Chapter 1

Introduction to computational details

1.1 Introduction

Noncovalent interactions play a significant role in determining the conformation of a molecular or a cluster system. They govern the physicochemical properties of the molecule appreciably and thereby impact the biological activities of the molecules [1–3]. The effect of conformational and noncovalent interaction is gauged by computing the normal modes of vibration, dipole moment, stabilization energy, and physical properties of the molecular system [4-5].

The work embodied in this thesis work explores conformational analysis of selected alkanethiol molecules and cluster systems based on sulfur centered hydrogen bonds. Following sections introduce briefly the concepts of conformational analysis, noncovalent interaction/bonding, and computational chemistry.

1.2 Conformational Analysis

The term 'conformation' was first introduced by the Nobel Laureate D. H. R. Barton [6]. In a very generic way, a conformer is a 3-dimensional arrangement of atoms of a molecule in the real space that shows two different structures of a molecule [7] which are also known as rotamers or conformational isomers or stereoisomers and these stereoisomers are interconvertible through rotation around a single bond. In 1933 Kuhn et al. introduced atropisomerism (means 'do not turn') term to rationalize the stereoisomers corresponding to the restricted rotations about the single bond by taking asymmetric molecular systems [8-9]. In 1936-37 Kemp et al. resolved the discrepancy in the measurement of entropy of ethane by applying the concept of restricted rotation corresponding to the carbon-carbon single bond (rotational barrier 3.00 kcal/mol) and predicted the staggered conformer at ground state instead of gauche conformer to be the most stable one [10-11]. Further, the concept of gauche and

staggered conformation was confirmed by Hessel et al. in 1947 by performing electron diffraction investigation on pyranose rings [12-13].

Conformational analysis got a fillip after the development of sophisticated spectroscopic techniques and increased capabilities of the computational chemistry, thereby making it a crucial tool for the investigation of complex reactions in organo-catalysis, metal-catalysis, development of newer catalysts etc. Elucidation of conformers using frequency domain spectroscopy has been traditionally well exploited. In this regards, many experimental techniques like X-ray [14], electron diffraction [15], infrared [16], Raman [17], UV [18], NMR [19], microwave spectra [20], photoelectron spectroscopy [21], supersonic molecular jet spectroscopy [22], optical rotatory dispersion (ORD) [23], and CD measurements [24] etc. are available, through which conformational study was explored. With the advent of femtosecond pulsed lasers, conformational analysis along with the associated dynamics involved also got unraveled in a seamless manner. Apart from experimental methods, computational chemistry tools have also been a very powerful method for studying the conformation of molecules [25].

Chemical aspects and electronic properties of a molecular system is intensely associated with the conformation that it attains. Therefore, the conformational analysis of a molecule helps in understanding the properties of molecules, like physical, chemical, biological, medicinal, and pharmaceutical etc. Several factors decide the conformation of a molecule and amongst them, the noncovalent interactions play a very significant role in making a molecule attain a particular conformation.

1.3 Non-covalent interaction

Non-covalent interactions (NC- interactions) are ubiquitous in solids, liquids, gases and defines the 3-dimensional crystal structure of the organic & organometallic moieties. The NC interactions play a crucial role in the stability of the molecules and molecular clusters and the function of the biomolecules and materials. It is a common place in most of the biological processes and is considered as the basic foundation for life processes [26]. Properties of supramolecular species are strongly affected by NC interactions.

1.3.1 Types of NCI or hydrogen bond

Generally, it is categorized into the following three classes [26]:

1). Very Strong NC-interaction (15-40 kcal/mol)

2). Strong NC-interaction (4-15 kcal/mol)

3). Weak NC-interaction (< 4 kcal/mol)

This classification is based on the complex energy or stabilization energy, which was well explored by Gautam R. Desiraju and Thomas Steiner in their book, '*The Weak Hydrogen Bond, In Structural Chemistry and Biology.'* Very strong NC-bonds are formed by the active participation of the donor and acceptor atom either the NC-bond is occurring in the same molecules or different molecules. Mobility of very strong NC-bond to strong NC-bond is a kind of transition from quasi-covalent to electrostatic nature of the bond. Weak NC-bonds are also electrostatic in nature but dispersive and charge transfer contributions are the dominating components but still, they are stronger than the van der Waals interactions [26].

1.3.2 Hydrogen bonding

In the history of Hydrogen bonding, the very first report came in the year 1920 by Latimer and Rodebush [27]. A hydrogen bond is a class of NC interactions that have so many fabulous years of study. The formation of the H-bond is a directional behavior just like other chemical bonds. This behavior of the Hydrogen bond enables them for extensive involvement in the conformation, and molecular association [26]. For hydrogen bonding, Pauling's work became a landmark; Pauling in his book [28] titled, '*The nature of chemical bond*' described the hydrogen bond in this manner (Chapter 12, Page 449): *"It was recognized some decades ago that under certain conditions an atom of a hydrogen bond is attracted by rather strong forces to two atoms, instead of only one, so that it may be considered to be acting as a bond between them"* and provided justification in favor of the hydrogen bonding [29]. Thus, it was formed by the electrostatic attraction of two atoms that have high electronegativity. Based on these, it is inferred that the hydrogen bond is purely ionic in nature. Pauling talked about the existence of only conventional hydrogen bonds [28]. Further, Pimentel and McClellan provided a broader aspect of hydrogen bonding in the year 1960 [30]. According to them, *"A hydrogen bond exists between a functional group A-H and an atom or a group of atoms B in the same or a different molecule when:*

1. *There is evidence of bond formation (association or chelation)*
2. *There is evidence that this new bond linking A-H and B involves the hydrogen atom already bonded to A."*

The hydrogen bonding definition of Pimentel and McClellan was not based on the concept of electronegativity and polarization of charge [30]. Most updated definition of hydrogen bonding came in the year 2011 from the IUPAC committee [31], according to whom, *"The hydrogen bond is an attractive interaction between a hydrogen atom from a molecule or a molecular fragment X-H in which X is more electronegative than H, and an atom or a group of atoms in the same or different molecule, in which there is evidence of bond formation."*

H-bond is the backbone of the living organism because H-bond plays a vital role in the aggregation of the monomer units of the amino acids, enabling the formation of the proteins and enzymes [32]. Based on the donor-acceptor atoms of the hydrogen bonds it is classified into two categories (conventional & unconventional H-bond) as shown in Fig.1.1

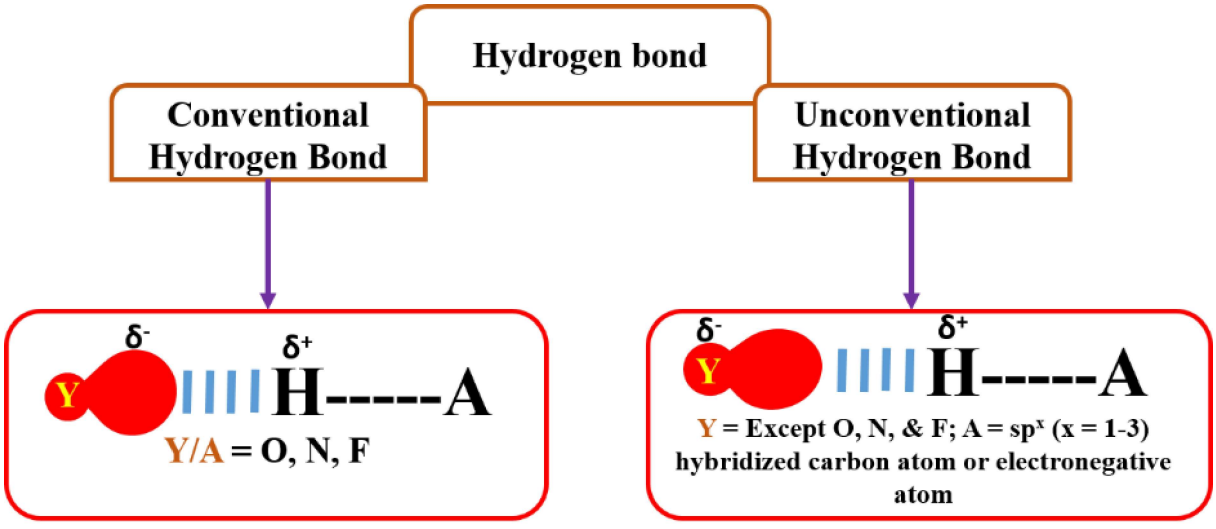

Figure 1.1 A schematic demonstration of different classes of Hydrogen bonding.

If a donor or acceptor atom or both atoms involved in H-bonding comprise of highly electronegative atoms like F, O, and N atom then it is referred to as a conventional Hydrogen

bond (CHB) e.g., F--H-F, O--H-O, N--H-N, N--H-O, whose bond energy ranges from 20 to 50 kcal/mol [26, 33-34]. On the other hand, if the donor or acceptor or both atoms involved in the H-bond are less electronegative atoms then it is referred to as an unconventional hydrogen bond (UHB) e.g., S--H-S, S--H-C, S--H-O, etc., whose bond energy is 1-20 kcal/mol [33-34]. Among the CHB and UHB it is the former that was explored relatively in a wider manner compared to the latter. Hydrogen bonds are generally described in terms of the d, D, r, θ, and θ' as shown in Fig. 1.2.

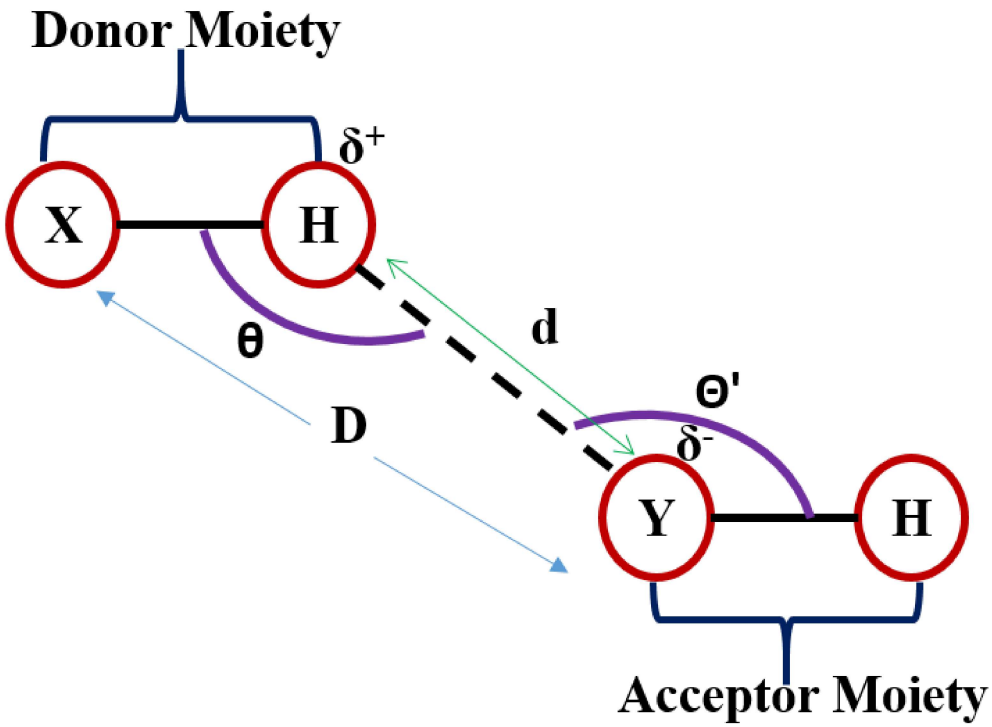

Figure 1.2 Pictorial presentation of the parameter that affects hydrogen bonding.

Hydrogen bond will be most stable when the value of θ is 180º [26]. In this interaction it is generally an involvement of the hydrogen atom and an electronegative atom. So it is essentially an interaction between acidic and basic moieties but it is not an acid-base reaction. In hydrogen bonding, $H^{\delta+}$ from X-H to Y-H group is only partially transferred but $H^{\delta+}$ is still connected with the X atom covalently.

1.3.3 Sulfur centered hydrogen bonds (SCHBs)

Sulfur also forms hydrogen bonds albeit weaker when compared to the oxygen based hydrogen bonds. This behavior of sulfur appears because of its lower electronegativity (2.58) and size. SCHBs are in a way UHB and just like the CHB, they also significantly affect crystal structure, solvent-solute interaction and stability of the molecular systems [35-36]. Early reports on hydrogen bond did not recognize the SCHBs as a valid H-bond until the work reported by Gordy et al., who confirmed the formation of hydrogen bond in sulfur-containing molecules in bases pyridine, α-picoline, and dibenzylamine [37]. Gordy et al. report opened the possibility of the exploration of the SCHBs. Consequent to this a series of studies were carried out that confirmed the involvement of the sulfur atom in the hydrogen bonding [26, 38-39]. Hoyer first detected the intramolecular hydrogen bonding in sulfur-containing compounds, N-acyl-thioureas in 1981 [40]. Biswal et al decoded the biological activity of methionine, tyrosine through experiments which were in favor of SCHBs. They utilized model molecules like dimethyl sulfide, phenol, p-cresol, 2-napththol and carried out laser induced fluorescence in supersonic jet spectroscopic technique [35]. Further, Thorley et al. studied S-F and S-O based hydrogen bond through computational tools on organic polymeric chain

system and confirmed the electrostatic nature of these hydrogen bonds [41]. Another report by Biswal et al. for SCHBs came in 2009, where they studied complexes of H_2S with indole and 3-methyl indole. Through this study they unambiguously attributed the stability of the modeled complexes to the S-H---π interactions [42]. In 2015 Biswal et al. reviewed SCHBs that were studied via laser spectroscopic and quantum-chemical calculations. They highlighted the nature and strength of donor-acceptor capability in isolated binary complexes of SCHBs [36]. Various investigations, both by experiment and computation, were done to unravel the nature, strength, and directionality of the SCHBs [43-50].

1.4 Computational and theoretical methods

Molecular properties of molecules can be accessed through their electronic structures which is rendered possible by solving the Schrodinger wave equation (for many-electron systems) which forms the core of quantum chemical calculations. Various types of quantum chemical methods are utilized such as semi-empirical methods (AM1, PM6, etc.), *ab initio* methods (HF, CCSD, CCSD(T), etc.), and density functional (BLYP, B3LYP, M06, etc.) methods. Out of these quantum chemical methods, the *ab initio* method and density functional theory are the most preferred and used methods for the elucidation of the modeled (i.e. guess) molecular system [4].

1.4.1 Born-Oppenheimer approximation

Born-Oppenheimer approximation came in 1927 after the combined efforts of Max Born and J. Robert-Oppenheimer. This is a basic concept that helps to resolve the complexity of the quantum chemical calculation during the assignment of quantum states of molecules. It

is known that an electron is 1000 times lighter than a nucleus. Consequent to this large difference in their weights, the nuclei are considered almost stationary compared to electrons in a molecular system under this approximation. Thus the Hamiltonian (for multielectronic and two nuclei) is free from the contribution of kinetic energy of nuclei, and is given as

$$\hat{H}_{elec}(r,R) = \frac{1}{2}\sum_i \nabla_i^2 - \left(\sum_i \sum_I \frac{Z_I}{|R_I - r_i|} + \sum_i \sum_{j<i} \frac{1}{|r_i - r_j|} + \sum_i \sum_{J<I} \frac{Z_I Z_J}{|R_I - R_J|}\right) \quad (1.1)$$

where ∇ is the Laplacian operator, Z_I and Z_J are the atomic number of 'I' and 'J' respectively, R_I and R_J is the position of nuclear I and J respectively, r_i is the position of electrons. The first term of the above equation represents the kinetic energy term of the electron and the second term contains mutual repulsion for each other and attraction of nuclei [51].

1.4.2 Density functional theory

Basic foundation of this theory was the work of Thomas and Fermi [52,53] and further followed by Thomas-Fermi-Dirac [54–56] followed by the seminal and revolutionizing works of P. Hohenberg and W. Kohn in 1964 [57] and of W. Kohn and L. J. Sham in 1965 [58].

1.4.2.1 Hohenberg-Kohn Theorems

Theorem: 1 The external potential $V_{ext}(r)$, and the total energy, is a unique functional of the electron density [n(r)] i.e. [n(r)] determine the exact number of electrons and total energy of the modeled system E[n(r)].

Thus, the energy functional E[n(r)] can be written in the form of the external potential V_{ext} (r) and electron density [n(r)]

$$E[n(r)] = \int n(r) V_{ext}(r) dr + F[n(r)] \quad (1.2)$$

Where F[n(r)] is a universal functional of the electron density [n(r)].

Hamiltonian of the system corresponding to energy functional can be written so that the expectation value gives the ground state energy in the following manner:

$$E[n(r)] = < \psi \hat{H} \psi > \quad (1.3)$$

Hamiltonian can be written as $\hat{H} = F + V_{ext}$

Where $F = T + V_{ee}$; where $V_{ee} = J + nonclassical\ term$

The potential energy due to repulsion between electrons V_{ee} consists of classical repulsion term J and non-classical exchange-correlation term. The electron operator F is the same for all electrons (N) so the $\hat{H}$ is defined by N and V_{ext} (r).

Theorem 2: The ground state energy can be obtained variationally i.e. the density that minimizes the total energy is the exact ground state i.e.

$$E_0 \leq E[n(r)] \quad (1.4)$$

Where E[n(r)] is the energy functional, and it is very similar to the variational principle of the wave function i.e. $E_0 \leq E(\psi')$.

1.4.2.2 Kohn-Sham Equation

Kohn-Sham proposed an indirect approach to treat many-body interacting systems in terms of a different auxiliary non-interacting one-body system [58]. This equation is based on the basic ansatz that the ground-state electron density of an interacting system will be as ground-state electron density of a non-interacting system with an effective potential V_{eff}. So, the Kohn-Sham Hamiltonian is

$$H_{KS} = -\frac{1}{2}\nabla^2 + V_{eff} \quad (1.5)$$

Kinetic energy for a non-interacting system of interest is expressed as

$$T[\rho] = \sum_i^N \langle \psi_i \left| -\frac{1}{2}\nabla_i^2 \right| \psi_i \rangle \qquad (1.6)$$

The above shows that the kinetic energy is dependent on the electron density of the non-interacting system and it is expressed as

$$\rho(r) = \sum_i^N \sum_{r,s} |\psi_i(r,s)|^2 \qquad (1.7)$$

which is the obtained after the solution of the Schrödinger equation for non-interacting electrons moving in Kohn-Sham effective potential, V_{eff} (r), which is given as

$$V_{eff}(r) = V(r) + \frac{\delta J[\rho]}{\delta\rho(r)} + \frac{\delta E_{XC}[\rho]}{\delta\rho(r)} \qquad (1.8)$$

$$= V(r) + \int \frac{\rho(r')}{|r-r'|}\, dr' + V_{XC}(r) \qquad (1.9)$$

where V_{XC} (r) is the exchange-correlation potential,

$$V_{XC}(r) = \frac{\delta E_{XC}[\rho]}{\delta\rho(r)} \qquad (1.10)$$

So, for a given V_{eff} (r), the electron density can be obtained by solving the N one-electron Schrödinger equation,

$$\left[-\frac{1}{2}\nabla^2 + V_{eff}(r)\right]\psi_i = \varepsilon_i \psi_i \qquad (1.11)$$

These equations are solved iteratively, till self-consistency is achieved. The total energy can be calculated from the following equation

$$E = \sum_i^N \varepsilon_i - \frac{1}{2}\int \frac{\rho(r)\rho(r')}{|r-r'|} dr dr' + E_{XC}[\rho] - \int V_{XC}(r)\rho(r) dr \qquad (1.12)$$

Likewise, the Hartree-Fock equation and Kohn-Sham equation provide a one-body equation corresponding to the many-body systems but, from a principle point of view, the

Kohn-Sham approach is quite different from the Hartree-Fock theory by rationalizing the effect of exchange-correlation on electrons. Furthermore, in the Kohn-Sham method, the exchange and correlation functional were not exactly incorporated. Resulting, various approximation methods were developed for the exchange-correlation functional because, exchange-correlation functional, $E_{XC}[\rho]$ in the Kohn-Sham equation plays a vital role in the accuracy of the results. Further, the exchange-correlation energy term $E_{XC}[\rho]$ is divided into two terms i.e. exchange contribution $E_X[\rho]$ and correlation contribution $E_C[\rho]$.

$$E_{XC}[\rho] = E_X[\rho] + E_C[\rho] \qquad (1.13)$$

So, in the DFT method, three exchange-correlation approximations were incorporated; that is

(i). Local Density Approximation (LDA)

(ii). Generalized Gradient Approximation (GGA)

(iii). Hybrid Functional

In LDA approximation, the local electron density is only used to define the approximate exchange-correlation functional. Examples of LDA methods are, SVWN, PVWN5, and CAPZ. In GGA-based methods, the gradient of the electron density is added in addition to local electron density to define the approximate exchange-correlation functional. Examples of the GGA method are BLYP, PW91, and B88. In hybrid-based functional, some percentage of the Hartree-Fock method is added in the exchange energy part along with the DFT method. Mixing of Hartree-Fock exchange provides a simple scheme for improving many molecular properties, like atomization energies, structural parameters, and vibrational frequencies which are poorly

described with regular exchange-correlation functional. Examples of Hybrid DFT methods are B3LYP, and M06-2X [4].

1.4.3 Møller-Plesset perturbation theory

Perturbation theory is in practice since the early days of quantum chemistry to get the electronic structure of atoms and molecules. In 1934, Møller and Plesset [59] expressed second-order perturbation theory into the Hartree-Fock method for getting the corrected electron pair correlation energy of the modeled system. Quantitatively, Møller-Plesset perturbation theory adds higher excitations to Hartree-Fock theory as a non-iterative correction. Perturbation theory is based upon dividing the Hamiltonian into two parts:

$$H = H^{(0)} + \lambda V \qquad (1.14)$$

where $H^{(0)}$ is the zero-order Hamiltonian and λV is the perturbation applied to $H^{(0)}$ (applied correction is small as compared to the $H^{(0)}$). Zeroth order Hamiltonian $H^{(0)}$ is expressed as the sum of the Fock operators in the following manner:

$$\hat{\mathrm{H}}^{(0)} = \sum_p \hat{\mathrm{F}}(p) = \sum_p \hat{\mathrm{h}}(p) + \sum_{p,i}[\hat{\mathrm{J}}_i(p) - \hat{\mathrm{K}}_i(p)] \qquad (1.15)$$

with $\hat{\mathrm{h}}(p)$ being the one-electron operator associated with the kinetic and nucleus-electron attraction part of electron p, $\hat{\mathrm{J}}_i(p)$, and $\hat{\mathrm{K}}_i(p)$ being the Coulomb and exchange operator that form the Hartree-Fock potential and are expressed in terms of spin orbitals ψ_i:

$$\hat{\mathrm{J}}_i(1)\psi_j(1) = [\int dr_2 \psi_i^*(2)\frac{1}{r_{12}}\psi_i(2)]\psi_j(1) \qquad (1.16)$$

$$\hat{K}_i(1)\psi_j(1) = [\int dr_2 \psi_i^*(2)\frac{1}{r_{12}}\psi_j(2)]\psi_i(1) \qquad (1.17)$$

(1 and 2 denote positions r_1 and r_2 of electrons 1 and 2)

The assumption that V is a small perturbation to $H^{(0)}$ suggests that the perturbed wavefunction and energy can be expressed as a power series in V. The usual way to do so in terms of the parameter λ is:

$$\psi = \psi^{(0)} + \lambda\psi^{(1)} + \lambda^2\psi^{(2)} + \lambda^3\psi^{(3)} + \cdots \qquad (1.18)$$

$$E = E^{(0)} + \lambda E^{(1)} + \lambda^2 E^{(2)} + \lambda^3 E^{(3)} + \cdots \qquad (1.19)$$

The perturbed wave function and energy are substituted back into the Schrödinger equation

$$H^{(0)} + \lambda V\left(\psi^{(0)} + \lambda\psi^{(1)} + \cdots\right) = (E^{(0)} + \lambda E^{(1)} + + \cdots)(\psi^{(0)} + \lambda\psi^{(1)} + \cdots) \qquad (1.20)$$

As zeroth-order wavefunctions, the HF wavefunction is taken $\varphi^{(0)} = \varphi^{HF}$. The corresponding eigenvalue is given by the sum of orbital energies:

$$E^{(0)} = E^{orb} = \sum_i^{occ} \varepsilon_i \qquad (1.21)$$

Hence, the HF wave function spans the model space P. The Q-space is spanned by the substitution function φ_s which are generated by single (S), double (D), triple (T), etc. excitations of electrons from occupied spin orbitals ψ_i, ψ_j, ……to virtual orbital ψ_a, ψ_b, … (ψ_p, ψ_q, … denote general spin-orbital). Functions $\varphi_s = \varphi_i^a, \varphi_{ij}^{ab}, \varphi_{ijk}^{abc}, \ldots$ are orthogonal to φ^{HF} which simplifies the calculation of different energy corrections. The MP perturbation operator is given by

$$\hat{\mathrm{V}} = \hat{\mathrm{H}} - \hat{\mathrm{H}}^{(0)} = \sum_{p\geq q} \frac{1}{r_{pq}} - \sum_{p,i}[\hat{\mathrm{J}}_i(p) - \hat{\mathrm{K}}_i(p)] \qquad (1.22)$$

Applying the perturbation operator, the first-order correction energy $E^{(0)}$ results as

$$E_{MP}^1 = < \varphi^{(0)}\left|\hat{\mathrm{V}}\right|\varphi^{(0)} > = \mathrm{V}_{00} = -\frac{1}{2}\sum_{ij} < ij||ij > \qquad (1.23)$$

Where double bar integral is an anti-symmetrized two-electron integral of the general type

$$< pq|rs > = \iint \psi_p^*(1)\psi_q^*(2)\frac{1}{r_{12}}[\psi_r(1)\psi_s(2) - \psi_s(1)\psi_r(2)]d\tau_1 d\tau_2 \qquad (1.24)$$

Where τ denotes integration over space and spin coordinate of the spin orbitals ψ.

Thus the HF energy is identical to the MP1 energy

$$E(HF) = E(MP1) = \sum_i^{occ} \varepsilon_i - \frac{1}{2}\sum_{ij}^{occ} < ij||ij > \qquad (1.25)$$

i.e. HF is correct up to first-order MPPT, which was first observed by Møller and Plesset. This holds also for dipole moment, electron density, and other one-electron properties and is known as the MP theorem.

The MP2 correlation energy correction is given by the following equation

$$E_{MP}^{(2)} = \sum_{s>0} \frac{V_{0S} - V_{S0}}{E_0 - E_S} \qquad (1.26)$$

where S substitutions ($\varphi = \varphi_i^a$) do not contribute because $< \varphi^{(0)}|\hat{V}|\varphi_i^a > = 0$as a consequence of the Brillouin theorem. Finite matrix elements V_{0s} are only obtained for D excitations ($\varphi_s = \varphi_{ij}^{ab}$) (T, Q, etc. excitations are excluded because of the Slater–Condon rules).

When applying Slater–Condon rules, the second-order correction takes the following form

$$E_{MP}^{(2)} = \frac{1}{4}\sum_{ij}^{occ}\sum_{ab}^{vir} < ij||ab > a_{ij}^{ab} \qquad (1.27)$$

with

$$a_{ij}^{ab} = (\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b)^{-1} < ab||ij > \qquad (1.28)$$

Being the amplitudes of the D excitations.

1.4.4 Coupled cluster methods

Coupled cluster (CC) is a numerical technique used for describing many-body systems. Its most common use is as one of several post-Hartree–Fock ab initio quantum chemistry methods in the field of computational chemistry, but it is also used in nuclear physics. Coupled cluster essentially takes the basic Hartree–Fock molecular orbital method and constructs multi-electron wave functions using the exponential cluster operator to account for electron correlation. Some of the most accurate calculations for small to medium-sized molecules use this method. The method was initially developed by Fritz Coester [60] and Hermann Kümmel [61] in the 1950s for studying nuclear physics phenomena but became more frequently used when in 1966 JiříČížek [62] and later together with Josef Paldus [63] reformulated the method for electron correlation in atoms and molecules. It is now one of the most prevalent methods in quantum chemistry that includes electronic correlation. CC theory is simply the perturbative variant of the Many Electron Theory (MET) of Oktay Sinanoğlu, which is the exact (and variational) solution of the many-electron problem, so it was also called "Coupled Pair MET (CPMET)". Therefore, the CC method is quite simpler and more accurate from a calculation point of view; so, the CC method is the best variant of MET and gives highly accurate results in comparison to experiments [64].

Coupled-cluster theory provides the exact solution to the time-independent Schrödinger equation

$$H|\Psi> = E|\Psi> \qquad (1.29)$$

The wave function of the coupled-cluster theory is written as an exponential ansatz:

$$|\Psi> = e^{T} |\Phi> \qquad (1.30)$$

where, $|\Phi>$ the reference wave function, which is typically a Slater determinant constructed from Hartree–Fock molecular orbitals, though other wave functions such as configuration interaction, multi-configurational self-consistent field, or Brueckner orbitals can also be used. T is the cluster operator which, when acting on $|\Phi>$, produces a linear combination of excited determinants from the reference wave function.

The cluster operator is written in the form,

$$T = T_1 + T_2 + T_3 + \ldots \qquad (1.31)$$

where T_1 is the operator of all single excitations, T_2 is the operator of all double excitations, and so forth. In the formalism of second quantization, these excitation operators are expressed as

$$T_1 = \sum_i \sum_a t_a^i \hat{\alpha}^a \hat{\alpha}_i \qquad (1.32)$$

$$T_2 = \frac{1}{4} \sum_{i,j} \sum_{a,b} t_{ab}^{ij} \hat{\alpha}^a \hat{\alpha}^b \hat{\alpha}_j \hat{\alpha}_i \qquad (1.33)$$

and for the general n-fold cluster operator

$$T_n = \frac{1}{(n!)^2} \sum_{i_1,i_2,i_3 \ldots\ldots i_n} \sum_{a_1,a_2,\ldots\ldots,a_n} t_{a_1,a_2,\ldots\ldots,a_n}^{i_1,i_2,i_3 \ldots\ldots i_n} \hat{\alpha}^{a_1} \hat{\alpha}^{a_2} \ldots \ldots \hat{\alpha}^{a_n} \hat{\alpha}_{i_n} \ldots \hat{\alpha}_{i_2} \hat{\alpha}_{i_1} \qquad (1.34)$$

The exponential operator e^T may be expanded as a Taylor series and if we consider only the T_1 and T_2 cluster operators of T, we can write:

$$e^T = 1 + T + \frac{1}{2!} T^2 + \cdots = 1 + T_1 + T_2 + \frac{1}{2} T_1^2 + T_1 T_2 + \frac{1}{2} T_2^2 + \cdots \qquad (1.35)$$

The classification of traditional coupled-cluster methods rests on the highest number of excitations allowed in the definition of *T*. The abbreviations for coupled-cluster methods usually begin with the letters "CC" (for coupled cluster) followed by

1. S – for single excitations (shortened to *singles* in coupled-cluster terminology)
2. D – for double excitations (*doubles*)
3. T – for triple excitations (*triples*)
4. Q – for quadruple excitations (*quadruples*)

The complexity of equations and the corresponding computer codes, as well as the cost of the computation increases sharply with the highest level of excitation. For many applications CCSD, while relatively inexpensive, does not provide sufficient accuracy except for the smallest systems (approximately 2 to 4 electrons), and often an approximate treatment of triples is needed. The most well-known coupled-cluster method that provides an estimate of connected triples is CCSD(T), which provides a good description of closed-shell molecules near the equilibrium geometry.

CCSD (single and double excitations) is convenient but the addition of disconnected triples (CCSDT) is very expensive. A perturbative estimate of the effect of triple excitations defines the CCSD(T) method, is sometimes called the “gold standard” [4].

1.4.5 Basis set

Electronic properties of molecular systems are stored in the form the molecular orbitals and these are constructed through mathematical functions [4]. Each molecular orbitals (ψ_i) constructed by the linear combinations of the n basis functions (φ_μ)

$$\psi_i = \sum_{\mu=1}^{n} C_{\mu i}\varphi_\mu \qquad (1.36)$$

where $C_{\mu i}$ is the molecular orbital expansion coefficients or molecular orbital coefficient.

Generally, basis functions are localized on the center of atoms and are often called atomic basis functions. The accuracy of the computed results depends upon the type of the basis function and the size of the basis sets. Generally, a basis function is a kind of mathematical function and the basis set is a set of basis functions that contains a specific set of parameters corresponding to the basis sets. In very general terms, a basis set is a set of functions (i.e. atomic-like functions) over which the molecular orbitals are created via a linear combination of the atomic orbitals (LCAO). Thus the success of the theoretical and computational calculations depends upon the appropriate basis set [4]. A finite set of basis functions is utilized to perform theoretical and computational calculations. So, for quantum chemical calculations, various basis sets are developed but only two basis sets are more common that is

(i) Minimal basis set

(ii) Split-valence basis set

Minimal basis sets are composed of the minimum number of basis functions for all electrons of the atoms in the molecular system and these are not much flexible to represent the distribution of electrons completely in atomic orbital (i.e. non-spherical electron distribution in the orbitals). So, to dissolve the overcoming of minimal basis sets, split-valence basis sets were derived which are based on the doubling of all basis functions. So, the Split-valence basis set is a combination of double basis sets i.e. split valence basis set (inner and valence) and the linear combination of the same type of orbitals with a different effective charge. The multiple

basis functions corresponding to each and every valence atomic orbitals are represented as double, triple, quadruple-zeta type basis etc. [4].

1.4.6 Interaction energy calculation

The interaction energy of a molecular cluster is obtained by the following equation

$$\Delta E_{int}(AB) = E_{AB}^{AB}(AB) - E_{A}^{A}(A) - E_{B}^{B}(B) \qquad (1.37)$$

where A and B are monomers; AB is the molecular cluster; $E_{AB}^{AB}(AB)$ represents the energy of the cluster that is obtained at a specific level of theory; $E_{A}^{A}(A)$ and $E_{B}^{B}(B)$ is the energy of the monomer A and B at the same level of theory that is used to calculate the energy of the molecular cluster [65].

During the formation of a cluster, both monomers approach each other; during this process monomers use basis functions of each other up to some extent that appears in the form of additional stabilization in the overall cluster. This additional stabilization is referred to as the basis set superposition which is corrected using [66] the following equation:

$$\Delta E_{Int}^{CP}(AB) = E_{AB}^{AB}(AB) - E_{A}^{AB}(A) - E_{B}^{AB}(B) \qquad (1.38)$$

here $\Delta E_{Int}^{CP}(AB)$ is the counterpoise (CP) corrected interaction energy of the molecular cluster; $E_{A}^{AB}(A)$ and $E_{B}^{AB}(B)$ are the energy of the monomer which is calculated at the same of level of theory that is used for the calculation of the molecular cluster.

1.4.7 Complete basis set (CBS) limit extrapolation

Complete basis set (CBS) accuracy provides an alternative path for computational chemists to more reliable approximate results of infinite larger basis sets by using results of finite larger basis sets without any further computational cost. Generally, CBS limit is useful

for post Hartree-Fock method where electron correlations are taken into account and these are compute-intensive calculations. The Helgaker method is generally utilized to extrapolate correlation energy with basis sets cc-pVNZ (where N = T, Q) [67].

$$E = a + bX^{-3} \qquad (1.39)$$

where X is the cardinal number i.e., four for quadruple-zeta sets and three for triple-zeta settings etc.

1.5 Computational tools and server credentials

All calculations of this thesis were carried out by using Gaussian16 suite. All data visualization and job submissions were performed by using Gaussview interface of Gaussian 16 software package [68]. MultiWFN software [69] was used for atoms in molecules (AIM) and noncovalent interactions (NCI-plot) calculations performed with the Gaussian16 generated wave functions. With the help of ORCA 4.2 software [70] local energy decomposition (LED) calculation was carried out. All Gaussian calculations were performed at institute server (i.e. IIT (BHU) Varanasi) that have following credentials-

Architecture	x86_64
CPU op-mode(s)	32-bit, 64-bit
Byte Order	Little Endian
CPU(s)	40
On-line CPU(s) list	0-39
Thread(s) per core	2

Core(s) per socket	10
Socket(s)	2
NUMA node(s)	2
Vendor ID	GenuineIntel
CPU family	6
Model	63
Model name	Intel(R) Xeon(R) CPU E5-2660 v3 @ 2.60GHz
Stepping	2
CPU MHz	2899.914
BogoMIPS	5203.73

1.6 Molecular properties

1.6.1 Quantum theory of atoms in molecules (QTAIM) theory

The Quantum theory of atoms in molecules (QTAIM) is one of the best method for unraveling atom-atom interactions in covalent or non-covalent interactions in molecules, and molecular cluster systems. This theory is totally dependent on the electron density or charge density of the molecule for exploring the nature of bonding in molecules or molecular systems. Generally, the electron density is given by the following equation:

$$\rho(r,X) = N \int \psi * (x;X)\psi(x;X)d\tau' \qquad (1.40)$$

where N is the number of electrons, x is the electronic coordinate, X is the nuclear coordinate and $d\tau'$ is the volume element of the system which is taken into consideration.

Strength of chemical bonds is rationalized by calculating the bond critical points (BCP). According to Koch and Popelier, BCP reveals the path of the hydrogen bonds and which has electron density in the range of 0.002-0.040 a.u. and Laplacian electron density in the range of the 0.024-0.139 a.u. [71] Moreover, the topological parameters of BCPs confirm the existence of the three different kinds of hydrogen bonds. So for this, Rozas et al. [33] set three conditions based on the Laplacian electron density ($\nabla^2\rho$); for the strong and covalent nature of the hydrogen bonds $\nabla^2\rho < 0$ and H < 0, for medium and partially covalent nature of the hydrogen bond $\nabla^2\rho > 0$ and H < 0, and for weak and electrostatic nature of the hydrogen bond $\nabla^2\rho > 0$ and H > 0. Some other typical relationships of Laplacian electron density were also found with energetic topological parameters at the BCPs. An important relation of Laplacian electron density is expressed by the Virial theorem

$$\frac{1}{4}\left(\nabla^2\rho(r_{BCP})\right) = 2G\left(r_{BCP)}\right) + V\left(r_{BCP)}\right) = H(r_{BCP)} \qquad (1.41)$$

where $G\left(r_{BCP)}\right), V\left(r_{BCP)}\right), H(r_{BCP)}$ are the kinetic energy (positive quantity), potential energy (negative quantity), and total electron energy densities respectively. With help of kinetic and potential electron density, the nature of the interaction is well explored. If, the potential electron density is higher than the kinetic electron density [i.e.$|V\left(r_{BCP)}\right)| > G\left(r_{BCP)}\right)$] then the interaction shows covalent behavior and if the interaction is not as per this relation they show non-covalent behavior. Further, if the ratio of the kinetic and potential electron density i.e. $\left(-G\left(r_{BCP)}\right)/V\left(r_{BCP)}\right)\right.$ is greater than the 1, then the nature of the interaction is truly non-covalent in nature [72,73].

1.6.2 Natural bond orbital (NBO)

NBO is a mathematical algorithm developed by Weinhold et al. [74], which helps to determine the strength of the formed hydrogen bond interactions in the molecular complexes. NBO suit modifies delocalized molecular orbitals into localized natural bond orbitals (NBOs). General NBOs are localized either on one center or two centers or maybe on three centers, which is dependent on the system. The NBOs for diatomic molecules is the linear combination of the NHOs that generates two NBOs, namely bonding NBO (or donor orbital) and antibonding NBO (acceptor orbital). Out of these NBOs, bonding NBO is occupied orbital and antibonding NBO is vacant orbital and accepts electrons from the donor NBO orbitals.

Delocalization of electrons are generally responsible for the formation of the hydrogen bonds (i.e. conventional and unconventional hydrogen bond) and the strength of hydrogen bonds are measured in second-order perturbation energy $E^{(2)}_{i\to j^*}$ which is as per the pomposity of the principle of the NBO analysis [74].

$$E^{(2)}_{i\to j^*} = q_i \frac{F(i,j^*)^2}{\varepsilon_{j^*}-\varepsilon_i} \qquad (1.42)$$

where i and j^* are the donor and acceptor NBO that are responsible for the transfer of the electron density into the system and the consequence of this donation appear in the form of the strength of the interactions, q_i is the occupancy of the donor orbital, F (i, j^*) is the off-diagonal NBO Fock-Matrix element and ε_{j^*} and ε_i are diagonal elements (orbital energies). As greater is the overlapping between donor and acceptor NBOs [74]; the greater will be the value of the

second-order perturbation energy and thus stronger will be the strength of the hydrogen bonding.

1.6.3 Frontier molecular orbital (FMO)

Frontier molecular orbital is the combination of two orbitals i.e. Highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO). As per Koopmans theorem, the ionization energy (I) and electron affinity (A) can be expressed in terms of the HOMO and LUMO orbital energies respectively [75]. For qualitative measurements, HOMO (electron donor) and LUMO (electron acceptor) orbitals are responsible and make sense regarding the chemical reactivity and kinetic stability of the system. Apart from this understanding it also helps to rationalize the physical and physicochemical properties by calculating global descriptors like electrophilicity index (ω), global hardness (η), global softness, electronic chemical potential (μ), electronegativity (χ), and maximum charge transfer capability index (ΔN_{max}).

$$\text{Electrophilicity } (\omega) = \frac{\mu^2}{2\eta} \qquad (1.43)$$

$$\text{Global hardness } (\eta) = \frac{E_{LUMO} - E_{HOMO}}{2} \qquad (1.44)$$

$$\text{Electronic chemical potential } (\mu) = \frac{E_{HOMO} + E_{LUMO}}{2} \qquad (1.45)$$

$$\text{Electronegativity } (\chi) = -\mu \qquad (1.46)$$

$$\text{Global softness } (S) = \frac{1}{\eta} \qquad (1.47)$$

$$\text{Maximum charge transfer capability index } (\Delta N_{Max}) = \frac{I + A}{2(I - A)} \qquad (1.48)$$

So once energy parameters are calculated from FMOs it further helps to understand the chemical reactivity, kinetic stability, and various physicochemical properties of the system. Generally, smaller the value of the H-L gap, smaller will be the stability, and more will be the reactivity.

1.6.4 Electrostatic potential (ESP) surface

Electrostatic potential maps are a three-dimensional distribution of the charge in a molecular system that allows visualizing differently charged regions of the molecule and once the distribution of charges is known, predict the chemical behavior of the molecule. According to Politzer, the distribution of molecular electrostatic potential charges is a kind of measure of the strength of the charges, nuclei, and electrons at a specific position. The electrostatic potential at any point, V(r), is the energy required to bring a positive charge from infinity to that point. As each pseudo atom in the refined model consists of the nucleus and the electron density distribution described by the multipole expansion parameters, the electrostatic potential may be calculated by the evaluation of

$$V(r) = \sum_A \frac{Z_A}{|R_A - r|} - \int \frac{\rho(r')dr'}{|r' - r|} \qquad (1.49)$$

where Z_A is the charge on nucleus A located at a distance R_A [76]. Thus, based on the ESP charges identifying the possible sites for the electrophilic and nucleophilic interactions.

1.6.5 Mullikan charge

Mullikan atomic charge depends upon the orbital where the electrons are found. Generally, Mullikan charges are the byproduct of quantum mechanical calculations and it is found during the optimization calculation at a certain level of theory. Quantum mechanical

calculations need a density matrix ($D_{i\,j}$) and overlap matrix ($S_{i\,j}$) for getting solutions to the wave function of the corresponding system. Mullikan charges are obtained by the following equation

$$P_{i\,j} = D_{i\,j}S_{i\,j} \qquad (1.50)$$

where P is the Mullikan population matrix, $D_{i\,j}$ is the density matrix, and $S_{i\,j}$ is the overlap matrix.

Mullikan charges are extremely sensitive to the size of the basis function of the calculation and strongly affect the values of the charges on the same atom. So comparisons of Mullikan charges are always done at the same level of theory because MC shows a big difference in charge of the same atom of the molecule with different levels of theory. Apart from this behavior of MC, it is very useful to interpret the reactive behavior of the molecules that are labile towards electron-rich or electron-deficient species [77,78].

1.6.6 Non-covalent interaction (NCI)

Non-covalent interaction plot is the best way to locate the presence of weaker interaction whether it will occur within the molecule or between two different molecules. Weaker interactions are well described by Bader in his theory i.e. quantum theory of atoms in molecules (QTAIM) along with the non-covalent interaction index [79]. Non-covalent interaction is dependent on the electron density which is assigned in terms of the reduced density gradient (RDG) and it is obtained by the following equation

$$RDG = s(r) = \frac{1}{2(3\pi^2)^{1/3}} \frac{|\nabla\rho(r)|}{\rho(r)^{4/3}} \qquad (1.51)$$

where ρ(r) is the electron density; $\nabla\rho(r)$ is the first derivative of the electron density.

With the help of the NCI index, non-covalent interactions are elucidated where QTAIM based bond critical points are of no help. Thus NCI plots are specifically applied to visualize the region where non-covalent interactions take place and where RDG approaches zero at regions with low electron density [79,80].

1.6.7 Objective of Thesis

This thesis addresses the following objective for resolving ambiguity regarding the number of conformers, global minima conformer, local minima conformers, and stability of the global and local minima conformers of the molecular system via exploring inter or intra molecular hydrogen bonding.

1). To revisit the conformational analysis and structural stability conformers of ethanethiol, and propylthiol (i.e. 1-propanethiol (1P), and 2-propanethiol (2P)) because literature has ambiguity regarding to the number of conformers and dominations of global and local minima of alkanethiol. These objectives are investigated and presented in chapters 2 and 3 of this thesis.

2). Revisiting the conformational analysis along with unraveling the nature and strength of sulfur centered hydrogen bonding in homo dimeric and trimeric clusters of methanethiol molecule. The driving force is the existing ambiguity regarding the number of minimum energy conformers and their stability. This objective is resolved and the details are presented in chapter 4 of this thesis.

3). To rationalize the strength and nature of sulfur centered hydrogen bonds in hetero dimeric, trimeric and tetrameric cluster system of the methanethiol with water, and the details are given in chapter 5 of this thesis.