CHAPTER – 3
COMPUTATIONAL METHODS

In the field of research, computational chemistry and molecular modeling is used for the modeling of biomolecules, drugs, organic and inorganic molecules etc. in order to understand and predict their behavior at the molecular level. Computational studies could eliminate the time-consuming, costly and impossible experimental studies.

3.1 Quantum chemical computation

The electronic wavefunctions describe the behavior of electrons in atoms and molecules [43] and characterize all of the physical properties of the molecular system. Quantum chemistry is useful in understanding molecules at the subatomic level. Quantum chemical computations are used to calculate thermodynamic properties of gases; to interpret molecular spectra; to calculate molecular properties theoretically. The quantum mechanical based electronic structure methods such as *ab initio*, semi-empirical and density functional theory were outlined below:

3.1.1 *Ab initio* method

Ab initio is the Latin term meaning "from the beginning". *Ab initio* method is an approximate quantum mechanical method which calculates the molecular structure solving the Schrodinger equation without including any experimental data. The most common *ab initio* calculation is the Hartree–Fock (HF) calculation, in which the primary approximation is the central field approximation. This method includes the net effect of Coulombic electron-electron repulsion in the calculation. Due to the

central field approximation, the energies from HF calculations are always greater than the exact energy and tend to a limiting value called the Hartree Fock limit [44].

The wavefunction must be described by some functional form, which is only known exactly for a few one-electron systems are the second approximation in HF calculations. The functions used most often are linear combinations of Slater type orbitals (STO) or Gaussian type orbitals (GTO). The wavefunction is formed from linear combinations of basis functions. Because of this approximation, most HF calculations give a computed energy greater than Hartree Fock limit. In general, *ab initio* calculations give very good qualitative results and for smaller molecular systems, it will give increasingly accurate quantitative results too [44].

3.1.2 Density functional theory

Density functional theory (DFT) is a computational quantum mechanical modeling method. This theory has been developed more recently than other *ab initio* methods. The electronic wavefunction of an n-electron molecule depends on 3n spatial and n spin coordinates [45] and contains more information but lacks in direct physical significance. This has prompted that the energy of a molecule can be determined from the electron density instead of a wavefunction. Kohn and Sham developed a practical application of this theory and formulated a method similar in structure to the HF method. In this formulation, the electron density is expressed as a linear combination of the basis functions similar to HF orbitals. A determinant is then formed from these functions and is called Kohn-Sham orbitals. The electron density from this determinant of orbitals is used to compute the energy. A density functional is then used to obtain the energy for the electron density. A functional is a function of

a function, it is electron density here [44]. BLYP refers to the joint use of the Becke correction and the LYP correlation correction, and was one of the first DFT methods. BLYP also forms the basis of an even more accurate functional, B3LYP [46]. The molecular geometry, harmonic force field and vibrational frequencies were computed using the B3 [47] exchange functional combined with the LYP [48] correlation functional resulting in the B3LYP density functional theory method.

3.2 Gaussian basis sets

A basis set is a collection of mathematical functions (basis functions), linear combinations of which yield molecular orbitals. It is used to describe the shape of the orbitals in an atom. A linear combination of atomic orbitals (LCAO) approximation constructs the molecular orbitals. A basis set must be specified while doing *ab initio* or density functional theory calculations [44]. Most quantum chemistry programs use Gaussian-type orbitals. Gaussian functions are closely related to exponential functions, which are of the form of exact solutions to the one electron Hydrogen atom. Minimal basis sets are used for very large molecules, qualitative results, and in certain cases quantitative results [49].

3.2.1 STO–3G minimal basis set

The smallest basis sets are called the minimal basis sets. This comprises only those functions required to accommodate all the electrons of the atom, while still maintaining its overall spherical symmetry. STO–3G is the most popular minimal basis set. Linear combination of three Gaussians have been used to approximate the Slater function STO–3G mean "Slater-type orbital three Gaussians". The values of the Gaussian exponents and the linear coefficients have been determined by least squares as best fits to Slater-type functions [49].

3.2.2 6–31G and 6–311G split valence basis sets

The simplest split valence basis sets are 3-21G and 6-31G. Each core atomic orbital in the 3-21G basis set is expanded in terms of three Gaussians (hence 3), while basis functions representing inner and outer shells of valence atomic orbitals are expanded in terms of two and one Gaussians respectively (hence 21) [49]. The 6-31G basis set splits each valence orbital into two parts, an inner shell and an outer shell. The basis function of the inner part is represented by three Gaussians, and that of the outer part by one Gaussian (hence 31). The core orbitals are each represented by one basis function, each composed of six Gaussians (hence 6) [50]. Additional valence shell splitting should lead to even greater flexibility. 6-311G basis sets split the valence functions into three parts instead of two, these being written in terms of three, one and one Gaussians, respectively [49].

3.2.3 6-31G* and 6-311G** polarization basis sets

The use of polarization basis functions is indicated by an asterisk (*). The simplest polarization basis sets 6-31G* and 6-311G* are constructed from 6-31G and 6-311G respectively. The polarization basis set 6-31G* refers to a 6-31G basis set with polarization functions on the heavy (non-hydrogen) atoms. Two asterisks on 6-31G** indicate the use of polarization functions on hydrogen and helium. The 6-31G (d) and 6-31G (d,p) are synonymous with 6-31G* and 6-31G** respectively [51]. Polarization functions are used because they often result in more accurate computed geometries and vibrational frequencies. The 6-31G** basis set is identical to 6-31G*, except that it provides p-type polarization functions for hydrogen. Similarly, 6-311G** is identical to 6-311G* except for its description of hydrogen [49].

3.3 Computer Software

The quantum chemistry software used for the molecular modeling is the 2009 version of the Gaussian 09W [52] suite of program package. The output from the Gaussian is examined graphically using the graphical user interface GaussView [53]. Even large molecules can be visualized using GaussView. The natural bond orbital (NBO) analysis is done to understand various second order interactions between the filled orbitals of one subsystem and vacant orbitals of another subsystem, which is a measure of intra-molecular delocalization or hyperconjugation. The NBO calculations were performed using the NBO 3.1 [54] program as implemented in Gaussian 09W [52] program package. The transformation of force field from Cartesian to symmetry coordinate, the subsequent normal coordinate analysis, calculation of potential energy distribution (PED) and IR and Raman intensities were done using MOLVIB program written by Sundius [55, 56]. Multiple scaling has been done based on scaled quantum mechanical (SQM) force field method to overcome the systematic error caused by the vibrational anharmonicity [57]. The contributions of the transitions in the UV-visible spectrum have been assigned with the help of GaussSum 2.2 [58]. The electron localization function (ELF) analysis has been done using a topological mesh modeling software Topmod [59] and the elf plot is visualized using Molekel visualization software [60].

3.4 Optimization of geometry

Geometry optimizations usually attempt to locate minima on the potential energy surface, in order to predict the equilibrium structures of molecular system. A potential energy surface is a mathematical relationship linking molecular structure and the resultant energy [61]. The energy of the molecule changes with its structure. The

process of calculating the geometry of molecular structure with minimum energy is geometry optimization or energy minimization. Geometry optimization is an iterative process. A geometry optimization starts with an input molecular structure, geometry of which is guessed based on the similar compounds. The initial guess structure of the molecule is specified by the internal coordinates (z-matrix), in which the position of each atom is described relative to other atoms in the system. The energy and its first derivatives with respect to all geometrical coordinates like bond lengths, bond angles and dihedral angles are calculated for the guess geometry. This process continues until the lowest-energy or optimized geometry is reached [49].

3.5 Normal coordinate analysis

Normal coordinate analysis (NCA) is commonly used as an aid in the interpretation of the vibrational spectra of large molecules. Vibrational mode assignments of polyatomic molecules have been made on the basis of relative intensities, line shape and potential energy distribution obtained from normal coordinate analysis. The results of NCA are helpful in the study of vibration-rotation and electronic-vibrational interactions.

The frequency of the normal vibration is determined by the kinetic and potential energies of the system. The kinetic energy and potential energy is expressed in terms of internal coordinates. The matrix representation for kinetic energy is G and F-matrix (force constant matrix) is the matrix representation for potential energy. The atomic masses, bond distances and overall structures of a molecule is known, for which normal coordinate analysis is to be carried out [62]. From the G-matrix and F-matrix, the vibrational frequencies are calculated by solving the secular equation (3.1).

$$|GF - E\lambda| = 0 \qquad\qquad \text{.......................... (3.1)}$$

where E is the unit matrix and λ is the root of the secular polynomial [63]. NCA is a mathematical procedure that gives the normal coordinates, frequencies and force constants. In order to get meaningful results, knowledge about the vibrational force field is necessary. As the number of atoms increases, the number of force constants also grows quadratically. Hence many approximations have to be employed in the calculation of harmonic force fields even for moderately large molecules [64]. A number of programs have been developed for the calculation of force field [65-67]. Gwinn [68] developed a program for normal coordinate analysis using mass-weighted Cartesian coordinates, which removes the redundancy problems arising when internal valence coordinates are used, as in Wilson's GF-method [69]. The molecular force fields are refined by a modified least squares method, as described by Sundius [70]. MOLVIB is used for the scaling of vibrational force fields using Pulay's method [71], the basis for scaled quantum mechanical (SQM) force field [72].

The Raman activities (S_i) calculated with Gaussian 09 program have been suitably adjusted by the scaling procedure and converted to Raman intensities (I_i) using the relationship eq. (3.2) derived from the basic theory of Raman scattering [73, 74].

$$I_i = \frac{f(\upsilon_0 - \upsilon_i)^4 S_i}{\upsilon_i[1 - \exp(-hc\upsilon_i / kT)]} \qquad\qquad \text{.......................... (3.2)}$$

where υ_0 is the exciting frequency (in cm^{-1} units), υ_i is the vibrational wavenumber of the ith normal mode. h, c and k are universal constants and f is the suitably chosen common scaling factor for all the peak intensities.

The internal force constants, which are needed in the scaling, are found by the following coordinate transformation [57]. In matrix form, the relationship between the internal coordinates s_i and the Cartesian displacement coordinates x_i of a molecule can be expressed as in eq. (3.3).

$$s = Bx \qquad \qquad \dots\dots\dots\dots\dots.. (3.3)$$

where B is a rectangular matrix, because the number of internal coordinates and number of Cartesian displacement coordinates are identical. With the L-matrix, the normal modes can be conveniently characterized and it gives the transformation from normal to internal coordinates. With the aid of B-matrix and the eigenvector matrix U, the L-matrix can be calculated as in eq. (3.4).

$$L = BM^{-1/2}U \qquad \qquad \dots\dots\dots\dots\dots.. (3.4)$$

Also the normal modes can be characterized by means of the potential energy distribution (PED). From the L-matrix and F-matrix in internal coordinates, PED is calculated according to eq. (3.5).

$$P_{ij} - L_{ij}^{\ 2} F_{ij}/\lambda_i \qquad \qquad \dots\dots\dots\dots\dots.. (3.5)$$

The PED gives the fractional contribution of the diagonal matrix elements of F to the normal modes and is used for a complete symmetry classification of the normal vibrations of the molecule [64].

3.6 Natural bond orbital (NBO) Analysis

Natural bond orbital analysis is a technique for studying hybridization and covalency effects in polyatomic wave functions. NBO analysis provides an efficient

method for studying intra- and inter-molecular bonding and interaction among bonds, and also provides a convenient basis for investigating charge transfer or conjugative interaction in molecular systems. In the NBO analysis the electronic wave functions are interpreted in terms of a set of occupied Lewis-type (bond or lone pair) and a set of unoccupied non-Lewis (antibond or Rydberg) localized NBO orbitals. Delocalization of electron density (ED) between these orbitals corresponds to a stabilizing donor-acceptor interaction. The second-order perturbation theory has been employed to evaluate the stabilization energies of all possible interactions between donor and acceptor orbitals in the NBO basis [75]. The result of interaction is a loss of occupancy from the concentration of electron NBO of the idealized Lewis structure into an empty non-Lewis orbital. For each donor (i) and acceptor (j), the stabilization energy E 2 associated with the delocalization i→j is estimated as in eq. (3.6).

$$E(2) = -n_\sigma \frac{\langle \sigma |F| \sigma \rangle^2}{\varepsilon_\sigma^* - \varepsilon_\sigma} = -n_\sigma \frac{F_{ij}^2}{\Delta E} \qquad \ldots\ldots\ldots\ldots\ldots.. (3.6)$$

where $\langle \sigma |F| \sigma \rangle^2$ or F_{ij}^2 is the Fock matrix element i and j NBO orbitals, ε_σ and ε_σ^* are the energies of σ and σ* NBOs and n_σ is the population of the donar σ orbital.

Larger the E(2) value, the interaction between electron donors and electron acceptors is more intensive and greater the extent of conjugation of the whole system. The intramolecular hyperconjugative interactions are formed by the orbital overlap between π(C–C) and geminal π*(C–C) bond orbitals which results intramolecular charge transfer (ICT) causing stabilization of the system.

3.7 Frontier molecular orbital (FMO) analysis

The two molecular orbitals which play an essential role in a wide range of chemical reactions of saturated or unsaturated compounds are referred as "frontier orbitals". They are highest occupied molecular orbital and lowest unoccupied molecular orbital abbreviated frequently by HOMO and LUMO respectively [76]. Both HOMO and LUMO are the molecular orbitals that take part in chemical stability. When electrons are transferred from a molecule, it is the electrons in the HOMO that are involved, because they are the most weakly held. When electrons are transferred to a molecule, they go into the LUMO, as it is the lowest energy orbital available [77]. The narrow energy gap between the HOMO and LUMO allows the promotion of electrons from HOMO to LUMO. HOMO and LUMO energies are very popular chemical descriptors, which determine the kinetic stability, chemical reactivity, optical polarizability and chemical softness-hardness of the material.

3.8 Hyperpolarizability

Hyperpolarizability is the nonlinear optical property of a molecule and is defined as the second-order electric susceptibility per unit volume. It describes how easily the electrons can be pushed or pulled around the molecule. The total static dipole moment (μ)eq. (3.7) and average linear polarizability (α) eq. (3.8) using the x, y, z components is defined as

$$\mu = \left[\mu_x^2 + \mu_y^2 + \mu_z^2 \right]^{1/2} \qquad\qquad \ldots\ldots\ldots\ldots\ldots\ldots.. (3.7)$$

$$\langle \alpha \rangle = 1/3 \left[\alpha_{xx} + \alpha_{yy} + \alpha_{zz} \right] \qquad\qquad \ldots\ldots\ldots\ldots\ldots\ldots.. (3.8)$$

The complete equation for calculating the magnitude of first order hyperpolarizability (β)eq. (3.9) using the x, y, z components from Gaussian 09 output is given as follows [78]:

$$\beta_{tot} = \left[\beta_x^2 + \beta_y^2 + \beta_z^2 \right]^{1/2} \qquad \dots\dots\dots\dots\dots.. (3.9)$$

where β_x, β_y and β_z are the components of first order hyperpolarizability and is given in eq. (3.10), (3.11) and (3.12) respectively.

$$\beta_x = \left[\beta_{xxx} + \beta_{xyy} + \beta_{xzz} \right]^2 \qquad \dots\dots\dots\dots\dots.. (3.10)$$

$$\beta_y = \left[\beta_{yyy} + \beta_{yzz} + \beta_{yxx} \right]^2 \qquad \dots\dots\dots\dots\dots.. (3.11)$$

$$\beta_z = \left[\beta_{zzz} + \beta_{zxx} + \beta_{zyy} \right]^2 \qquad \dots\dots\dots\dots\dots.. (3.12)$$

First order hyperpolarizability (β) is a third rank tensor, described by a $3\times3\times3$ matrix. The 27 components of the 3D matrix can be reduced to 10 components by the Kleinman symmetry [79]. The output from Gaussian 09 provides 10 components of the matrix as β_{xxx}, β_{xyy}, β_{xzz}, β_{yyy}, β_{yzz}, β_{yxx}, β_{zzz}, β_{zxx}, β_{zyy} and β_{xyz}. The theoretical determination of hyperpolarizability can be used for understanding the relationship between the molecular structure and nonlinear optical properties.

3.9 Mulliken population analysis

Mulliken population analysis computes charges by dividing orbital overlap evenly between the two atoms involved. The sum over the Mulliken charges of all atoms equals the total number of electrons in the system. According to Mulliken formalism [80], the total population is a sum of net atomic populations and overlap

populations. In Mulliken population analysis, partial atomic charges were estimated from the density functional theory calculations based on the linear combination of atomic orbitals molecular orbital method (LCAO-MO). Mulliken population analysis strongly depends on the basis set and hence cannot provide physically meaningful absolute magnitudes of atomic charges.

3.10 Harmonic oscillator model of aromaticity (HOMA)

The term aromatic usually applied to benzene-like structures. Benzene is a six-membered ring structure in which six π electrons are delocalized around the ring. Aromaticity has an important effect on geometry, reactivity of functional groups, stability and magnetic properties of organic compounds. Different aromaticity indices are used to measure the aromaticity of the ring. The aromaticity is measured based on bond lengths by the harmonic oscillator model of aromaticity (HOMA) [81, 82] in eq. (3.13).

$$HOMA = 1 - \frac{\alpha}{n} \sum_{i=1}^{n} \left(R_{opt} - R_i \right)^2 \qquad \ldots\ldots\ldots\ldots\ldots\ldots (3.13)$$

where n is the number of bonds taken into the summation and α is a normalization constant (for CC bonds $\alpha = 257.7$; for CN bonds $\alpha = 93.52$) chosen to give $HOMA = 0$ for a model of nonaromatic system and $HOMA = 1$ for the full aromatic system. R_{opt} is the optimal bond length (for CC bonds $R_{opt} = 1.388$ Å; for CN bonds $R_{opt} = 1.334$ Å) and R_i is the running bond length.

3.11 Electron localization function analysis

Electron localization function (ELF) was first introduced by Becke and Edgecombe [83], is a measure of finding electrons in atomic and molecular systems. The expression for ELF is given in eq. (3.14).

$$\eta(r) = \frac{1}{1+\left(\dfrac{D_\sigma}{D_\sigma^0}\right)^2} \qquad\qquad \dots\dots\dots\dots\dots..\ (3.14)$$

where D_σ and D_σ^0 represent the curvature of the electron pair density for electrons of identical spins for the actual system and a homogeneous electron gas with the same density respectively. ELF $\eta(r)$ has values between 0 and 1, where 1 corresponds to perfect localization. The small values of ELF indicate a space region where electrons are delocalized, while high values correspond to a region of the space where the electrons are localized. In a molecule there are two types of basins: core basins (C) and valence basins (V). The valence basins are characterized by the number of atomic valence shells to which they participate, thus there are monosynaptic, disynaptic, trisynaptic basins and so on. The ELF has been used to explain structure, stability of organic, inorganic and organometallic compounds [84].