, V. V. Speybroeck, M. Waroquier, P. Geerlings and F. D. Proft, *Org. Lett.*, 2007, **9**, 2721.
167 P. W. Ayers, J. S. M. Anderson and L. J. Bartolotti, *Int. J. Quantum Chem.*, 2005, **101**, 520.
168 P. W. Ayers, J. S. M. Anderson, J. I. Rodriguez and Z. Jawed, *Phys. Chem. Chem. Phys.*, 2005, **7**, 1918.
169 P. R. Campodónico, C. Pérez, M. Aliaga, M. Gazitúa and R. Contreras, *Chem. Phys. Lett.*, 2007, **447**, 375.
170 S. Liu, T. Li and P. W. Ayers, *J. Chem. Phys.*, 2009, **131**, 114106.
171 P. W. Ayers, S. Liu and T. Li, *Chem. Phys. Lett.*, 2009, **480**, 318.
172 The quantity ρ is of course of much interest, being directly accessible experimentally and readily visualizable—just the classical density of the electronic system. The properties of ρ should be mentioned: see ref. 31,173–178.
173 R. F. W. Bader, *Acc. Chem. Res.*, 1975, **8**, 34.
174 V. H. Smith Jr and I. Absar, *Israel J. Chem.*, 1977, **16**, 87.
175 R. Bader, Y. Tal, S. Anderson and T. Nguyen-Dang, *Israel J. Chem*, 1980, **19**, 8.
176 R. F. W. Bader and C. Chang, *J. Phys. Chem.*, 1989, **93**, 2946.
177 R. F. W. Bader, *Atoms in Molecules: A Quantum Theory*, 1990.
178 P. Nasertayoob and S. Shahbazian, *J. Mol. Struct. (THEOCHEM)*, 2008, **869**, 53.
179 R. G. Parr and W. Yang, *J. Am. Chem. Soc.*, 1984, **106**, 4049.
180 An explicit expression for the Fukui function in terms of K-S orbital vide ref. 307.
181 P. W. Ayers and M. Levy, *Theor. Chem. Acc.*, 2000, **103**, 353.
182 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.
183 W. T. Yang, Y. K. Zhang and P. W. Ayers, *Phys. Rev. Lett.*, 2000, **84**, 5172.
184 P. W. Ayers, *J. Math. Chem.*, 2008, **43**, 285.
185 J. P. Perdew, R. G. Parr, M. Levy and J. L. Balduz, Jr., *Phys. Rev. Lett.*, 1982, **49**, 1691.
186 W. Yang and W. J. Mortier, *J. Am. Chem. Soc.*, 1986, **108**, 5708.
187 R. K. Roy, S. Pal and K. Hirao, *J. Chem. Phys.*, 1999, **110**, 8236.
188 P. W. Ayers, *J. Chem. Phys.*, 2000, **113**, 10886.
189 P. Fuentealba, P. Perez and R. Contreras, *J. Chem. Phys.*, 2000, **113**, 2544.
190 R. F. Nalewajski and R. G. Parr, *Proc. Natl. Acad. Sci. U.S.A.*, 2000, **97**, 8879.
191 R. K. Roy, K. Hirao and S. Pal, *J. Chem. Phys.*, 2000, **113**, 1372.
192 R. K. Roy, K. Hirao, S. Krishnamurty and S. Pal, *J. Chem. Phys.*, 2001, **115**, 2901.
193 R. F. Nalewajski, *Phys. Chem. Chem. Phys.*, 2002, **4**, 1710.
194 J. Oláh, C. Van Alsenoy and A. B. Sannigrahi, *J. Phys. Chem. A*, 2002, **106**, 3885.
195 P. Bultinck, R. Carbo-Dorca and W. Langenaeker, *J. Chem. Phys.*, 2003, **118**, 4349.
196 M. Mandado, C. Van Alsenoy and R. A. Mosquera, *J. Phys. Chem. A*, 2004, **108**, 7050.
197 L. J. Bartolotti and P. W. Ayers, *J. Phys. Chem. A*, 2005, **109**, 1146.
198 M. Mandado, R. A. Mosquera, A. M. Grania and C. Van Alsenoy, *Tetrahedron*, 2005, **61**, 819.
199 M. Mandado, C. Van Alsenoy and R. A. Mosquera, *J. Phys. Chem. A*, 2005, **109**, 8624.
200 M. Mandado, C. Van Alsenoy and R. A. Mosquera, *Chem. Phys. Lett.*, 2005, **405**, 10.
201 R. G. Parr, P. W. Ayers and R. F. Nalewajski, *J. Phys. Chem. A*, 2005, **109**, 3957.
202 P. W. Ayers, *Theor. Chem. Acc.*, 2006, **115**, 370.
203 F. L. Hirshfeld, *Theor. Chim. Acta*, 1977, **44**, 129.
204 P. W. Ayers, R. C. Morrison and R. K. Roy, *J. Chem. Phys.*, 2002, **116**, 8731.
205 R. S. Mulliken, *J. Chem. Phys.*, 1955, **23**, 1833.
206 J. P. Foster and F. Weinhold, *J. Am. Chem. Soc.*, 1980, **102**, 7211.
207 A. B. Rives and F. Weinhold, *Int. J. Quantum Chem. Symp.*, 1980, **14**, 201.

208 A. E. Reed and F. Weinhold, *J. Chem. Phys.*, 1983, **78**, 4066.
209 A. E. Reed, R. B. Weinstock and F. Weinhold, *J. Chem. Phys.*, 1985, **83**, 735.
210 R. K. Roy, N. Tajima and K. Hirao, *J. Phys. Chem. A*, 2001, **105**, 2117.
211 P. W. Ayers, *Proc. Natl. Acad. Sci. USA*, 2000, **97**, 1959.
212 R. F. Nalewajski, *Int. J. Mol. Sci.*, 2002, **3**, 237.
213 P. Bultinck, C. Van Alsenoy, P. W. Ayers and R. Carbo-Dorca, *J. Chem. Phys.*, 2007, **126**, 144111.
214 S. Saha, R. Roy and P. Ayers, *Int. J. Quantum Chem.*, 2009, **109**, 1790.
215 P. W. Ayers, *Phys. Chem. Chem. Phys.*, 2006, **8**, 3387.
216 J. Melin, P. W. Ayers and J. V. Ortiz, *J. Phys. Chem. A*, 2007, **111**, 10017.
217 K. S. Min, A. G. DiPasquale, J. A. Golen, A. L. Rheingold and J. S. Miller, *J. Am. Chem. Soc.*, 2007, **129**, 2360.
218 J. Cioslowski, M. Martinov and S. T. Mixon, *J. Phys. Chem.*, 1993, **97**, 10948.
219 W. Yang, Y. Zhang and P. W. Ayers, *Phys. Rev. Lett.*, 2000, **84**, 84.
220 P. W. Ayers, F. De Proft, A. Borgoo and P. Geerlings, *J. Chem. Phys.*, 2007, **126**, 224108.
221 P. Bultinck, S. Fias, C. Van Alsenoy, P. W. Ayers and R. Carbo-Dorca, *J. Chem. Phys.*, 2007, **127**, 34102.
222 P. Bultinck and R. Carbo-Dorca, *J. Math. Chem.*, 2003, **34**, 67.
223 P. W. Ayers, C. Morell, F. De Proft and P. Geerlings, *Chem. Eur. J*, 2007, **13**, 8240.
224 A. K. Chandra and M. T. Nguyen, *Int. J. Mol. Sci.*, 2002, **3**, 310.
225 M. K. Harbola, P. K. Chattaraj and R. G. Parr, *Israel J. Chem.*, 1991, **31**, 395.
226 B. G. Baekelandt, A. Cedillo and R. G. Parr, *J. Chem. Phys.*, 1995, **103**, 8548.
227 M. H. Cohen, M. V. Ganduglia-Pirovano and J. Kudrnovsky, *J. Chem. Phys.*, 1995, **103**, 3543.
228 R. K. Roy, S. Krishnamurti, P. Geerlings and S. Pal, *J. Phys. Chem. A*, 1998, **102**, 3746.
229 M. Berkowitz and R. G. Parr, *J. Chem. Phys.*, 1988, **88**, 2554.
230 M. Berkowitz, S. K. Ghosh and R. G. Parr, *J. Am. Chem. Soc.*, 1985, **107**, 6811.
231 S. K. Ghosh and M. Berkowitz, *J. Chem. Phys.*, 1985, **83**, 2976.
232 P. K. Chattaraj, D. R. Roy, P. Geerlings and M. Torrent-Sucarrat, *Theor. Chem. Acc.*, 2007, **118**, 923.
233 R. G. Parr and J. L. Gazquez, *J. Phys. Chem.*, 1993, **97**, 3939.
234 S. K. Ghosh, *Chem. Phys. Lett.*, 1990, **172**, 77.
235 J. L. Gazquez, in *Struct. Bonding 80*, ed. K. D. Sen, Springer Verlag, Berlin, 1993, p. 27.
236 W. Langenaeker, F. de Proft and P. Geerlings, *J. Phys. Chem.*, 1995, **99**, 6424.
237 S. Liu, F. De Proft and R. G. Parr, *J. Phys. Chem.*, 1997, **101**, 6991.
238 T. Gál, *J. Phys. A: Math. Gen.*, 2002, **35**, 5899.
239 M. Torrent-Sucarrat, M. Duran and M. Sola, *J. Phys. Chem. A*, 2002, **106**, 4632.
240 T. Gál, *J. Math. Chem.*, 2007, **42**, 661.
241 M. Torrent-Sucarrat, P. Salvador, P. Geerlings and M. Solà, *J. Comput. Chem.*, 2007, **28**, 574.
242 P. W. Ayers and R. G. Parr, *J. Chem. Phys.*, 2008, **128**, 184108.
243 M. Torrent-Sucarrat, P. Salvador, M. Solà and P. Geerlings, *J. Comput. Chem.*, 2008, **29**, 1064.
244 For unconstrained local hardness vide ref. 59.
245 P. Mignon, S. Loverix and P. Geerlings, *Chem. Phys. Lett.*, 2005, **401**, 40.
246 A. S. Ozen, F. De Proft, V. Aviyente and P. Geerlings, *J. Phys. Chem. A*, 2006, **110**, 5860.
247 M. Torrent-Sucarrat and P. Geerlings, *J. Chem. Phys.*, 2006, **125**, 244101.
248 A. Borgoo, M. Torrent-Sucarrat, F. D. Proft and P. Geerlings, *J. Chem. Phys.*, 2007, **126**, 234104.
249 P. Mignon, P. Geerlings and R. Schoonheydt, *J. Phys. Chem. C*, 2007, **111**, 12376.
250 M. Torrent-Sucarrat, F. De Proft, P. Geerlings and Paul W. Ayers, *Chem.—Eur. J.*, 2008, **14**, 8652.
251 M. Torrent-Sucarrat, F. D. Proft, P. W. Ayers and P. Geerlings, *Phys. Chem. Chem. Phys.*, 2010DOI: 10.1039/b919471a.
252 S. Saha and R. Roy, *Phys. Chem. Chem. Phys.*, 2008, **10**, 5591.
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. 242.
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 Scientic, 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 G. Klopman, *J. Am. Chem. Soc.*, 1968, **90**, 223.

307 W. Yang, R. G. Parr and R. Pucci, *J. Chem. Phys*., 1984, **81**, 2862.
308 L. H. Thomas, *Proc. Cambridge Philos. Soc*., 1927, **23**, 542.
309 E. Fermi, *Z. Phys*., 1928, **48**, 73.
310 P. A. M. Dirac, *Proc. Cambridge Philos. Soc*., 1930, **26**, 376.
311 A. Aizman, R. Contreras and P. Perez, *Tetrahedron*, 2005, **61**, 889.
312 R. K. Roy, P. Bagaria, S. Naik, V. Kavala and B. K. Patel, *J. Phys. Chem. A*, 2006, **110**, 2181.
313 M. Torrent-Sucarrat, P. Geerlings, S. B. Liu and R. G. Parr, *Polish J. Chem*., 1998, **72**, 1737.
314 M. Torrent-Sucarrat, J. M. Luis, M. Duran and M. Sola, *J. Mol. Struct. (Theochem)*, 2005, **727**, 139.
315 P. Mignon, S. Loverix, F. De Proft and P. Geerlings, *J. Phys. Chem. A*, 2004, **108**, 6038.
316 Some recent studies are focused in this direction and Morell *et al.*[a] introduced a new dual descriptor which exhibits nucleophilicty as well as electrophilicity. This new index is defined in terms of the variation of hardness with respect to the external potential and is written as the difference between nucleophilic and electrophilic Fukui functions, thus being able to characterize both types of reactive behaviors. Also the studies by Morell and coworker[b,c] showed why the dual descriptor is a powerful local reactivity indicator for regioselectivity. Ref: (*a*) C. Morell, A. Grand and A. Toro-Labbé, *J. Phys. Chem. A*, 2005, **109**, 205; (*b*) ref. 223; (*c*) C. Cárdenas, N. Rabi, P. W. Ayers, C. Morell, P. Jaramillo and P. Fuentealba, *J. Phys. Chem. A*, 2009, **113**, 8660; (*d*) V. Labet, C. Morell, J. Cadet, L. A. Eriksson and A. Grand, *J. Phys. Chem. A*, 2009, **113**, 2524.
317 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.
318 P. D. Lawley and P. Brookes, *Nature (London)*, 1961, **192**, 1081.
319 C. Nagata, A. Imamura, H. Salto and K. Fukui, *Gann*, 1963, **54**, 109.
320 A. Dipple, P. Brookes, D. S. Mackintosh and M. P. Rayman, *Biochemistry*, 1971, **10**, 4323.
321 J. J. Danneberg and M. Tomasz, *J. Am. Chem. Soc*., 2000, **122**, 2062.
322 K. S. Gates, T. Nooner and S. Dutta, *Chem. Res. Toxicol*., 2004, **17**, 839.
323 A. Loveless, *Nature (London)*, 1969, **223**, 206.
324 B. Singer, *Nature (London)*, 1976, **264**, 333.
325 B. Singer and J. M. Essigmann, *Carcinogenesis*, 1991, **12**, 949.
326 N. Guex and M. C. Peitsch, SWISS-MODEL and the Swiss-Pdb-Viewer: An environment for comparative protein modeling, *Electrophoresis*, 1997, **18**, 2714; http://www.expasy.org/spdbv/.
327 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.
328 W. J. Fan, R. Q. Zhang and S. Liu, *J. Comput. Chem*., 2007, **28**, 967.
329 C. Barrientos-Salcedo, D. Arenas-Aranda, F. Salamanca-Gomez, R. Ortiz-Muniz and C. Soriano-Correa, *J. Phys. Chem. A*, 2007, **111**, 4362.
330 R. Rosal, M. Pincus, P. W. Brandt-Rauf, R. L. Fine, J. Michl and H. Wang, *Biochemistry*, 2004, **43**, 1854.
331 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.
332 Y. Li and J. N. S. Evans, *J. Am. Chem. Soc*., 1995, **117**, 7756.
333 Y. Li and J. N. S. Evans, *Proc. Natl. Acad. Sci. U.S.A*., 1996, **93**, 4612.
334 M. Berkowitz, *J. Am. Chem. Soc*., 1987, **109**, 4823.
335 P. K. Chattaraj, *J. Phys. Chem. A*, 2000, **105**, 511.
336 P. Srinivasan and D. Sprinson, *J. Biol. Chem*., 1959, **234**, 716.
337 A. B. DeLeo and D. B. Sprinson, *Biochem. Biophys. Res. Commun*., 1968, **32**, 873.
338 H. Floss, D. Onderka and M. Carroll, *J. Biol. Chem*., 1972, **247**, 736.

339 D. L. Anton, L. Hedstrom, S. M. Fish and R. H. Abeles, *Biochemistry*, 1983, **22**, 5903.
340 C. E. Grimshaw, S. G. Sogo, S. D. Copley and J. R. Knowles, *J. Am. Chem. Soc.*, 1984, **106**, 2699.
341 L. Hedstrom and R. Abeles, *Biochem. Biophys. Res. Commun.*, 1988, **157**, 816.
342 F. M. Unger, *Adv. Carbohydr. Chem. Biochem.*, 2001, **57**, 207.
343 A. Baeten, F. De Proft and P. Geerlings, *Int. J. Quantum Chem.*, 1996, **60**, 931.
344 F. De Proft, W. Langenaeker and P. Geerlings, *J. Phys. Chem.*, 1993, **97**, 1826.
345 A. Baeten, F. De Proft, W. Langenaeker and P. Geerlings, *J. Mol. Struct. (Theochem)*, 1994, **306**, 203.
346 F. De Proft, S. Amira, K. Choho and P. Geerlings, *J. Phys. Chem.*, 1994, **98**, 5227.
347 W. Langenaeker, N. Coussement, F. De Proft and P. Geerlings, *J. Phys. Chem.*, 1994, **98**, 3010.
348 A. Baeten, F. De Proft and P. Geerlings, *Chem. Phys. Lett.*, 1995, **235**, 17.
349 S. Damoun, W. Langenaeker, G. Van de Woude and P. Geerlings, *J. Phys. Chem.*, 1995, **99**, 12151.
350 F. De Proft, W. Langenaeker and P. Geerlings, *Tetrahedron*, 1995, **51**, 4021.
351 F. D. Proft, W. Langenaeker and P. Geerlings, *Int. J. Quantum Chem.*, 1995, **55**, 459.
352 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.
353 A. Baeten, D. Maes and P. Geerlings, *J. Theor. Biol.*, 1998, **195**, 27.
354 R. Bott, M. Ultsch, A. Kossiakoff, T. Graycar, B. Katz and S. Power, *J. Biol. Chem.*, 1988, **263**, 7895.
355 A. Russell and A. Fersht, *Nature*, 1987, **328**, 496.
356 A. J. Russell, P. G. Thomas and A. R. Fersht, *J. Mol. Biol.*, 1987, **193**, 803.
357 J. Wells and D. Estell, *Trends Biochem. Sci.*, 1988, **13**, 291.
358 P. Mignon, J. Steyaert, R. Loris, P. Geerlings and S. Loverix, *J. Biol. Chem.*, 2002, **277**, 36770.
359 R. T. Raines, *Chem. Rev.*, 1998, **98**, 1045.
360 J. Steyaert, *Eur. J. Biochem.*, 1997, **247**, 1.
361 S. Loverix and J. Steyaert, *Methods Enzymol.*, 2001, **341**, 305.
362 S. Krishnamurty and S. Pal, *J. Phys. Chem. A*, 2000, **104**, 7639.
363 P. Rivas, G. Zapata-Torres, J. Melin and R. Contreras, *Tetrahedron*, 2004, **60**, 4189.
364 G. Roos, P. Geerlings and J. Messens, *J. Phys. Chem. B*, 2009, **113**, 13465.
365 S. Shinkai, N. Honda, Y. Ishikawa and O. Manabe, *J. Am. Chem. Soc.*, 1985, **107**, 6286.
366 O. Tapia, R. Cardenas, J. Andres and F. Colonna-Cesari, *J. Am. Chem. Soc.*, 1988, **110**, 4046.
367 P. A. Karplus, M. J. Daniels and J. R. Herriott, *Science*, 1991, **251**, 60.
368 F. Fieschi, V. Nivière, C. Frier, J.-L. Décout and M. Fontecave, *J. Biol. Chem.*, 1995, **270**, 30392.
369 J. Andres, V. Moliner, V. S. Safont, L. R. Domingo and M. T. Picher, *J. Org. Chem.*, 1996, **61**, 7777.
370 J. J. Tanner, B. Lei, S.-C. Tu and K. L. Krause, *Biochemistry*, 1996, **35**, 13531.
371 R. Castillo, J. Andres and V. Moliner, *J. Am. Chem. Soc.*, 1999, **121**, 12140.
372 G. Roos, S. Loverix, F. De Proft, L. Wyns and P. Geerlings, *J. Phys. Chem. A*, 2003, **107**, 6828.
373 G. Roos, J. Messens, S. Loverix, L. Wyns and P. Geerlings, *J. Phys. Chem. B*, 2004, **108**, 17216.
374 G. Roos, L. Buts, K. van Belle, E. Brosens, P. Geerlings, R. Loris, L. Wyns and J. Messens, *J. Mol. Biol.*, 2006, **360**, 826.
375 G. Roos, N. Foloppe, K. Van Laer, L. Wyns, L. Nilsson, P. Geerlings and J. Messens, *PLoS Comput. Biol.*, 2009, **5**, 1.
376 I. Zegers, J. C. Martins, R. Willem, L. Wyns and J. Messens, *Nat. Struct. Biol.*, 2001, **8**, 843.
377 J. Messens and S. Silver, *J. Mol. Biol.*, 2006, **362**, 1.
378 M. S. Bennett, Z. Guan, M. Laurberg and X. D. Su, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 13577.
379 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.
380 X. Guo, Y. Li, K. Peng, Y. Hu, C. Li, B. Xia and C. Jin, *J. Biol. Chem.*, 2005, **280**, 39601.

381 J. Messens, G. Hayburn, A. Desmyter, G. Laus and L. Wyns, *Biochemistry*, 1999, **38**, 16857.
382 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.
383 J. L. Martin, *Structure*, 1995, **3**, 245.
384 K. Assemat, P. M. Alzari and J. Clément-Métral, *Protein Sci.*, 1995, **4**, 2510.
385 Y. Li, Y. Hu, X. Zhang, H. Xu, E. Lescop, B. Xia and C. Jin, *J. Biol. Chem.*, 2007, **282**, 11078.
386 G. Roos, S. Loverix, E. Brosens, K. Van Belle, L. Wyns, P. Geerlings and J. Messens, *Chem. Bio. Chem.*, 2006, **7**, 981.