Chapter 1

Background of the Theoretical Methods and Abstract of the Thesis

Contents

1.1 General Introduction

Computational chemistry is an emerging field, where computer is used as an 'experimental' tool to generate data, by which one may gain insight and rationalize the behavior of a large class of chemical systems. The increase in the threshold of this subfield of theoretical chemistry is due to the advancements in computer technology, availability of practical algorithms for theoretical methods and success in explaining the problems.[1] It is developed into an important tool in almost all areas of chemistry. It is an eminent tool for both theoreticians and experimentalists in analyzing the chemistry related problems. The three key steps involved in solving chemistry related problems through computational methods are (a) building an appropriate chemical model for the real system, (b) calculation of total energies and its derived properties of the system and (c) analyzing the properties obtained from calculation. Accuracy of these calculations depends on the factors such as (a) closeness of the model system to the real one, (b) type of theoretical method and (c) evaluating the quality of the obtained results in comparison to experiments. Thus the inclusion of theoretical and experimental results, knowing the limitation of theory and why the numbers are produced the way they are helps to give a better understanding on the problems.[1,2]

There are many properties which can be studied using computational methods. Those include the structure, bonding, stability, reactivity, rates of reactions, interactions, excitation, thermal, mechanical and magnetic properties, phase transitions etc for various classes of systems (clusters, solids and biomolecules). The structure and bonding interactions of the molecules, clusters and biomolecules are the interest of this thesis work. Structure can be determined experimentally from various spectroscopic techniques

such as X-ray, electron and neutron diffraction, mass, microwave and NMR etc.[2b] The structures are available in Cambridge Crystallographic Data Centre or Cambridge Structural Database[3a] for small inorganic and organic molecules where as biomolecules are present in Protein and Nucleic acid Data Banks.[3b] On the other hand theoretically one can probe the structure by calculating the potential energy surface of the small molecules using the computational quantum mechanical methods and for biomolecules using molecular mechanical methods such as homology modeling and molecular dynamics. The theoretical methods that are used in the thesis are briefly discussed in the following sub sections. The outline of the problems discussed in the next four chapters is also presented.

1.2 Overview of Theoretical Methods

There are various levels of theoretical methods and can be broadly classified as quantum mechanical (QM), molecular mechanical (MM) and hybrid quantum and molecular mechanical (QM/MM).[1,4] The choice of these methods depends on the size of the system, accuracy and property that need to be calculated. Quantum mechanics can further be wave function and density functional based methods.[4]

1.3 Quantum Mechanical Methods[4]

The phenomenon such as black body radiation, Compton and photoelectric effect and specific heats of microscopic objects were unable to explain by Newtonian or classical mechanics. The quantum mechanics explains the motion of microscopic particles and their properties.[4] The fundamental equation of quantum mechanics that explains the state of the physical system is Schrödinger equation.

$$H\ \Psi(r,t) = i\hbar\frac{\partial}{\partial t}\Psi(r,t) \tag{1.1}$$

$$H = T + V = -\frac{\hbar^2}{2m}\nabla^2 + V \tag{1.2}$$

where H is Hamiltonian operator that includes both kinetic (T) and potential energy (V) operators, $\Psi(r,t)$ is the wave function which represents the state of system, i = $\sqrt{-1}$ and $\hbar$ is h/2π where h is Planck's constant. The equation 1.1 is the non-relativistic time dependent Schrödinger equation.[5] The equation which gives the stationary state of the particle is time independent Schrödinger equation (1.3). In this thesis only time independent form was used.

$$H\Psi(r) = E\Psi(r) \tag{1.3}$$

In principle, Schrödinger equation can be written for any problem but it can be exactly solvable only for particles in a box, hydrogen and hydrogen like atoms etc. In order to solve for many body systems various approximations are needed and are illustrated here.

1.3.1 The Born-Oppenheimer Approximation

This is the first approximation[6] that simplifies the solution of Schrödinger equation for many body systems. This considers electronic and nuclear motion are separable. As the motion of nucleus is negligible compared to electron motion (M_{nuceli} = 1837 m_e) electron readjusts its position before the nucleus change its position. This leads to electronic Schrödinger equation (1.4) that describes the motion of electron in the field of fixed nuclei.

$$H_{el}\Psi_{el}(r,R) = E_{el}(R)\Psi_{el}(r,R) \tag{1.4}$$

where $H_{el} = -\frac{\hbar^2}{2m_e}\sum_i^n \nabla_i^2 - \sum_A^N \sum_i^n \frac{Z_A e^2}{r_{iA}} + \sum_j^n \sum_{i>j}^n \frac{e^2}{r_{ij}} + \sum_A^N \sum_{B>A}^N \frac{Z_A Z_B e^2}{R_{AB}}$ (1.5)

$$H_{el} = -\frac{1}{2}\sum_i^n \nabla_i^2 - \sum_A^N \sum_i^n \frac{1}{r_{iA}} + \sum_j^n \sum_{i>j}^n \frac{1}{r_{ij}} + \sum_A^N \sum_{B>A}^N \frac{1}{R_{AB}} \text{ in atomic units} \quad (1.6)$$

The first term in the Hamiltonian corresponds to kinetic energy of the electron and the next three terms represent the potential energy due to electron-nuclei, electron-electron and nuclei-nuclei interactions. Solving the equation 1.4 at different nuclear coordinates describe the potential energy surface for the system. The Born-Oppenheimer approximation is supported from spectral studies. Limitation of this approximation is that it cannot explain whenever electronic motion is coupled to nuclear motion such as Jahn Teller, Renner and Peierl's structural distortions, excited states, highly charged systems, as well as occurrences of charge- and spin-density waves and superconductivity.[3g] The methods that solve the equation 1.4 using only the fundamental constants (electronic mass m_e, charge *e*, Planck's constant *h* etc) and do not include any experimental data are called *ab initio* or first principle electronic structure methods. However, solving the electronic Schrödinger equation (1.4) exactly for many body systems is difficult due to the electron-electron interactions.

1.3.2 Hartree-Fock Method

Hartree-Fock (HF) method[4,7] is an iterative variational self consistent procedure to calculate the single Slater determinant. This is the foundation for much of modern molecular orbital theory and starting point for many other electron correlation methods. In HF method each electron's motion is described by a single-particle function (orbital) which does not depend explicitly on the instantaneous motions of the other electrons.

Thus many electron Schrödinger equation is simplified as many one electron equations (1.7).

$$\left(H_i^{core} + V_i^{eff}\right)\phi_i = \varepsilon_i \phi_i \tag{1.7}$$

where $H_i^{core} = -\frac{1}{2}\nabla_i^2 - \frac{Z}{r_i}$ in a.u is the average kinetic energy and nuclear electron attraction, V_i^{eff} is the effective potential of electron-electron interaction,

$$V_i^{eff} = \sum_{j \neq i}^{N} \int \frac{\phi_j^*(j)\phi_j(j)}{r_{ij}} d\tau \tag{1.8}$$

ϕ_i is the spatial one electron wave function called molecular orbital and ε_i is the one electron energy. The total Hamiltonian of equation 1.7 by expanding V_i^{eff} in terms of coulomb J_{ij} and exchange operator K_{ij} gives the Fock operator[7b] (1.9).

$$F_i = -\frac{1}{2}\nabla_i^2 - \sum_i^N \frac{Z_\alpha}{r_{i\alpha}} + \sum_{i=1}^{N/2}\sum_{j=1}^{N/2} 2J_{ij} - K_{ij} \tag{1.9}$$

where $J_{ij} = \int \phi_i(1)\phi_j(2)\left(\frac{1}{r_{12}}\right)\phi_i^*(1)\phi_j^*(2)d\tau$ (1.10)

$$K_{ij} = \int \phi_i(1)\phi_j(2)\left(\frac{1}{r_{12}}\right)\phi_i^*(2)\phi_j^*(1)d\tau \tag{1.11}$$

J_{ij} represents the classical repulsion between electrons and K_{ij} represents the exchange interaction that is coming out from the anti symmetric nature of the wave function. These equations are called as Hartree-Fock equations. The total electronic energy by HF method for a closed shell system is obtained by summing up these one electron energies.

$$E_{HF} = \sum_{i=1}^{N/2}\left[2\varepsilon_i + \sum_{j=1}^{N/2} 2J_{ij} - K_{ij}\right] + V_{NN} \tag{1.12}$$

HF equations are mathematically complicated nonlinear equations which can be solved directly only for simple systems through numerical integration. For a molecular system, the equations can be derived by expanding the unknown molecular orbital ϕ_i in terms of given basis set χ is $\phi_i = \sum_{u=1}^{n} c_{ui}\chi_u$ (1.13)

where c_{ui} is the orbital coefficient. This is called linear combination of atomic orbital (LCAO)- molecular orbital (MO) method.[8] The initial set of orbital coefficients is solved by using the principles of variational theorem. It states that expectation value of energy E_ϕ calculated using any trial function of electron coordinates will be greater than the energy for exact wave function.

$$E_\phi = \frac{\int \phi^* H \phi d\tau}{\int \phi^* \phi d\tau} \geq E_0 \qquad (1.14)$$

Substitution of ϕ (1.13) in above equation and introducing the resonance integral H_{uv} and overlap integral S_{uv} gives

$$\sum_{u=1}^{N} c_u \sum_{v=1}^{N} c_v S_{uv} E = \sum_{u=1}^{N} c_u \sum_{v=1}^{N} c_v H_{uv} \qquad (1.15)$$

where $H_{uv} = \int \phi_u^* H \phi_v d\tau$ and $S_{uv} = \int \phi_u^* \phi_v d\tau$

The best wave function is obtained by minimizing the energy with respect to orbital coefficients c_{ui}

$$\frac{\partial E}{\partial c_{ui}} = 0 \qquad (1.16)$$

This gives N simultaneous linear and homogenous equations which can be written in determinantal form called as secular determinant (1.17).

$$\begin{vmatrix} H_{11}-ES_{11} & H_{12}-ES_{12} & \ldots & H_{1n}-ES_{1n} \\ H_{21}-ES_{21} & H_{22}-ES_{22} & \ldots & H_{2n}-ES_{2n} \\ \ldots & & & \\ H_{n1}-ES_{n1} & H_{n2}-ES_{n2} & \ldots & H_{nn}-ES_{nn} \end{vmatrix} = 0 \qquad (1.17)$$

The solution of this n^{th} order determinant will give n roots i.e., a set of n energy values, $E_1, E_2, E_3, \ldots, E_n$ and a set of n wave functions, $\phi_1, \phi_2, \ldots, \phi_n$ for the system. Antisymmetric nature of complete wave function Ψ_{el} leads to the following Slater determinantal form of the wave function,

$$\Psi_{el} = \frac{1}{\sqrt{n!}} \begin{vmatrix} \phi_1(1) & \phi_2(1) & \ldots & \phi_n(1) \\ \phi_1(2) & \phi_2(2) & \ldots & \phi_n(2) \\ \ldots & & & \\ \phi_1(n) & \phi_2(n) & \ldots & \phi_n(n) \end{vmatrix} \qquad (1.18)$$

where, ϕ_i is the spin orbital of i^{th} particle.

Variation of the total energy (1.15) when carried out with respect to the coefficients c_{ui} leads to a set of algebraic equations and can be written in matrix form as Roothan equations[9] (1.19).

$$FC = SCE \qquad (1.19)$$

where F is Fock matrix, C is matrix representing molecular orbital coefficients, S is the overlap matrix for overlap between orbitals and $E = \varepsilon_i \delta_{ij}$ where ε_i are one electron Fock energies.

$$F = H_{uv}^{core} + \sum_{u=1}^{N}\sum_{v=1}^{N} P_{\lambda\sigma}\left[(uv|\lambda\sigma) - \frac{1}{2}(u\lambda|v\sigma)\right] \qquad (1.20)$$

where H_{uv}^{core} is the matrix representing energy of the single electron in the field of nuclei and $P_{\lambda\sigma} = 2\sum_{i=1}^{occ} C_{\lambda i}^{*} C_{\sigma i}$ is the density matrix.

A rigorous solution of HF-Roothan equations without any empirical parameters is known as *ab initio* molecular orbital theory. Since *F* depends on the coefficients of *C*, the HF-Roothan equations are solved by choosing an initial set of *C* and calculating matrix *F*. Then a new set of *C* is obtained from equations 1.19. This process is repeated until they are consistent. Hence this method is referred to as self consistent field (SCF) method. Basis set employed plays a crucial role in this method. Thus a better expectation value of energy can be obtained with larger basis set. The limit at which there is no change in the HF energy further by increasing the basis set is known as HF limit.[9d] HF methods predict the ground state and related properties of most molecules. This method also gives 99% of total energy of the system. The bond breaking and making process are poorly determined by this method due to poor estimation of the correlation of electron pairs of opposite spin.

1.3.2.1 Basis Sets[4]

It is a set of mathematical functions used to represent the molecular orbitals for a system. Basis set assign a group of basis function to each atom within a molecule to approximate its orbitals.[4] Depending upon the functional form basis set can be (a) Slater type[10a] (b) Gaussian type[10b] (c) Effective core potential[10c] (d) Plane wave[10d] etc.

(a) Slater type orbitals (STO): Atomic orbitals are best represented by STO's. The functional form of STO resembles the hydrogen like orbitals.

$$\lambda_{\zeta,n,l,m}(r,\theta,\varphi) = Nr^{(n-1)}e^{-\zeta r}Y_l^m(\theta,\varphi) \qquad (1.21)$$

where ζ orbital exponent is $\frac{Z-S}{a_o} = \frac{Z_{eff}}{a_o}$, S is screening constant, N is the normalization constant, $Y_{lm}(\theta,\varphi)$ contains all the angular information needed to describe the wave function (spherical harmonics) and r, θ φ represent the polar coordinates. STO's

gives accurate solutions for atomic, diatomic and simpler linear molecules. For other molecules evaluation of many centre two electron integrals is very difficult and is time consuming computations.

(b) Gaussian type orbitals (GTO): This is introduced by Boys[10b] and its functional form in Cartesian coordinates is $g_{\zeta,l_x,l_y,l_z}\left(x,y,z\right) = Nx^{l_x}y^{l_y}z^{l_z}e^{-\zeta r^2}$ (1.22)

where ζ represents the orbital exponent and these are primitive gaussians. Actual basis functions or contracted Gaussian functions χ_u are linear combinations of such primitive Gaussians $\chi_u = \sum_p d_{up} g_p$ (1.23)

where d_{up} are contraction functions and g_p are primitive gaussians.

Unlike STO's there is r^2 dependence in the exponential for GTO (1.22). Thus at nucleus it leads to zero slope and fails to represent the region near the nucleus. In practice, the functional behavior of an STO is reproduced by using number of GTO's with different orbital exponents. GTO's are the popular basis functions used in electronic structure calculations. These are called STO-KG basis sets where K is the number of primitive Gaussian functions to fit an STO. They are more popular and developed by John A. Pople. STO-3G is a minimal basis set that contains minimum number of basis functions needed for each atom (3 Gaussians to fit an STO). Since the valence orbitals are involved in chemical bonding, each valence atomic orbital can be better described by increasing the basis functions. This type of separate treatment for core and valence atomic orbitals is referred as split valence basis set. Depending on the number of basis functions used for valence atomic orbital they can be double zeta (DZ), triple zeta (TZ) etc. To describe the expansion or contraction of orbitals that is caused due to environment

or for polarization related properties, a high angular momentum functions like p-functions for H and d- functions for heavy atoms are added. This type changes the shape of orbital and is termed as polarized basis set. Diffused basis set is also adding the functions to heavy and H atoms to give large region of space orbital. This is needed especially for describing the systems where electrons are far away from nucleus such as in anions, molecules with lone pairs, systems with excited states and low ionization potential etc. General representation of Pople type basis set is

$$\text{k-nlm++G(d,p) or k-nlm++G**} \qquad (1.24)$$

The basis sets developed for dealing correlated systems are correlation consistent basis sets. They are developed by Dunning and represented generally as cc-PVXZ type (X= D, T, Q).

A 6-311++G (d,p) basis set was used for Boradiphosphole ($BC_2P_2H_3$) in chapter 2. It includes six Gaussians for core orbitals, valence orbital splits into three shells composed of three, one and one Gaussians and each heavy and hydrogen atom consists of diffuse and polarization functions. While for triple-decker sandwich complexes ($CpMP_6MCp$) in chapter 3 double zeta (DZ), triple zeta valence (TZV) and triple zeta valence polarization (TZVP) basis sets and chapter 4 for pentapeptide calculation 6-31+G(d) basis set were used.

(c) Effective core potential (ECP): In the case of atoms with high atomic number the number of orbitals increases with increase in electrons. This leads to more computational cost. Since the core orbital energy levels do not change significantly due to chemical bonding they are represented by average potential. Effective core potential is one of the pseudo potential whose functional form is

$$\mathrm{ECP}(r) = \sum_{i=1}^{M} d_i r^{n_i} e^{-\xi_i r^2} \tag{1.25}$$

where d_i is the coefficient for each term, r is the distance from the nucleus with a power of n_i for the i^{th} term and ξ_i is an exponent of the i^{th} term. ECP's takes very less computational time as basis functions are drastically reduced due to core potential and involves only calculation for valence electrons. This basis sets are widely used for transition metals. Los Alamos National Laboratory double-zeta (LANL2DZ) basis set is one such basis set developed by Hay and Wadt.[10e] This basis set is used in chapter 3 for transition metals in triple-decker sandwich complexes.

(d) Plane wave: Unlike other functions they are completely delocalized and therefore not to be ascribed for atoms. The functional form is

$$\chi_i(r) = e^{ikr} \tag{1.26}$$

where k is the momentum vector and r run through the Bravais lattice.

The right choice of basis set is important for quantum mechanical methods. One should be aware of Basis set superposition error (BSSE)[10f] especially in calculating interaction (dimerization) energies which are overestimated. This is due to the expanded larger basis functions description for the wave function of the complex than the component. One of the methods to rectify this problem is counterpoise.[10g]

1.3.2.2 Semi Empirical Methods

Semi empirical approaches[4] start from HF equations, where the core electrons are not included for the calculation. They use a minimal basis set (STO-3G) and replace the atomic nuclei plus core electron by ion cores of charges Z_v. There are a range of such methods which differ basically on the treatment of matrix elements of Fock operator

(1.20).[11] The methods can be broadly classified as two types (a) Involves only one electron terms of Hamiltonian and (b) Involves both one and two electron Hamiltonian terms. Hückel and extended Hückel methods involve only one electron Hamiltonians (1.27).[12] The former method involves only π electrons and latter includes all valence electrons. They are non SCF methods. The results obtained from these methods are useful for quick qualitative determination of wave function and predicting the geometric and energy trends.

$$H = \sum_{i=1}^{N} h_{eff}(i) \qquad (1.27)$$

On the other hand Pariser-Parr-Pople (PPP),[13a,b] Intermediate Neglect of Differential Overlap INDO, MNDO[13c] methods, Austin model 1 (AM1)[13d] and Parametric model 3 (PM3)[13e] involve Hamiltonian that has both one and two electron terms which are replaced by parameters. These methods differ in the extent to which zero differential overlap approximation (ZDO) is invoked in evaluating electron repulsions integrals and the way parameters are specified.[12d] ZDO is an approach to the systematic neglect of the small-in-value electron repulsion integrals. It assumes that all the products of atomic orbitals $\chi_\mu\chi_\nu$ are set to zero and the overlap integral $S_{\mu\nu} = \delta_{\mu\nu}$ (where $\delta_{\mu\nu}$ is the Kronecker delta). The ZDO approximation greatly simplifies the computation of wave functions by eliminating many of two-electron integrals. Thus neglecting of integrals and parameterization makes these calculations applied for larger systems at low cost of CPU time. AM1 and PM3 are popularized semi empirical methods used for larger systems and in hybrid QM/MM methods.

1.3.3 Post Hartree-Fock Methods

The major drawback of HF method is neglecting the electron correlation between the motions of electrons of anti-parallel spin. Hence it results severe underestimation of bond dissociation energies and deficiency in describing the compounds exhibiting Jahn-Teller distortions. Even the single determinant HF wave function may not always describe the full symmetry of the wave function. Therefore, Hartree-Fock method (single determinant approach) is unsuccessful in predicting the properties where many body correlation effects contribute significantly. The difference in the HF energy (E_{HF}) and the exact non-relativistic energy (E_{exact}) is called correlation energy (E_{corr}).

$$E_{cor} = E_{exact} - E_{HF} \tag{1.28}$$

The methods that are developed after HF explicitly treat the correlation effects and are known as Post Hartree-Fock or electron correlation methods.[4,14] These methods should have size extensivity and size consistent properties especially when calculating, ionization potential, electron affinity, dissociation energies and oligomer gaps.

(a) Size extensivity: when a method scales correctly with the number of particles in the system as the exact energy does, then it is said to be size-extensive.

(b) Size consistent: when there are two non-interacting systems A and B. If a given method calculates the energy of the complex system A-B as equal to the sum of the energy of A plus the energy of B taken by themselves ($E(A - B) = E(A) + E(B)$) then it is size consistent.

Most common electron correlation methods are Møller-Plesset Perturbation Theory,[15] Configuration Interaction[16] and Coupled Cluster Theory.[17]

1.3.3.1 Many Body Perturbation Theory

Here the correlation is added as perturbation to HF wave function. Hamiltonian for perturbation theory is expressed as

$$H = H^{(0)} + \lambda H^{(1)} + \lambda^2 H^{(2)} + ... \quad (1.29)$$

The wave function and energy are also expanded in power series of perturbation.

$$\psi = \psi^{(0)} + \lambda\psi^{(1)} + \lambda^2\psi^{(2)}.... \quad (1.30)$$

$$E = E^{(0)} + \lambda E^{(1)} + \lambda^2 E^{(2)}.... \quad (1.31)$$

This leads Schrödinger equation to

$$H^{(0)}\psi^{(0)} = E^{(0)}\psi^{(0)} \quad (1.32)$$

$$\left(H^{(0)} - E^{(0)}\right)\psi^{(1)} = \left(H^{(1)} - E^{(1)}\right)\psi^{(0)} \quad (1.33)$$

These solutions give different order of correction to energy. Depending on the truncation of the series, it is Møller-Plesset 2 (MP2), MP3 and MP4 etc. These methods are size consistent even though they are non-variational.[15] Among MP_n levels, MP_4 is considered to attain 90-95% of correlation effect. The limitation of these methods is that they are computationally demanding for larger systems.

1.3.3.2 Configuration Interaction (CI)

This method solves variationally a multi determinant wave function, which includes the excited states for the description of an electronic state.[16]

$$\Psi = c_o\psi_o + \sum_{ic} c_i^c\psi_i^c + ... \quad (1.34)$$

where Ψ is usually an electronic ground state of the system, ψ_i's are called Configuration state functions (CSF's) which are Slater determinants containing products

of spin orbitals with ψ_0 being the Hartree-Fock wave function (1.18) and c_0 and c_1 are the coefficients of determinants.

The method is constructed by starting with HF equations and new determinants are made by promoting the electrons from occupied to unoccupied orbitals. Based on the number of excitations used to make each determinant, CI calculations can be (a) CIS involves single excitations, (b) CISD involves single and double excitations and (c) Full CI includes all possible excitations. This method calculates energy very accurately at each level. The difference between the E_{HF} and the full CI energy is the basis set correlation energy. The drawbacks of this method are (a) truncated CI are size inconsistent and (b) limited to small molecules as cost of CPU time is very high (N^8). Multi configuration self consistent field (MCSCF) is also a multi determinant approach where both coefficients of basis functions and determinants are varied. This method is desirable whenever HF wave function gives a poor qualitative description for the system. While Complete Active Space Self Consistent Field (CASSCF) is an MCSCF method where all combinations of active space orbitals are included.

1.3.3.3 Coupled Cluster Theory

This method uses exponential form of wave function as linear combinations of many determinants.[17]

$$\Psi = e^{T}\Psi_{HF} \tag{1.35}$$

where Ψ is the exact non-relativistic ground state molecular electronic wave function, Ψ_{HF} is the normalized ground state Hartree-Fock equation, the operator e^{T} is defined by Taylor series expansion and the cluster or excitation operator $T = T_1 + T_2 \ldots T_n$.

$$e^{T} = 1 + T + \frac{T^2}{2!} + \frac{T^3}{3!} + \ldots \tag{1.36}$$

$$T_1\Psi_0 = \sum_{a=n+1}^{\infty}\sum_{i=1}^{n} t_i^a \Psi_i^a \tag{1.37}$$

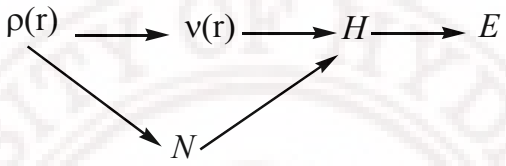
where T_1 is a one particle excitation operator, Ψ_i^a is a singly excited Slater determinant with occupied spin orbital χ_i replaced by the virtual spin orbital χ_a and t_i^a is a numerical coefficient whose value depend on *i* and *a* that are determined form equation 1.35. Even in this method based on the excitations involved in T form of operator they can be CCD, CCSD etc. Unlike CI, CC calculation of singles, doubles refer to all interactions associated with that quantity of particles (one and two body interactions).

This is one of the best method for estimating the electron correlation and predicting the molecular properties. This method is size extensive but size consistency depends on reference HF wave function. The drawback of this method is the determination of the cluster amplitudes *t* for all of the operators included in the particular approximation. This method is also limited to small molecules only. QCISD is an intermediate method of CC and CI developed by Pople and co-workers.[18] It is widely used for small molecules and gives good results for calculating correlation energies. All the post Hartree-Fock methods are computationally very expensive and can be handled in solving only for small size systems.

1.3.4 Density Functional Theory

In HF and Post HF methods the energy and its derived properties are obtained from wave function. For a many body system, wave function is complex due to the increase in number of variables (3 spatial and a spin coordinates for each electron). Hence they cannot be applied for larger systems. While density functional theory derives

energies and related properties from the electron probability density $\rho(x,y,z)$, which depends only on 3 variables. This was proposed by Thomas-Fermi[19] for non-interacting uniform electron gas, which was later proved by Hohenberg-Kohn theorems.[20a,b] The theorems states that the ground state energy of an N electron system is functional of the electron density $\rho(r)$, and energy is minimum when evaluated with the exact ground state density.



Consequently energy is functional of density $E = E_v[\rho]$ (1.38)

where $E_v[\rho] = T[\rho] + V_{ee}[\rho] + V_{ext}[\rho]$ (1.39)

and $T[\rho]$ is the electronic kinetic energy functional, V_{ee} is the electron-electron interaction potential functional and V_{ext} is the electron nuclei potential energy interaction called as external potential given as

$$V_{ext} = \int \upsilon(r)\rho(r)dr = \sum_{A=1}^{N} \frac{Z_A}{r_{iA}} \quad (1.40)$$

The first two terms of equation 1.39 are unknown as they are independent of external potential and can be determined by minimizing the energy with respect to variations in $\rho(r)$ (second Hohenberg-Kohn theorem, 1.41).

$$E_v[\rho] \geq E_v[\rho_o] \quad (1.41)$$

The real application of these theorems in solving chemical problems came into light only after Kohn-Sham formalism.[20c] This formalism has given an idea for (a)

determining the density without finding the wave function and (b) determining the properties from the electron density.

$$\left[-\frac{1}{2}\nabla_i^2 + \upsilon_{eff}(r)\right]\psi_i = \varepsilon_i\psi_i \tag{1.42}$$

where first two terms represent the one electron Hamiltonian, ψ_i is Kohn-Sham orbital, ε_i is the Kohn-Sham orbital energy and υ_{eff} is the effective potential given as

$$\upsilon_{eff}(r) = \upsilon_{ext}(r) + \upsilon_{coul} + \upsilon_{xc} \tag{1.43}$$

where υ_{ext} is external potential (nuclei-electron attraction), υ_{coul} is the columbic repulsion of electron- electron, $\upsilon_{xc} = \frac{\delta E_{xc}[\rho(r)]}{\delta\rho}$ is exchange correlation potential and E_{xc} is the exchange correlation energy. These equations are also solved iteratively as HF equations and the total energy is given as

$$E_{DFT} = \sum_{i=1}^{N}\varepsilon_i + V_{NN} \tag{1.44}$$

The advantages of DFT over HF method are (a) includes correlation effects (as the Hamiltonian of the Kohn-Sham equation 1.42 considers both exchange and correlation part in the form of functional υ_{xc} where as Fock operator of Hartree-Fock equation 1.9 has only exchange part (K_{ij})) and (b) for large system also density depends only on three variables. Although DFT has advantages the crucial quantities E_{xc} and υ_{xc} decide the liability of the results. The E_{xc} is an unknown term and is approximated. E_{xc} is usually divided into separate parts, referred to as the exchange and correlation parts given as

$$\begin{aligned} E_{xc} &= (T[\rho] - T_s[\rho]) + (V_{ee}[\rho] - V_{coul}[\rho]) \\ E_{xc} &= E_x + E_c \end{aligned} \tag{1.45}$$

Various approximations are used to obtain the E_{xc}. These include local density, gradient density and hybrid methods.

1.3.4.1 Local Density Approximation (LDA)

This is the simplest approximation proposed by Hohenberg and Kohn.[19a] LDA assumes that density can be treated locally as a uniform electron gas where the density is a slowly varying function with position. The E_{xc} in LDA is given as

$$E_{xc}^{LDA} = \int \rho(r)\varepsilon_{xc}(\rho)dr \tag{1.46}$$

$$\upsilon_{xc} = \varepsilon_{xc}[\rho(r) + \rho(r)\frac{\partial \varepsilon_{xc}[\rho]}{\partial \rho}] \tag{1.47}$$

where $\varepsilon_{xc}[\rho]$ is the exchange correlation energy per electron in a homogenous electron gas with electron density ρ. In more general case where α and β densities are not equal then LDA is replaced by local spin density approximation (LSDA). This method is not very successful in calculating systems involving weak intermolecular interactions and van der Waals attractions.

1.3.4.2 Generalized Gradient Approximation (GGA)

In this method exchange and correlation energy are dependent both on electron density $\rho(r)$ and its gradient ∇_{ρ}.[19a,b] They are useful when electron density varies with the position. General form of the GGA is

$$E_x^{GGA}[\rho^{\alpha},\rho^{\beta}] = \int f[\rho^{\alpha}(r),\rho^{\beta}(r),\nabla_{\rho^{\alpha}(r)},\nabla_{\rho^{\beta}(r)}]dr \tag{1.48}$$

These functionals tend to improve the total and atomic energies in comparison to LDA. The GGA functional used in this thesis is Becke- Perdew functional.[21a,b]

1.3.4.3 Hybrid Functionals

These functionals include the mixture of HF and DFT exchange along with DFT correlation. As the exchange part can be completely evaluated by HF method. The general form is

$$E_{xc}^{Hybrid} = c_{HF}E_x^{HF} + c_{DFT}E_{xc}^{DFT} \tag{1.49}$$

The most common hybrid functional used in the thesis (chapter1 and 2) is B3LYP which uses a combination of the three-parameter Becke exchange functional[21c] along with the Lee-Yang-Parr non local correlation functionals.[21d] Its functional form is

$$E_{xc}^{B3LYP} = (1-c_0-c_x)E_x^{LSDA} + c_0E_x^{HF} + c_xE_x^{B88} + (1-c_c)E_c^{VWN} + c_cE_c^{LYP} \tag{1.50}$$

where c's are constants (c_o= 0.20, c_x = 0.72 and c_c = 0.81) to give best fit experimental atomization energies.

1.3.5 Composite Methods

This method uses a series of *ab initio* calculations with added empirical corrections. The Gaussian (Gn)[22a] and complete basis set (CBS)[22b] are two such composite methods. These methods are used to accurately determine the thermodynamic data such as enthalpy of formation, molecular atomization energies etc. These methods are also limited for smaller size systems. In this thesis G3B3 composite method was used.[22c] The total energy at 0 K is

$$E_{G3B3} = E_{MP4/6-31G(d)} + E_{QCI} + E_{Plus} + E_{2df,p} + E_{G3large} + E_{SO} + E_{Hlc} + E_{ZPE} \tag{1.51}$$

where $E_{QCI} = E_{QCISD(T)/6-31G(d)} - E_{MP4/6-31G(d)}$ (1.52)

$$E_{plus} = E_{MP4/6-31+G(d)} - E_{MP4/6-31G(d)} \tag{1.53}$$

$$E_{2df,p} = E_{MP4/6-31G(2df,p)} - E_{MP4/6-31G(d)} \tag{1.54}$$

$$E_{G3Large} = E_{MP4(Full)/G3Large} - E_{MP4/6-31G(2df,p)} - E_{MP4/6-31G(d)} \quad (1.55)$$

For molecules $E_{Hlc} = -An_{\beta} - B(n_{\alpha} - n_{\beta})$ and atoms $E_{Hlc} = -Cn_{\beta} - D(n_{\alpha} - n_{\beta})$ (1.56)

where n_{α} and n_{β} represent the number of α and β electrons, *A* and *C* represent the corrections for paired valence electrons and *B* and *D* is for unpaired valence electrons. E_{SO} = spin orbit correction for atomic species, E_{ZPE} = zero point energy correction.

The steps involved for calculating E_{G3B3} energy are (a) Equilibrium structure is initially obtained from hybrid B3LYP level with 6-31G(d) basis, and the zero point energy correction (E_{zpe}) is obtained from the frequency calculation at same level of theory. (b) Equilibrium geometry is refined further at higher level such as MP4(full)/6-31G(d) to obtain all the electron correlation energies. (c) A series of single point energy calculations are carried out at higher levels and energy is modified by series of corrections as E_{QCI} (for correlation effects beyond fourth order of perturbation theory), E_{plus} (for the correction of diffuse functions of basis set), $E_{2df,p}$ (for the correction of polarization functions), $E_{G3large}$ (for some core polarizations as well as multiple set of valence polarizations). (d) High level corrections are added to account for the remaining deficiencies in energy calculations. All electronic structure calculations are done using Gaussian[23] and ADF[24] program packages.

1.4 Molecular Mechanical Methods

Molecular mechanical methods use classical mechanics to model and predict the properties of molecular systems. In this method electrons are not explicitly examined, but are assumed to find an optimal distribution about the nuclei. The energy of the system is calculated as a function of the nuclear positions only. It considers atoms as spheres or balls those have net charges and bonds as springs. The interactions that held together this

collection of particles in molecule are simple harmonic forces. The sum of energies arising from various types of the forces that are described in terms of individual potential functions gives the molecular potential energy (V) or steric energy of molecule. Another fundamental assumption in this method is transferability of the functional form and parameters, since molecules tend to be composed of units which are structurally similar in different molecules. The method is invariably used to carry out calculations on systems containing large number of atoms such as biomolecules, polymers etc in predicting structural properties.[2,25]

1.4.1 Force Fields

A force field is used to describe the potential energy of a system of particles (typically but not necessarily atoms). The functional form of a force field[2] is

$$V(R^N) = \sum_{bonds}^{N} \frac{k_i}{2}(l_i - l_{i,0})^2 + \sum_{angles}^{N} \frac{k_i}{2}(\theta_i - \theta_{i,0})^2 + \sum_{torsions}^{N} \frac{V_n}{2}(1 + \cos(n\omega - \gamma))$$
$$+ \sum_{cross}^{N} \frac{k_i}{2}(\theta_i - \theta_{i,0})\left[(l_1 - l_{1,0}) + (l_2 - l_{2,0})\right] +$$
$$+ \sum_{i=1}^{N}\sum_{j=i+1}^{N}\left(4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}}\right) \qquad (1.57)$$

where V denotes the potential energy which is a function of the positions (R) of N particles. The first term in equation 1.57 is bond stretching energy, which represents the interaction between pairs of bonded atoms modeled by a harmonic potential where bond length l_i deviates from the reference (equilibrium) value $l_{i,0}$. The second term is the bond bending energy which is the summation of valence angles in the molecule modeled using a harmonic potential. The third term is a torsional potential that models how the energy changes when bond rotates. These three terms describe a single aspect of molecular shape

and are called as valence terms. The fourth term is the bond-stretch cross term that describes how the equilibrium bond lengths tend to shift as bond angles change. The last term in equation 1.57 represents the non bonded valence term. This is calculated between all pairs of atoms (*i* and *j*) in different molecules or in the same molecule separate by at least three bonds. This term is modeled using a Lennard-Jones potential for van der Waals interactions and Coulombic potential term for electrostatic interactions. Since the constants (k_i, V_n and σ_{ij}) of equation 1.57 are parameterized using the spectroscopic data or quantum mechanical calculations they are also called as empirical force field methods. MM calculations needs initial atomic coordinates and atom types. Atom type gives the information of atom's atomic number and molecular environment of each atom in a molecule.

In general force fields are of two types (a) Class I or harmonic and (b) Class II. Class I force fields are usually applicable for larger systems such as DNA and proteins (AMBER, CHARMM, GROMOS, OPLS). It includes the harmonic potential for the stretch and bending energy (valence terms), Lennard-Jones potential for van der Waals energy. It has no cross terms. On the other hand Class II force fields (MM2, MM3, CFF etc) are applicable for small or medium sized organic systems. It includes a number of cross terms, at least as cubic expansions of stretch and bending energy and exponential potential for van der Waals energy. Force fields can be united-atom or all-atom type. In united atom hydrogens are not treated explicitly and thus save computational time.

Various force fields that are in use are MM2, MM3, CHARMM, AMBER, ESFF, CFF, GROMOS, CVFF, Tripos, OPLS, MMFF, UFF etc. They basically differ in (a) functional form of each term, (b) number of cross terms used and (c) the data used to

parameterize the constants. The performance of force field is more sensitive to non-bonded and torsional terms. The parameters of one type of force field are not usually transferable to another type. AMBER and UFF force fields are used in chapter 4 and Tripos force field is used in chapter 5. Assisted Model Building with Energy Refinement (AMBER)[26a] is specifically parameterized for proteins and nucleic acids. It is name of a force field and molecular mechanics program also. It uses only five bonding (valence) and non-bonding terms with sophisticated electrostatic treatment and excludes cross terms. Universal Force Field (UFF)[26b] is applicable for all elements in the periodic table but most widely used for systems containing inorganic elements. It uses four valence terms and excludes electrostatic term. The number of parameters are relatively small in UFF hence cannot achieve accuracy compared to other highly parameterized force fields as MMFF94. The Tripos force field (SYBYL)[26c] is applicable to small organic molecules and proteins and is used for drug design. It uses four valence terms and an electrostatic term.

The least computational cost of molecular mechanics allows its use in procedures such as molecular dynamics, conformational energy searching, and docking that require large numbers of energy evaluations. The following sections discusses about molecular dynamics and docking aspects that were used in chapter 4 and 5.

1.4.2 Molecular Dynamics

It is a simulation of the time-dependent behavior of atomic and molecular system such as vibrational or Brownian motion. It provides details of system changes from one configuration or conformation to another. This method is generally used to investigate structure, dynamics and thermodynamics of the systems with a feasible time of

computation. To compute the time evolution of the system, we need to solve the time dependent Schrödinger equation for the N-particle wave function (1.1) or approximate the system as classical particles and use classical mechanics. Finding the wave function for the system more than 4 or 5 atoms is extremely difficult using time dependent Schrödinger equation and hence is not applicable for larger systems. On the other hand molecular mechanical methods are efficiently used to understand the complex and dynamic phenomenon that occur in biological process such as protein stability, conformational changes, protein folding, ion transport in membranes etc.

The molecular dynamic (MD) simulations[27] are based on Newton's second law, F=ma where F is the force exerted on the particle, 'm' is the mass and 'a' is the acceleration of the particle. Integration of the equations of motion then yields a trajectory that describes the positions, velocities and acceleration of particles that vary with time. The macroscopic observables such as pressure, energy, heat capacities, diffusion coefficients, radial distribution functions etc are determined from this trajectory using statistical mechanics. Unlike Monte Carlo method (another simulation technique), MD is deterministic method and once the positions and velocities of each atom are known, the state of the system can be predicted at any time in the past or the future. For the calculation of trajectory one only needs initial positions of the atoms and initial distribution of velocities. The initial velocities are often chosen randomly from a Maxwell-Boltzmann or Gaussian distribution. For a given temperature the probability that an atom i has a velocity v_x in the x direction at a temperature T is

$$p(v_{ix}) = \left(\frac{M_i}{2\pi k_B T} \right)^{\frac{1}{2}} \mathrm{e}^{-\left[\frac{1 M_i v_{ix}^{\,2}}{2 k_B T} \right]} \tag{1.58}$$

Further, velocities and accelerations are determined by the gradient of the potential energy (V) as

$$F = m\frac{d^2x}{dt^2} = \nabla_i V \tag{1.59}$$

The potential energy is a function of atomic positions (3N) for all the atoms in the system. There is no analytical solution to the equation of motion due to complicated nature of this function and are solved numerically. The numerical algorithms that integrate the equations of motion are Verlet, Leap-Frog, Velocity-Verlet, Predictor-corrector and Beeman's.[2a] The characteristics of the integrator are (a) time reversibility and (b) phase space volume should be preserved. All these algorithms assume that the positions, velocities and accelerations can be approximated by Taylor series expansion. Among the four, best is Velocity-Verlet algorithm as it needs less disk storage and there is no compromise in precision, CPU time. The system positions, velocities and accelerations are obtained from this algorithm as

$$r(t+\delta t) = r(t) + v(t)\delta t + \frac{1}{2}a(t)\delta t^2 \tag{1.60}$$

$$v(t+\delta t) = v(t) + \frac{1}{2}\left[a(t) + a(t+\delta t)\right]\delta t \tag{1.61}$$

Computing forces is the most time consuming process in MD. The calculation of the non-bonded interactions especially electrostatic interaction play significant role on the results. The three methods to compute long range coulomb forces are Ewald summation; Particle Mesh Ewald and Cell multipole.[2a] Depending on the evaluation of observables or property there are different types of ensembles used in MD calculations. An ensemble is a collection of all possible systems which have different microscopic states but have an

identical macroscopic or thermodynamic state. The different ensembles are (a) Micro canonical ensemble (NVE) corresponds to isolate system where number of atoms, volume and energy are fixed. (b) Canonical ensemble (NVT) where number of atoms, volume and temperature are fixed. (c) Isobaric-Isothermal ensemble (NPT) where number of atoms, temperature and pressure are fixed. (d) Grand Canonical ensemble (μVT) where chemical potential μ, temperature and volume are fixed. For maintenance of fixed pressure various types of barostat (Berendsen and Nosé-Hoover) and for fixed temperature thermostats (Anderson, Berendsen etc) are used. The molecular dynamics calculations mainly depends on force field functional form, evaluation of long range forces, time step etc. Time step should be one order of magnitude less than the time scale of the shortest motion of system. AMBER, CHARMM, GROMOS are some of the MD programs. AMBER9[28] was used for analyzing the average number of water molecules that are surrounding the Phosphodiesterase 4 (PDE4) enzyme (chapter 4).

1.4.3 Docking

It is the process of finding the best fit of two molecules in 3 dimensional space.[29] This method can be applied when X-ray structure or theoretical model is known for the molecule. Hence is known as a structure based design method. It helps to discover or design molecules that can interact with biochemical targets such as proteins or DNA. Based on the type of input molecules docking is of two types. Docking of two macromolecules like proteins or DNA-protein is macromolecular docking. Because of large degrees of freedom the energetically favorable complexes are evaluated based on geometric properties such as shape complementarity. Other type is small molecule (ligand or substrate) docked to macromolecule (DNA or protein). This substantially

differs from macromolecular docking in the fact that conformational flexibility of the ligand molecule is of importance along with geometrical properties to decide the low-energy complexes. Thus, depending on the flexibility of ligand molecular docking can be rigid or flexible type. The steps involved in docking are (a) conformational search (posing) and (b) scoring the obtained conformations which finally gives a rational idea of best fit of molecule.

There are several types of conformational searching methods[2,29] used in docking programs. They are (a) geometric search method such as descriptor matching, grid search, fragment based type and (b) energy search method such as simulation annealing or MD. They can also be divided as (a) systematic search which is deterministic, exhaustive and extensive and (b) stochastic search which is random, non deterministic (simulation annealing and evolutionary algorithms) method. Descriptor based method analyze the receptor for regions of likely complementarity (eg: DOCK). It is a very fast method and provides a reasonable sampling of particular region of receptor site. Grid search method samples six degrees of freedom of orientation space (soft docking). They find the neighborhood of the correct solution that cannot be guaranteed with discrete sampling method. The step size used in the search determines the time of search and accuracy of result. Fragment–joining method identifies regions of high complementarity by docking functional groups (as fragments) into receptors (eg: FLEX). It is an incremental construction algorithm which is deterministic and exhaustive. Thus overcomes the rigid ligand issues and also when ligand has more number of flexible or rotable bonds.

In energy search methods the configurational and conformational aspects of docking are taken care because there are multiple minima and complex topography of molecular potential surfaces. Simulation annealing developed by Goodsell and Olson uses Metropolis algorithm to find the low energy complexes (eg: AUTODOCK).[30] It is a quite efficient one in considering the number of degrees of freedom. Evolutionary algorithms are genetic and Lamarckian genetic algorithms and are very much popularized in the recent times. This method mimics the evolutionary process where the individuals are considered as configurations in search space and fitness function decides which individuals survive and produce off springs. In this thesis Lamarckian genetic algorithm[31] that is available in AUTODOCK[30] was used (chapter 5). It is very successful in reproducing the co crystal structures that have many rotable bonds.

After conformational sampling, ranking them with respect to binding energy to a given protein is critical. The function that ranks the conformations and gives the binding affinity estimation is scoring. The prediction of the best fit of molecule (true positives) in the receptor pocket depends mainly on scoring function. There are three types of scoring functions used in several docking programs. They are (a) force field, (b) empirical and (c) knowledge based scoring.[2a,29] Force Field scoring can be directly applied to protein-ligand complexes. It is very good in selecting docking modes but difficult to compare different molecules on this basis. In force field scoring function entropy and solvation effects are not considered. Empirical scoring function approximates the binding affinity as a sum of weighted interactions (hydrogen-bonding, hydrophobic, ionic interactions and binding affinity) that contribute to binding effect. These functions are calibrated with a set of experimental binding affinities obtained from the known protein-ligand complexes.

Inconsistency of this data will cause problems. While knowledge based scoring functions is based on inverse formulation of Boltzmann law. It is used to derive the sets of atom-pair potentials by favoring preferred contacts and penalizing repulsive interactions.

Even though molecular mechanical methods are applied widely to understand for the complex process in biomolecules, they are mainly limited in understanding the electronic structural changes (bond breaking, bond forming, charge transfer and electronic excitations etc) and in application to metallo-proteins due to the lack of effective parameters.

1.5 Quantum Mechanical/Molecular Mechanical Methods

The fundamental idea in this method is that to divide a large condensed phase system into two regions as QM and MM.[32] The reactive chemical event is contained within the QM region, while the surrounding condensed phase is modeled by MM. The Hamiltonian and energy of the whole system are given as

$$H = H_{QM} + H_{MM} + H_{QM/MM} \tag{1.62}$$

$$E = E_{QM} + E_{MM} + E_{QM/MM} \tag{1.63}$$

There are many levels of accuracy used for the QM region such as DFT or *ab initio* HF method when there are small number of atoms and semi empirical methods such as PM3 or AMI or empirical valence bond (EVB) for considerable large number of atoms. On the other hand empirical Hamiltonian can be any typical force fields such as AMBER, GROMOS, CHARMm and UFF etc.[32b]

The key aspects of QM/MM method is the interactions between the QM and MM regions. There are two types of interactions[32] (a) In order to account the influence of surrounding solvent on the enzyme of QM region (or ligand in protein active site),

electrostatic and van der Waals interactions of QM/MM interface must be included and (b) In enzymatic reaction of the large biomolecules where QM/MM interface involves covalent bonds needs special treatment. The capable way of handling both bonding (bond stretching, bond bending and internal rotation) and non bonded interactions (electrostatic and van der Waals interactions) limits the success of this method.

The electrostatic interaction which is key element in coupling is treated in two ways. They are mechanical embedding and electrostatic embedding. In mechanical embedding QM region is calculated quantum mechanically in absence of MM region and treats the interactions of QM/MM interface at MM level. Hence it is simpler but less accurate due to the lack of accurate set of electrostatic MM parameters for atoms in QM region and ignoring potential perturbation of the electronic structure of QM region by charge distribution of MM region. In electrostatic or electronic or electric embedding the QM region is calculated quantum mechanically in presence of MM region and treats the interactions of QM/MM interface by including certain one electron Hamiltonian terms. As a result the electronic structure of the QM region adjusts to the charge distribution in the MM region. It is more accurate and computationally expensive. The limitation is constructing appropriate one electron term.[32]

There are two most successful ways to treat the QM/MM interface when a covalent bond is involved. They are link atom and frozen orbital approaches.[31] In link atom approach, the link atom is used to saturate the dangling bond of the QM/MM interface. This link atom is usually a hydrogen atoms or a parameterized atom that mimics the properties of the original bond. They are treated explicitly in QM calculation but do not interact with the MM atoms. It is widely used but introducing the artificial link

atoms increases extra degrees of freedom and also the QM/MM energy is more complicated. In frozen orbital approach, localized frozen molecular orbitals are used at QM/MM interface. The localized orbitals are determined by calculations on small model compounds and assumed to be transferable. This is more fundamental and accurate than link atom approach as the frozen orbitals provide the QM description of the charge distribution at QM/MM interface and are better parameterized.

The QM/MM methods are not only quite successful in explaining the complex enzymatic reactions in biomolecules but are also useful in understanding the substitution and solvation effects in large inorganic molecules.[33] A ONIOM[34] two layer QM/MM method was used to understand the solvation effect in the PDE4 enzyme (chapter 4).

1.6 A Brief Outline of the Chapters

The electronic structure and bonding aspects of clusters such as boradiphosphole and triple-decker sandwich complexes with a P_6 as middle ring are studied in chapter 2 and chapter 3. The chemical bonding was understood by applying isolobal analogy and electron counting rules. The isolobal analogy introduced by R. Hoffmann[35] constructs a link between inorganic (organo-metallic) and organic chemistry. Later it is extended to BH and CH fragments in carboranes, BH and Si, and P and CR fragments in phosphorus-organic compounds.[35-37] The isolobal fragments have same number, symmetry, energy and occupancy of frontier orbitals which leads to existence of similar structural analogs. Application of the CH and P isolobal connection to the cyclopentadienyl anion (Cp^-) has extended the utility of Cp^- type ligands. The various *neutral* and *monoanionic* phospholes (E*t*$Bu_2C_2P_2$ where E= P^-, As^-, Sb^-, S, Se and Te) which are isostructural and isoelectronic

to cyclopentadienyl anion (Cp^-) ligand are of considerable importance because of their ability to stabilize monovalent metals.

However, *cationic* ($Et Bu_2C_2P_2$ where E = P^+, As^+ Sb^+) and *neutral* phospholes ($Et Bu_2C_2P_2$ where E = BPh, Sn, Ge and Pb), analog of Cp^+, which are synthesized in recent years have different structures. The Cp^+ which is a 4 π electron system with 20 valence electrons (5C-C, 5C-H bonds) can exist in triplet D_{5h} or vinylcyclopropenyl-type structure. All heterophospholes except $PhBtBu_2C_2P_2$ which has a tricyclopentane-type structure exist in *nido*-type geometry. To understand the structural anomaly a potential energy surface of *neutral* and *dianionic* boradiphosphole ($R'BC_2P_2R_2$) were explored in the second chapter.

Optimization of classical and non-classical *neutral* and *dianionic* $BC_2P_2H_3$ isomers at the hybrid DFT (B3LYP/6-311++G (d,p)) and composite (G3B3) methods gave a variety of geometric and positional isomers as minima. Vinylcyclopropenyl-type structure was found to be lowest energy structure for $BC_2P_2H_3$ at B3LYP/6-311++G (d,p) level of theory. However, at the G3B3 method (that includes correlation and basis set effects) two more low-energy isomers are obtained for $BC_2P_2H_3$ (atricyclopentane-, a *nido*-type). This shows that *neutral* $BC_2P_2H_3$ isomers in comparison to the isolobal Cp^+ and $H_nC_nP_{5-n}^+$ isomers has a competition between 2 π delocalized vinyl cyclopropenyl- and cluster-type structures (*nido* and tricyclopentane). Substitution of H on C by *t*Bu and H on B by Ph in $BC_2P_2H_3$ showed that lowest energy isomer to be experimentally known tricyclopentane structure rather than *nido*-type as observed in $tBu_2C_2P_3^+$. The lower preference for *nido* boradiphosphole was due to the poor overlap of BH capping orbital with phosphorus $3p_z$ orbitals of basal ring. This makes the bond-stretch isomer

tricyclopentane-type structure of lower energy. On the other hand $BC_2P_2H_3^{2-}$ at both levels and with similar substituents (C-*t*Bu and B-Ph) favors cyclic planar geometry as lowest energy structure. The factors that affect stability of 3- 4- 5- membered ring and acyclic geometrical and positional isomers of *neutral* and *dianionic* $BC_2P_2H_3$ isomers are (i) relative bond strengths (ii) availability of electrons for the empty 2p-boron orbital, and (iii) steric effects of the *t*Bu groups in the $HBC_2P_2tBu_2$ systems.[38]

In the third chapter structural diversity that is present in P_6 middle ring of triple-decker sandwich ($CpMP_6MCp$) clusters were studied. The P_6 ring existed as (a) symmetrical and planar for M=Mo, W where valence electron count [VEC] is 28, (b) distorted with four long and two short P-P bonds and yet planar for M=V (26 VEC), (c) distorted with two long and four short P-P bonds and yet planar for M=Nb (26 VEC) and (d) nonplanar chair like conformation for M=Ti (24 VEC). The traditional valence electron count (VEC) for $CpMP_6MCp$ includes 10 π electrons from 2 Cp, 6 π electrons from P_6 and valence electrons of two oxidized metal atoms. We have applied *mno* rule, which states the number of electron pairs required for stable condensed polyhedral structure is $m + n + o$, where m is the number of polyhedra, n is the number of vertices and o is the number of single atom bridges. Although this was devised primarily for applications in condensed polyhedral boranes,[39a,b] it is also applicable to metallocenes and multiple-decker sandwiches.[39c] According to *mno* rule $CpMP_6MCp$ needs 25 skeletal electron pairs or 50 electrons (m=3, n=18, o=2 plus two for the two nido arrangements) or 18 VEC to be stable. So the structures with 28 VEC or 60 skeletal electrons (if all VE of the metal and σ electrons that is 20 from ten C-C bonds of the two Cps, 12 from 6 P-P bonds are included) have 10 electrons more than required by *mno* rule. The nature of the

interactions exhibited by these extra electrons and how they affect the overall structure were investigated by studying structure and bonding of $CpMP_6MCp$ triple-decker sandwich complexes, ranging from 18–28 valence electrons (VE) with M = Sc, Y, Ti, Zr, Hf, V, Nb, Ta, Cr, Mo, and W. To model 18 and 20 VEC triple-decker complexes a $C_3B_3H_6$ and $P_3B_3H_3$ cationic and anionic middle rings are used. A density functional theory (DFT) and hybrid HF-DFT methods were employed.

The distortion of P_6 middle ring and other finer structural details in 28, 26, 24 and 22 VEC was controlled by filling of additional electrons to varied type of molecular orbitals. In 28 VE complexes complete filling of $2a^*$ and $2b^*$ orbitals lead to planar symmetrical P_6 middle ring, while the occupancy in either $2a^*$ or $2b^*$ alone explains the in-plane distortions (asymmetric) in 26 VE complexes. In comparison to 28 VE complexes, the puckering of P_6 middle ring in 24 VE complexes was due to the greater stabilization of 5a and the extra stabilization of the +4 oxidation state of Ti. The quintet state of 22 VE complexes was planar as $2a^*$ and $2b^*$ is half filled. Similar geometrical and bonding patterns observed between $CpScP_6ScCp$ and $C_2P_3H_2ScC_3P_3H_3ScC_2P_3H_2$ support the carbon–phosphorus analogy further. The modeled 20 VE systems ($CpScC_3B_3H_6ScCp^-$ and $CpScP_3B_3H_3ScCp^-$) with 52 skeletal electrons (20 VE) have in plane distortion of the middle rings. The modeled 18 VE systems ($CpScC_3B_3H_6ScCp^+$ and $CpScP_3B_3H_3ScCp^+$) with 50 skeletal electrons as stipulated by the *mno* rule were planar middle P_6 ring triple-decker sandwich complexes.[40]

The next two chapters discusses about molecular insights to the inhibitor selectivity of the phosphodiesterase enzyme. Designing small molecules (inhibitors) that bind specifically to a family of enzyme (target) is an important step in drug discovery

development. Phosphodiesterases (PDEs) are metalloenzymes which hydrolyze phosphodiesterase bond of cyclic adenosine 3',5'-monophosphate (cAMP) and cyclic guanosine 3',5'-monophosphate (cGMP) into the corresponding 5'-nucleotides (AMP and GMP) in various cells. They play a vital role in the regulation of various physiological functions. Among the 12 PDE isozymes, cAMP specific enzymes PDE4 and PDE7 were studied. PDE4 is a keen target for the inflammatory diseases such as asthma, chronic obstructive pulmonary disease (COPD), allergic rhinitis, type II diabetes and rheumatoid arthritis. However, development of PDE4 drugs is hampered due to the side effects such as nausea, vomiting, dyspepsia, headache and emesis. The several strategies to overcome the side effects are (a) selectivity towards low affinity form over low affinity rolipram binding site, (b) targeting broader PDE family, (c) low blood brain barrier drugs, (d) disease activated drugs and (e) subtype selectivity of PDE4. PDE4 is coded by four genes named as A, B, C and D and are called as subtypes. Chapter 4 is about how to achieve PDE4 subtype selectivity. For this 35 PDE4B and 4D X-ray crystal structures were investigated. A characteristic secondary structural difference (3_{10}-helix versus turn) was found in M-loop region which is near to the active site pocket. Twelve of PDE4B have a 3_{10}-helix and the four PDE4B and nineteen PDE4D strucutres have a turn in the M-loop region. An experimental variable pH seems to be controlling this structural variation rather than the sequential differences in the PDE4.

At lower pH there is a possibility of protonation, thermodynamics of protonation and deprotonation of the aspartic acid, cysteine side chains and amide bonds of M-loop (SPMCD) residues were calculated. The pentapeptide residue (SPMCD) models were investigated in the gas phase and in the explicit solvent using ONIOM method at

B3LYP/6-31+G* and B3LYP/6-31+G*:UFF levels of theory respectively. The isodesmic equations of the various protonation states show that the turn containing structure is thermodynamically more stable when proline or cysteine is protonated. The preference for the turn structure on protonation, (pH =6.5-7.5) is due to an increase in the number of the hydrogen bonding and electrostatic interactions gained by surrounding environment such as adjacent residues and solvent molecules. The average number of solvent molecules (H_2O) available in the 3_{10}-helix and turn containing structures is also analyzed by molecular dynamics (MD) simulation with explicit water for 10ns under NPT conditions.[41] Further analysis of molecular properties such as hydrophobicity and MOLCAD surface for residues of the M-loop region have shown that larger hydrophobic group consisting ligands can be better accommodated in the Q2 pocket of PDE4B.[42] Thus the variations in local (physiological) pH at the point of interaction and hydrophobic groups of Q2 pocket can be added parameters in optimizing the specificity of PDE4 inhibitors.

Chapter 5 is focused on dual selectivity of cAMP specific PDE4 and PDE7 enzymes (targeting broader family). The reason is both of them are abundant in inflammatory cells and are considered as therapeutic targets for treating various inflammatory diseases with minimal side effects. To understand the basis of inhibitor selectivity X-ray structures of PDE4 and PDE7 were analyzed systematically. We have found that some crucial residues are altered in the Q1 and Q2 pockets of active site of these proteins. Later the substrate cAMP, nonselective isobutylmethyl xanthine (IBMX), PDE4 specific rolipram and PDE7 specific BRL-50481 inhibitors were docked to the co-crystal structures of PDE4 and PDE7 using Lamarckian genetic algorithm in

AUTODOCK. The docked energies obtained were well correlated to its specificity of binding. The important interactions with substrate cAMP are conserved in both enzymes indicating its selectivity to catalyze cAMP. The small difference in docking energies of rolipram analogs and spiroquinazolinones cannot explain the selectivity as external conditions, water and peripheral residues can also influence binding. In the case of a PDE4-specific inhibitor such as rolipram, the lower affinity towards PDE7 is due to the variation of two residues in the Q2 pocket (a) Ser^{442}/Ile^{412} residue causes steric repulsions with cyclopentoxy group of rolipram (b) Met^{411}/Thr^{381} residue reduces the hydrophobic interactions with the rolipram.

These alterations of residues also reduced the size of Q2 pocket in PDE7 compared to PDE4. On the other hand variation of Tyr^{403}/Ser^{373} residue in the Q1 pocket of PDE7 affects the hydrogen bonding with Gln^{413} residue and increases the size of pocket. Further the influence of Q1 and Q2 pocket residues in inhibitor selectivity was studied by docking the less hydrophobic substituents containing ligands such as modified rolipram (where cylcopentoxy group is replaced by CH_2CH_2OH) and pyrazole analogs to PDE4 and PDE7. They all indicate an improved binding towards the PDE7 and comparable lowest docked energies with respect to the PDE4. Thus size and nature of Q1 and Q2 pockets play crucial role for the inhibitor binding to PDE4 and PDE7.

1.7 References

1. (a) Jensen, F. *Introduction to Computational Chemistry*, John Wiley & Sons, Inc., New York, **1999**. (b) Foresman, J. B.; Frisch, A. *Exploring Chemistry with Electronic Structure Methods,* Gaussian Inc., Pittsburgh, USA, **1996**. (c) Young, D. *Computational Chemistry: A Practical Guide for Applying Techniques to Real*

World Problems, John Wiley & Sons, Inc., New York, **2001**. (d) Schaeffer III, H. F. *Electronic Structure of Atoms and Molecules*, Addison-Wesley, Massachusetts, USA, **1972**. (e) Clark, T. A *Handbook of Computational Chemistry*, John Wiley & Sons, Inc., New York, **1985**. (f) Lewars E. G. *Computational Chemistry: Introduction to the Theory and Applications of Molecular and Quantum Mechanics*, Kluwer Academic Publishers, Massachusetts, USA, **2003**.

2. (a) Leach, A. R. *Molecular modeling Principles and Applications*, Pearson Education Limited, **1996**. (b) Hinchliffe, A. *Molecular modeling for Beginners*, John Wiley & Sons, Inc., New York, **2003**.
3. (a) http://www.ccdc.cam.ac.uk/products/csd/ (b) http://www.rcsb.org (c) http://ndbserver.rutgers.edu
4. (a) Levine, I. N. *Quantum Chemistry*, 5th Edition, Prentice Hall, NJ, **2000**. (b) Prasad, R. K. *Quantum Chemistry*, New Age International Limited, New Delhi, **2006**. (c) McQuarrie, D. A. *Quantum Chemistry*, Oxford University Press, California, USA, **1983**. (d) Chandra, A. K. *Introductory Quantum Chemistry*, Tata Mc-Graw Hill Publishing Co., New Delhi, **198**8. (e) Lowe, J. P. *Quantum Chemistry*, Academic Press, New York, **1978**. (f) Pople, J. A. (Nobel lecture) *Angew. Chem. Int. Ed*., **1999**, *38*, 1894-1902. (g) Barden, C. J.; Schaefer III, H. F. *Pure. Appl. Chem*., **2000**, *72*, 1405-1423.
5. Schrödinger, E. *Ann. Physik*., **1926**, 79, (new*384)*, 361-376.
6. Born, M.; Oppenheimer, J. R. *Ann. Physik*., **1927**, *84*, (new *389)*, 457-484.
7. (a) Hartree, D. R. *Proc. Cambridge Phil. Soc*., **1928**, *24*, 89-110. (b) Fock, V. A. *Z. Phys*., **1930**, *61*, 126-148. (c) Fock, V. A. *Z. Phys.,* **1930**, *62*, 795-805.

8. Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab-initio Molecular Orbital Theory*, John Wiley & Sons, Inc., New York, **1986**.

9. (a) Roothaan, C. C. J. *Rev. Mod. Phys*., **1951**, *23*, 69-89. (b) Roothaan, C. C. J. *Rev. Mod. Phys*., **1960**, *32*, 179-185.

10. (a) Slater, J. C. *Phys. Rev*., **1930**, *36*, 57-64. (b) Boys, S. F. *Proc. Roy. Soc., (London)*, **1950**, *A200*, 542-554. (c) Frenking, G.; Seijo, L.; Bohme, M.; Dapprich, S.; Ehlers, A. W.; Jonas, V.; Neuhaus, A.; Otto, M.; Stegmann, R.; Veldkamp, A.; Vyboishchikov, S. F.; *Rev. Comput. Chem*., **1996**, *8*, 63-68. (d) Ihm, J.; Zunger, A.; Cohen, M. L. *J. Phys*., **1979**, *C12*, 4409-4422. (e) Hay P. J.; Wadt, W. R. *J. Chem. Phys*., **1985**, *82*, 270-283. (f) Liu, B.; McLean, A. D. *J. Chem. Phys*., **1973**, *59*, 4557-4558. (g) Boys, S. F.; Bernardi, F. *Mol. Phys.,* **1970**, *19*, 553-566.

11. (a) Sadlej, J. *Semi-empirical Methods of Quantum Chemistry*, Ellis Harwood, Chichester, **1985**. (b) Zerner, M. C. *Rev. Comput. Chem*., **1991**, *2*, 313-320. (c) Kohanoff, J. *Electronic Structure Calculations for Solids and Molecule*, Cambridge University Press, New York, **2006**. (d) Pople, J. A.; Beveridge, D. L. *Approximate Molecular Orbital Theory*, Mc-Graw Hill Publishing Co., NewYork, **1970**.

12. (a) Hückel, E. *Z. Phys*., **1931**, *70*, 204-286. (b) Hoffmann, R. *J. Chem. Phys*., **1963**, *39*, 1397-1412.

13. (a) Pariser, R.; Parr, R. G. *J. Chem. Phys*., **1953**, *21*, 466-471. (b) Pople, J. A. *Trans. Faraday Soc*., **1953**, *49*, 1375-1385. (c) Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc*., **1977**, *99*, 4899-4907. (d) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E.

F.; Stewart, J. J. P. *J. Am. Chem. Soc*., **1985**, *107*, 3902-3909. (e) Stewart, J. J. P. *J. Comput. Chem*., **1989**, *10*, 209-220.

14. (a) Hurley, A. C. *Electron Correlation in Small Molecules*, Academic Press, London, **1977**. (b) Wilson, S. *Electron Correlation in Molecules*, Clarendon Press, Oxford, **1984**. (c) Raghavachari, K.; Anderson, J. B. *J. Phys. Chem*., **1996**, *100*, 12960-12973.
15. (a) Moller, C.; Plesset, M. S. *Phys. Rev*., **1934**, *46*, 618-622. (b) Bartlett R. J. *Ann. Rev. Phys. Chem*., **1981**, *32*, 359-401.
16. (a) Hylleraas, E. A. *Z. Physik*., **1928**, *48*, 469-494. (b) Hylleraas, E. A. *Z. Physik*., **1929**, *54*, 347-366.
17. (a) Čižek, J. *J. Chem. Phys*., **1966**, *45*, 4256-4266. (b) Čižek, J. *Int. J. Quantum Chem*., **1977**, *5*, 359-379. (c) Bartlett, R. J.; Musial, M. *Rev. of Modern Phys*., **2007**, *79*, 291-352.
18. Pople, J. A.; Head-Gordan, M.; Raghavachari K. *J. Chem. Phys*., **1987**, *87*, 5968-5975.
19. (a) Parr, R. G.; Yang, W. *Density Functional Theory of Atoms and Molecules,* Oxford University Press, Oxford, **1989**. (b) Dreisler, R. M.; Gross, E. K. V. *Density Functional Theory: An Approach to the Quantum Many-body Problem*, Springer-Verlag, Berlin, **1990**. (c) Baerends, E. J.; Gritsenko, O. V. *J. Phys. Chem. A*., **1997**, *101*, 5383-5402. (d) Geerlings, P.; De Proft, F.; Langenaeker, W. *Chem. Rev*., **2003**, *103*, 1793-1873. (e) Kohn, W.; Becke, A. D.; Parr, R. G. *J. Phys. Chem*., **1996**, *100*, 12974-12980.

20. (a) Hohenberg, P.; Kohn, W.; *Phys. Rev*., **1964**, *136*, B864-B871. (b) Hohenberg, P.; Kohn, W.; Sham, L. J. *Advances in Quantum Chemistry*, Vol. **21**, Academic Press, **1990**. (c) Kohn, W.; Sham, L. J. *Phys. Rev*., **1965**, *140*, A1133-A1138. (d) Kohn, W. *Rev. Mod. Phys***., 1999**, *71*, 1253-1266.

21. (a) Becke, A. D. *Phys. Rev. A*., **1988**, *38*, 3098-3100. (b) Perdew, J. P. *Phys. Rev. B*., **1986**, *33*, 8822-8824. (c) Becke, A. D. *J. Chem. Phys*., **1993**, *98*, 5648-5652. (d) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B*., **1988**, *37*, 785-789.

22. (a) Pople, J. A.; Head-Gorden, H.; Fox, D. J.; Raghavachari, K.; Curtiss. L. A. *J. Chem. Phys*., **1989**, *90*, 5622-5629. (b) Ochterski, J. W.; Petersson, G. A.; Monotgomery, J. A.; *J. Chem. Phys*., **1996**, *104*, 2598-2619. (c) Baboul, A. G.; Curtiss, L. A.; Redfern, P. C. Raghavachari, K. *J. Chem. Phys*., **1999**, *110*, 7650-7657.

23. Frisch, M. J. *et al.* **Gaussian 03**, Revision C.02, D.03 version, Gaussian, Inc., Wallingford CT.

24. Baerends, E. J. *et al.* **ADF2003.01**, SCM, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands.

25. (a) Burkert, U.; Allinger, N. L. *Molecular mechanics*, ACS Monograph, Washington, D. C, **1982**, *177*. (b) Allinger, N. L. *J. Am. Chem. Soc*., **1977**, *99*, 8127-8134. (c) Schlecht, M. F. *Molecular Modeling on the PC*, John Wiley & Sons, Inc., New York, **1998**. (d) Rappè, A. K.; Casewit, C.J. *Molecular Mechanics across Chemistry*, University Science Books, Sausalito, **1997**.

26. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz Jr., K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem.*

Soc., **1995**, *117*, 5179-5197. (b) Rappé, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard III, W. A.; Skiff, W. M. *J. Am. Chem. Soc*., **1992**, *114*, 10024-10035. (c) Clark, M.; Cramer III, R. D.; Opdenbosch, N. V. *J. Comput. Chem*., **1989**, *10*, 982-1012.

27. (a) Karplus, M.; McCammon, J. A.; *Nature Struct. Bio*., **2002**, *9*, 646-652. (b) Gunsteren, W. F.v. *et al., Angew. chem. Int. Ed*., **2006**, *45*, 4064-4092. (c) Adcock, S. A.; McCammon, J. A. *Chem. Rev*., **2006**, *106*, 1589-1615.

28. Case, D. A.; Darden, T. A.; Cheatham, III, T. E.; Simmerling, C.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Pearlman, D. A.; Crowley, M.; Walker, R. C.; Zhang, W.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Wong, K. F.; Paesani, F.; Wu, X.; Brozell, S.; Tsui, V.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Mathews, D. H.; Schafmeister, C.; Ross, W. S.; Kollman. P. A. **AMBER9**, 9th ed.; University of California: San Francisco, CA.

29. (a) Mannhold, R.; Kubinyi, H.; Timmerman, H. *Methods and Principles in Medicinal Chemistry*, John Wiley & Sons, Inc., New York, **2001**. (b) Halperin, I.; Ma, B.; Wolfson, H.; Nussinov, R. *Proteins: Structure, Function, and Genetics*, **2002**, *47*, 409-449. (c) Schneider, G.; Böhm, H-J. *Drug Disc. Today,* **2002**, *7*, 64-70. (d) Kuntz, I. W.; Meng, E. C.; Shoichet, B. K. *Acc. Chem. Res.,* **1994**, *27*, 117-123.

30. **AUTODOCK,** 10550, North Torrey Pines Road, La Jolla, CA, 92037-1000 USA.

31. (a) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. *J. Comput. Chem*., **1998**, *19*, 1639-1662. (b) Hu, X.; Balaz, S.; and Shelver, W. H. *J. Mol. Graph. Model.*, **2004**, *22*, 293-307.

32. (a) Thiel, W. *Curr. Opin. in Chem. Bio*., **2007**, *11*, 182-187. (b) Friesner, R. A.; *Drug Disc. Today: Technologies*, **2004**, *1*, 253-260. (c) Monard, G.; Merz Jr, K. M. *Acc. Chem. Res*., **1999**, *32*, 904-911. (d) Lin, H.; Truhlar, D. G. *Theor. Chem. Acc*., **2007**, *117*, 185-199.

33. (a) Ryde, U., *Curr. Opin. in Chem. Bio*., **2003***, 7,* 136-142. (b) Bo, C.; Maseras, F. *J. Chem. Soc. Dalton Trans*., **2008**, 2911-2919.

34. (a) Dapprich, S.; Komáromi, I.; Suzie Byun, K.; Morokuma K.; Frisch, M. J. *J. Mol. Struct. (THEOCHEM)*, **1999**, *462*, 1-21. (b) Vreven, T.; Morokuma, K.; Farkas, Ö.; Schlegel, H. B.; Frisch M. J. *J. Comput. Chem*., **2003**, 24, 760-769. (c) Verven, T.; Byun, K. S.; Komáromi, I.; Dapprich, S.; Montgomery Jr., J. A.; Morokuma, K.; Frisch, M. J. *J. Chem. Theory Comput*., **2006**, *2*, 815-826.

35. Hoffmann, R. (Nobel Lecture) *Angew. Chem. Int. Ed*., **1982**, *21*, 711-724.

36. (a) Prasad, B. V.; Jemmis, E. D.; Tsuzuki, S.; Tanabe, K. *Proc. Ind. Acad. Sci (Chem. Sci)*, **1990**, *102*, 107-115. (b) Prasad, B. V.; Jemmis, E. D.; Tsuzuki, S.; Tanabe, K. *J. Phys. Chem*., **1990**, *94*, 5530-5535. (c) Giju, K. T.; Jemmis, E. D. *The Encyclopedia of Computational Chemisty,* John Wiley & Sons, Chichester, **1998**, *2*, 1449-1455.

37. Dillon, K. B.; Mathey, F.; Nixon, J. F.; *Phosphorus: The Carbon Copy: From Organophosphorus to Phospha-organic Chemistry*, John Wiley & Sons, Chichester, **1998**.

38. Usharani, D.; Poduska, A.; Nixon, J. F.; Jemmis, E. D. *Accepted in Chem. Euro. J.*

39. (a) Jemmis, E. D.; Balakrishnarajan, M. M.; Pancharatna, P. D. *J. Am. Chem. Soc*., **2001**, *123*, 4313-4323. (b) Jemmis, E. D.; Balakrishnarajan, M. M.; Pancharatna,

P. D. *Chem. Rev*., **2002**, *102*, 93-144. (c) Jemmis, E. D.; Jayasree, E. G.; Parameswaran, P. *Chem. Soc. Rev*., **2006**, *35*, 157-168.

40. Usharani, D.; Prasad, D. L. V. K.; Nixon, J. F.; Jemmis, E. D. *J. Comput. Chem*., **2007**, *28*, 310-319.
41. Usharani, D.; Srivani, P.; Narahari Sastry, G.; Jemmis, E. D.; *J. Chem. Theory Comput*., **2008**, *4*, 974-984.
42. Srivani, P.; Usharani, D.; Jemmis, E. D.; Narahari Sastry, G. *Curr. Pharma. Design,* **2008,** *14*, 3854-3872.