CHAPTER III

SPECTROSCOPIC TECHNIQUES USED FOR STRUCTURAL ELUCIDATION AND QUANTUM CHEMICAL COMPUTATION

INTRODUCTION

Spectroscopy is the study of the interaction between matter and electromagnetic radiation. Spectroscopy is a broad field and many sub-disciplines exist, each with numerous implementations of specific spectroscopic techniques. The combination of atoms into molecules leads to the creation of unique types of energetic states and therefore unique spectra of the transitions between these states. Molecular spectra can be obtained due to electron spin states (electron paramagnetic resonance), molecular rotations, molecular vibrations and electronic states. Rotations are collective motions of the atomic nuclei and typically lead to spectra in the microwave and millimeter-wave spectral regions. Vibrations are relative motions of the atomic nuclei and are studied by both Infrared and Raman spectroscopy. Electronic excitations are studied using visible and ultra violet spectroscopy as well as Fluorescence spectroscopy. Vibrational spectroscopy is the collective term used to describe two analytical techniques–infrared and Raman spectroscopy. Infrared and Raman spectroscopy are non–destructive, non–invasive tools that provide information about the molecular composition, structure and interactions within a sample. These techniques measure vibrational energy levels which are associated with the chemical bonds in the sample. The sample spectrum is unique, like a fingerprint, and vibrational spectroscopy is used for identification, characterization, structure elucidation, and reaction monitoring, quality control and quality assurance.

Modern spectroscopic techniques such as FT-IR, FT-Raman, UV-visible, have proven to be an exceptionally powerful technique for solving many drug molecules, food preservative molecules and the natural products because of the availability of sophisticated instrumentation methods. The present work mainly considers various spectroscopic techniques particularly relevant to food preservative activity. Various spectroscopic techniques and different electronic structure theory used for the research work are explained briefly in this chapter.

Instrumentation used for structural optimizations are:

(i) Infra red spectroscopy

(ii) Raman spectroscopy

(iii) UV-Visible spectrometer

3.1 MOLECULAR VIBRATION

As a consequence of elasticity the atom in a molecule vibrate about some mean position and they do not remain in a fixed relative position. The frequency of molecular vibrations ranges from 10^{13} to 10^{14} Hz corresponding to wave numbers of 300 to 3000 cm^{-1}. The change in bond length and bond angle occurs due to the stretching and bending vibration of the molecules. The energy corresponds to the lowest state is called the zero point energy. The molecule vibrates even at the lowest vibrational level where the vibrational energy is not zero. In molecular spectroscopy molecule has as many degrees of freedom as that of the total degrees of freedom of its individual atoms. Each atom has three degrees of freedom in the Cartesian coordinates (x, y, z), necessary to describe its position with respect to a fixed point in the molecule. A molecule of N atoms has 3N degrees of freedom. Of the 3N

degrees of freedom for non-linear molecules, 3 degrees of freedom describe translation and 3 describe rotation and the remaining 3N-6 degrees of freedom describe vibrational degrees of freedom. Linear molecules have 3N-5 degrees of freedom, since only 2 degrees of freedom are required to describe rotation. Of the 3N-6 vibrational modes, (N-1) modes are bond stretching vibrations and the other (2N-5) modes are angle-bending vibrations [27]. The atoms of subgroups of a molecule vibrate almost independently with same frequency so as to generate group frequencies, which are highly characteristic. Due to resonance, when two groups have comparable frequencies the group frequency may be shifted from the expected values.

3.2 INFRA RED SPECTROSCOPY

Infrared spectroscopy is a non destructive technique for material analysis by studying the interaction of infrared radiation with matter as a function of photon frequency. Fourier transform infrared spectroscopy provides specific information about the vibration and rotation of the chemical bonding and molecular structures, making it useful for analyzing organic materials and certain inorganic materials. An infrared spectrum represents a fingerprint of a sample with absorption peaks which correspond to the frequencies of vibrations between the bonds of the atoms making up the material. Because each material is a unique combination of atoms, no two compounds produce the exact infrared spectrum. Therefore, infra red spectroscopy can result in a positive identification (qualitative analysis) of every different kind of material. IR region of electromagnetic spectrum is generally subdivided into 3 regions such as near IR (12,500-4000 cm^{-1}), mid IR (4000-400 cm^{-1}) and far IR (400-50 cm^{-1}).

The mid IR is the region most commonly employed in standard laboratory investigations, as it covers most of the vibrational transitions. The mid IR region provides more information upon the structures of compounds and consequently it is much used as a procedure for identifying organic compounds for which it remains a form of functional group finger printing. Thus IR spectroscopy is an important and popular tool for structural elucidation and compound identification.

3.2.1 Selection Rules

In IR spectroscopy, the necessary condition for the absorption of a quantum of radiation 'hυ' by the molecule should be equal to the energy difference between two states represented by the wave functions 'ψ_i' and 'ψ_j'. The transition between these states under the influence of electromagnetic radiation depends on the interactions of electric field of the radiation with the electric dipole moment of the molecule.

According to quantum mechanical theory of a molecular system, the probability of transition from a state i to the state j is proportional to the square of the transition moment.

$$\mu_{ij} = \int \psi_i^* \mu \psi_j d\tau \qquad (3.1)$$

Where, 'μ' is the dipole moment of the molecule.

The dipole moment of a molecule is a function of the normal coordinates 'Q_k' of the vibrational mode and can be expanded in a Taylor series as:

$$\mu = \mu_0 + \left(\frac{\partial \mu}{\partial Q_k}\right)_0 Q_k + ... \qquad (3.2)$$

On neglecting higher terms

$$\mu_0 = \mu_0 \int \psi_i^* \psi_j d\tau + \left(\frac{\partial \mu}{\partial Q_k} \right)_0 \int \psi_i^* Q_k \psi_j d\tau \qquad (3.3)$$

The first term vanishes and the condition for the second term to be non-zero is:

(i) $\left(\frac{\partial \mu}{\partial Q_k} \right)_0$ must be finite atleast for one component of the dipole moment. That is, for a mode of vibration to be infrared active; the vibrational motion of that mode must give rise to change in dipole moment.

(ii) The integral $\int \psi_i^* Q_k \psi_j d\tau$ must be finite, which is possible only if the vibrational quantum number change $\Delta V = \pm 1$ under harmonic approximation and for a harmonic oscillator $\Delta V = \pm 1, \pm 2, \pm 3. .$

3.2.2 Fourier Transform Infrared Spectrometer

FT-IR spectroscopy has dramatically improved the quality of infrared spectra and minimized the time required to obtain data. FT-IR relies on the fact that most of the molecules absorb light in the infrared region of the electromagnetic spectrum. This absorption corresponds specifically to the bonds present in the molecule. The fundamental requirement for infrared activity, leading to absorption of infrared radiation, is that there must be a net change in dipole moment during the vibration for the molecule or the functional group under study. The average modern infrared instrument records spectra from an upper limit of around 4000 cm^{-1} (by convention) down to 400 cm^{-1} as defined by the optics of the instrument (commonly based on potassium bromide, KBr). The resultant absorption spectrum from the bond natural vibration frequencies indicates the presence of various chemical bonds and functional groups present in the sample.

3.2.2.1 Instrumentation

IR spectrometer consists of source of radiation, sample holder, detector and recorder. In IR spectrometer, the Nernst glower acts as a source of radiation. A spindle of rare earth oxide such as thorium, cesium or zirconium of length 2.5 cm and 0.25 cm diameter is enclosed within the glower. The conductivity of the filament is induced by the preheating of glower. Emissivity of the glower is inversely proportional to the wavelength and the maximum peak appears close to 1.4 μm at a temperature of 1600-1700˚C. The radiation coming out of the source enters into the core optics system, where it diverges into two beam by the mirrors M1 and M2. Of the two beam one passes through the sample and other acts as the reference beam. The two beams enter the rotating sector where the sample beam passes or the reference beam gets reflected with a definite interval to the monochromator slit. The grating or the prism acts as the monochromator and splits the polychromatic radiation into its component wavelength. Intensity of resolution depends upon the slit width and the quality of the mirror used. For the region 650-4000 cm^{-1} rock salt prisms are used. The radiations enter the detectors. Depending upon the modes of operation, the types of detectors available are the Thermopiles, Bolometer and Golay cell. The commonly used detector is the Golay cell. The Golay cell consists of a blackened infrared absorbing material at one end and a flexible silvered diaphragm at other end. Xenon gas fills the chamber. The radiation entering the IR window is absorbed by the blackened plate. Expansion of the diaphragm takes place due to the heat conduction of the gas. Alternating radiation produces alternating voltage with an approximate response time of 20 Ms. The low intensity signal is amplified by the amplifier and is recorded by the recorder.

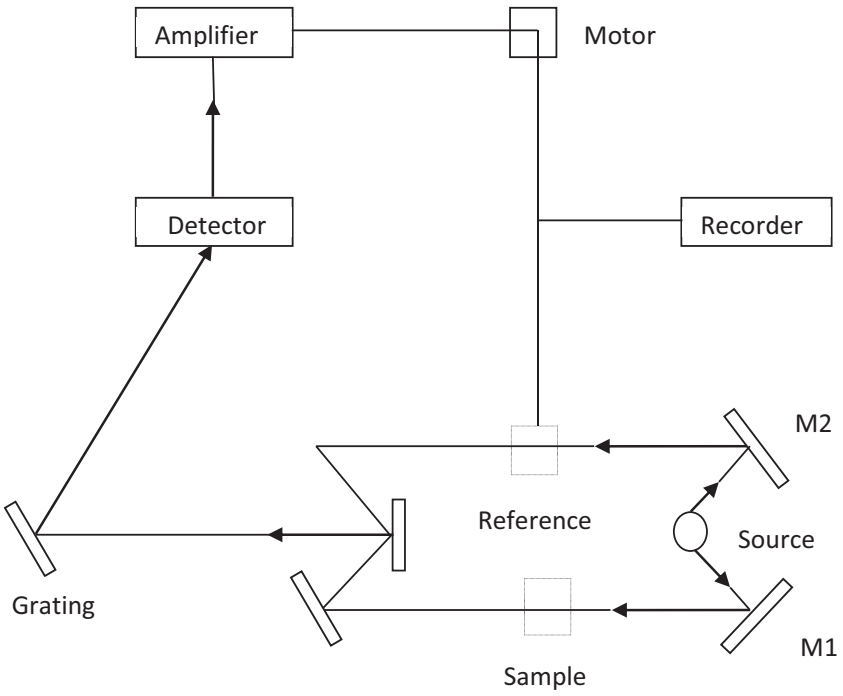

Fig. 3.1 Schematic diagram of FT-IR Spectrometer

3.2.2.2 Sample handling

Sample handling is considered as an important technique in IR spectroscopy. One of the advantages of infrared spectosopy is that it can be applied to all molecules. There are various methods of sample preparation involved in FT-IR. For the liquid samples a drop of the sample is placed in between the two infrared transmitting plates of sodium chloride salt. The drop forms a thin film between the plates. For solid samples KBr pellet techniques is used. The solid sample is grinded with potassium bromide to form a very fine powder and pressed into thin pellet of suitable size and placed in a sample holder for recording the spectrum. For chemical analysis the most preferred region is the mid infrared region (4000-400 cm^{-1}).

3.3 RAMAN SPECTROSCOPY

When electromagnetic radiation of energy content hυ irradiates a molecule, the energy may be transmitted, absorbed, or scattered. In Tyndall effect the radiation is scattered by particles (smoke, or fog, for example). The molecules scatter light in Rayleigh scattering. No change in wavelength of individual photons occurs either in Tyndall or Rayleigh scattering. This effect had been theoretically predicted by Smekal before the successful experimental demonstration of the effect by Raman and is therefore referred to as the Smekal-Raman effect in the German literature.

In Raman spectrometer the sample is irradiated with an intense source of monochromatic radiation usually in the visible part of the spectrum. Generally this radiation frequency is much higher than the vibrational frequencies but it is lower than the electronic frequencies. The radiation scattered by the sample is analyzed in the spectrometer. Rayleigh scattering can be looked on as an elastic collision between the incident photon and the molecule. Since the rotational and vibrational energy of the molecule is unchanged in an elastic collision, the energy and therefore the frequency of the scattered photon is the same as that of the incident photon. This is the strongest component of the scattered radiation. The Raman effect can be looked on as an inelastic collision between incident photon and the molecule where as a result of the collision the vibrational or rotational energy of the molecule is changed by an amount ΔE_m. In order that energy may be conserved, the energy of the scattered photon, $h\upsilon_s$, must be different from the energy of the incident photon $h\upsilon_i$, by an amount equal to ΔE_m:

$$h\upsilon_i - h\upsilon_s = \Delta E_m \qquad (3.4)$$

If the molecule gains energy then ΔE_m is positive and υ_s is smaller than υ_i, giving rise to Stokes lines in the Raman spectrum. This terminology arose from Stokes rule of fluorescence which stated that fluorescent radiation always occurs at lower frequencies than that of the exciting radiation. If the molecule loses energy, then ΔE_m is negative and υ_s is larger than υ_i, giving rise to antistokes line in Raman spectrum.

3.3.1 Classical Theory of Raman Effect

An insight into the phenomenon of Raman scattering is possible on the basis of classical considerations. When a molecule is placed in an electric field, polarization of the medium takes place as the negatively charged electron cloud is being attracted towards the positive pole and the positively charged nuclei being attracted toward the negative pole. The polarization P induced is propotional to the applied electric field E. i.e,

$$P = \alpha E \tag{3.5}$$

The constant of proportionality α is the polarizability of the molecule. In general, it is a tensor with nine components.

When radiation of frequency υ_0 is allowed to fall on a molecule, each molecule experiences a varying electric field

$$E = E_0 \cos 2\pi \upsilon_0 t \tag{3.6}$$

Let Q be the normal coordinate associated with a particular mode of vibration of frequency υ_m of the molecule. In harmonic approximation, Q can be written as

$$Q = Q_0 \cos 2\pi \upsilon_m t \tag{3.7}$$

The probability α can be expressed as a Taylor series in the normal coordinate Q as

$$\alpha = \alpha_0 + \left(\frac{\partial \alpha}{\partial Q}\right)_0 Q + ... \quad (3.8)$$

Where, 'α_0' is the polarizability at the equilibrium position. The term $\left(\frac{\partial \alpha}{\partial Q}\right)_0$ is the rate of change of 'α' with respect to 'Q' evaluated at the equilibrium position. The induced dipole moment is therefore:

$$P = \alpha_0 E_0 \cos 2\pi \upsilon_0 t + \frac{1}{2}\left(\frac{\partial \alpha}{\partial Q}\right)_0 Q_0 E_0 \left[\cos 2\pi(\upsilon_0 + \upsilon_m)t + \cos 2\pi(\upsilon_0 - \upsilon_m)t\right] \quad (3.9)$$

The first term represents Rayleigh scattering and the second term represents anti-stokes ($\upsilon_0+\upsilon_m$) and stokes ($\upsilon_0-\upsilon_m$) lines of Raman scattering. For Raman active vibrations, the rate of change of polarizability should be non zero.

$$\text{i.e.} \quad \left(\frac{\partial \alpha}{\partial Q}\right)_0 = 0 \quad (3.10)$$

Thus, we get the important rule that a molecule will be Raman active only if it causes a change in the component of the polarizability. Though the classical theory correctly describes the frequencies $\upsilon_0 \pm \upsilon_m$ of the Raman lines, it fails to give the correct intensities. Only a quantum mechanical theory would be able to predict the intensities.

3.3.2 Quantum Theory of Raman Effect

In the quantum picture, radiation has both particle and wave nature. In explaining Raman scattering, incident radiation of frequency υ_0 is considered as a stream of particles (photons) undergoing collision with the molecules. If the collision is perfectly elastic, there will not be any exchange of energy between the photons and the molecules. However, there will be exchange of energy between the two if

the collision is inelastic. The molecules can gain or lose energy equal to the energy difference ΔE between any two of its allowed states. If the molecule gains energy, the scattered photons will have frequency $\upsilon_0 - \upsilon_m$, where $\upsilon_m = \Delta E/h$ (stokes line). On the other hand if it loses energy the scattered photon will have frequency $\upsilon_0 - \upsilon_m$ (anti-stokes line). The different process giving rise to Rayleigh, Stokes and anti-Stokes lines are illustrated in Fig. 3.2.

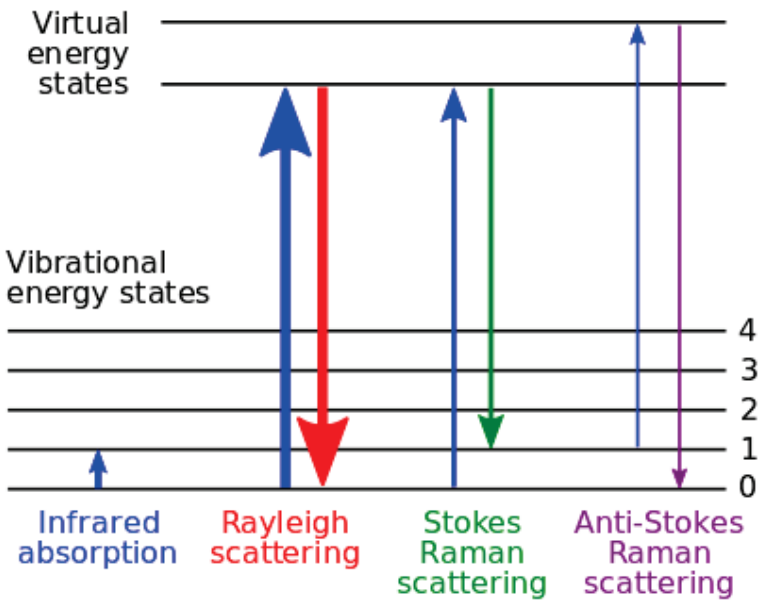

Fig. 3.2 Schematic representations of Raman lines

When a system interacts with a radiation of frequency υ_0 it may make an upward transition to a virtual state of the system. Most of the molecules of the system return to the original state from the virtual state giving the Rayleigh scattering. However, a very small fraction returns to states of higher and lower energies giving rise to Stokes and anti-Stokes lines respectively. If the virtual state of the system coincides with a real state of the system, it will lead to Resonance Raman effect. It

may be pointed out that the annihilation of the incident photon and the creation of the scattered photon are simultaneous.

The intensity of spectral line depends on number of factors, the most important being the initial population of the state from which the transition originates. The Stokes line originating from $\upsilon = 1$ gives the Raman shift υ_m. Based on Boltzmann distribution for the population in states, the intensity ratio of Stokes to anti-Stokes is given by

$$\frac{I_s}{I_{a.s}} = \exp \frac{h\upsilon_m}{kT} \tag{3.11}$$

Where k is the Boltzmann constant and T is the temperature in Kelvin. A mire rigorous theory, taking other factors into account, gives

$$\frac{I_s}{I_{a.s}} = \frac{(\upsilon_0 - \upsilon_m)^4}{(\upsilon_0 + \upsilon_m)^4} \exp \frac{h\upsilon_m}{kT} \tag{3.12}$$

Anti-Stokes lines have much less intensity than Stokes lines.

3.3.3 Polarizability Tensor and Polarizability Ellipsoid

Intense laser beam (electric field) persuades the molecule to endure distortion, since the positively charged molecules are shifted in the direction of the field and electrons in the opposite direction, producing an induced dipole moment. In matrix form, the polarizability tensor can be written in terms of the components of the field vectors as:

$$\begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix} = \begin{pmatrix} \alpha_{xx}\alpha_{xy}\alpha_{xz} \\ \alpha_{yx}\alpha_{yy}\alpha_{yz} \\ \alpha_{zx}\alpha_{zy}\alpha_{zz} \end{pmatrix} \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix} \tag{3.13}$$

The polarizability tensor provides valuable information regarging the polarization properties of vibrational modes [28] and the dependence of vibrational intensity on polarization direction [29]. Observed transition is Raman active only if any of the components in the polarizability tensor has to be changed during vibration. For normal Raman scattering, polarizability tensor should be symmetric, i.e., $\alpha_{xy}= \alpha_{yx}$, $\alpha_{xz}= \alpha_{zx}$ and $\alpha_{yz}= \alpha_{zy}$. Polarizability ellipsoid is another valuable tool for predicting the polarization properties of internal vibrations of sample molecules [30]. A plot representing $\frac{1}{\sqrt{\alpha_i}}$, where '$\alpha_i$' is the value of α in the i^{th} direction from the centre of gravity of a molecule in all directions, a three dimensional surface is obtained, which is the polarizabitity ellipsoid. During normal vibration, the change in size, shape or orientation of the polarizability ellipsoid represents active mode. Whenever the size of the ellipsoid changes the diagonal element α_{xx}, α_{yy}, and α_{zz} change simultaneously. The diagonal element change with different rates for any modification in the shape of polarizability ellipsoid and the off diagonal elements vary for orientational changes.

3.3.4 Depolarization Ratio

In addition to intensity and frequency information, Raman measurements provide an additional parameter, the depolarization ratio that is useful in determining the structure of the molecule. Depolarization ratio can be obtained by considering a molecule irradiated by a plane polarized light from Y-direction whose electric field 'E' is in the Z-direction shown in Fig. 3.3. The scattered beam by the molecule is

observed in the X–direction and I_y ($I_{parallel}$), I_z ($I_{perpendicular}$) are respectively the minimum and maximum intensity observed through the analyzer.

The depolarization ratio is given by,

$$\rho = \frac{I_{perpendicular}}{I_{parallel}} \tag{3.14}$$

which measures the degree of polarization. Theoretically, depolarization ratio is given by [31]

$$\rho = \frac{3g^s + 5g^a}{10g^a + 4g^s} \tag{3.15}$$

$$g^o = \frac{1}{3}\left[\alpha_{xx} + \alpha_{yy} + \alpha_{zz}\right]^2 \tag{3.16}$$

Where,

$$g^s = \frac{1}{3}\left[\left(\alpha_{xx} - \alpha_{yy}\right)^2 + \left(\alpha_{yy} - \alpha_{zz}\right)^2 + \left(\alpha_{zz} - \alpha_{xx}\right)^2\right] + \frac{1}{2}\left[\left(\alpha_{xy} - \alpha_{yx}\right)^2 + \left(\alpha_{yz} + \alpha_{yz}\right)^2 + \left(\alpha_{xy} + \alpha_{zx}\right)^2\right]$$

$$g^a = \frac{1}{2}\left[\left(\alpha_{xy} - \alpha_{yx}\right)^2 + \left(\alpha_{yx} - \alpha_{zy}\right)^2 + \left(\alpha_{xz} - \alpha_{zx}\right)^2\right]$$

For normal Raman scattering, the polarizability tensor is symmetric and therefore, $g^0 = 0$. For total symmetric vibration $g^0 > 0$ and $g^2 \geq 0$. The vibration is polarized and 'p' takes the value range with $0 \leq \rho \leq 3/4$. The non symmetric vibrations are depolarized and therefore, $g^0 = 0$ and $g^s > 0$. Then $\rho = 3/4$. Calculated depolarization ratio with all these values provide valuable information regarding the symmetry of vibration which helps in band assignments and vibrational analysis [32, 33].

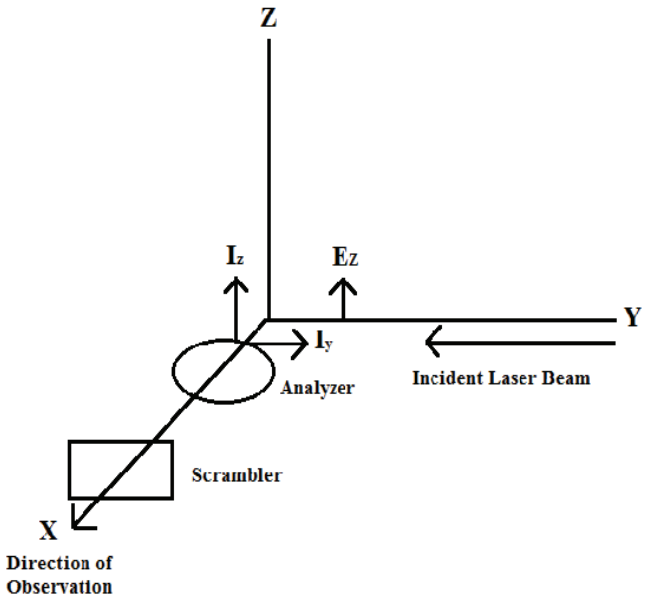

Fig. 3.3 Depolarization measurements

3.3.5 FT-Raman Instrumentation

Raman spectrometer basically employ either dispersive or Fourier transform techniques for recording the spectra. Fourier transform Raman spectroscopy system have been available since 1987. Unlike dispersive Raman spectrometers, FT-Raman spectrometers have mainly Jacquinot (high throughput) and Felgett (multiplex) advantages apart from that they normally produce fluorescence free background on weak Raman signals and high resolution with good wave number accuracy. The schematic of a Raman spectrometer is shown in Fig. 3.4

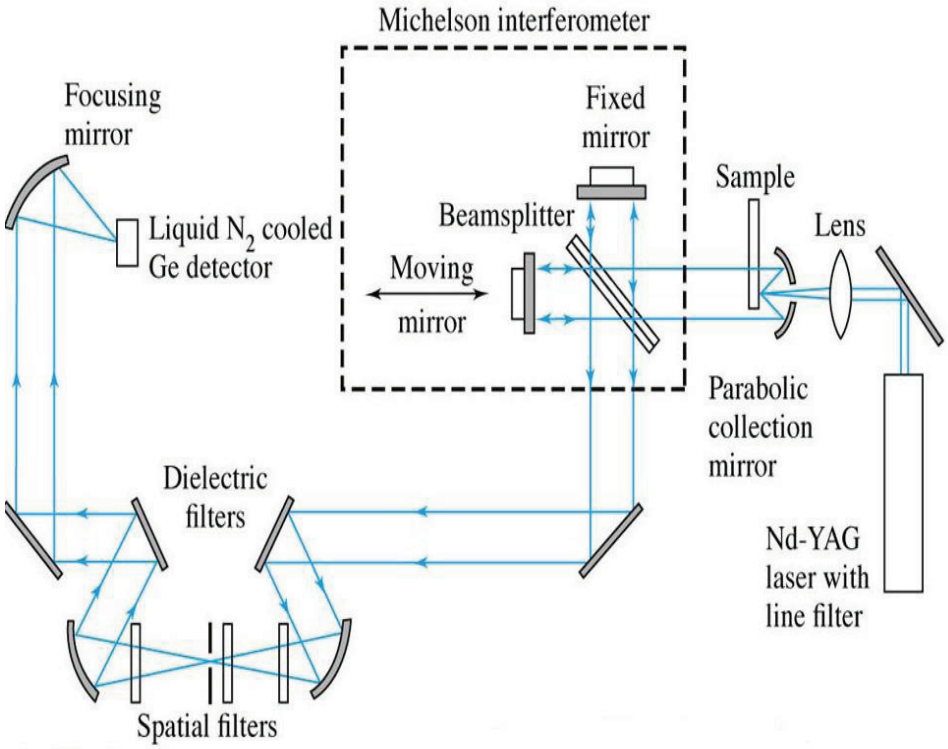

Fig. 3.4 Schematic of a Raman spectrophotometer

Initially the sample is illuminated with a monochromatic laser beam which interacts with the molecules of the sample and originates a scattered light having a frequency different from that of incident light. The scattered light arising from this inelastic collision is used to construct a Raman spectrum. Much of the scattered radiation has a frequency which is equal to frequency of incident radiation and constitutes Rayleigh scattering. Only a small fraction of scattered radiation has a frequency different from frequency of incident radiation and constitutes Raman scattering. When the frequency of incident radiation is higher than the frequency of scattered radiation, Stokes lines appear in Raman spectrum. But when the frequency of incident radiation is lower than the frequency of scattered radiation, anti-Stokes

lines appear in Raman spectrum. Scattered radiation is usually measured at right angle to incident radiation [34].

Stokes shifted Raman bands involve the transitions from lower to higher energy vibrational levels and therefore, Stokes bands are more intense than anti-Stokes bands and hence they are measured in conventional Raman spectroscopy while anti-Stokes bands are measured with fluorescing samples because fluorescence causes interference with Stokes bands [35]. Since Raman scattering due to water is low, water is an ideal solvent for dissolving samples and glass is used for optical components (mirror, lens, sample cell) in Raman spectrophotometer.

A Raman spectrum is presented as an intensity-versus-wavelength shift and is recorded over a range of 4000–10 cm^{-1}. Raman spectrophotometers can be dispersive or non-dispersive such that dispersive type Raman spectrophotometers uses prism or grating while non-dispersive Raman spectrophotometers uses an interferometer such as Michelson interferometer.

Wide range of lasers such as Argon ion laser (488 and 514.5 nm), Krypton ion laser (530.9 and 647.1 nm), Helium–Neon (He–Ne) (632.8 nm), Near Infrared (IR) diode lasers (785 and 830 nm), Neodymium–Yttrium Aluminum Garnet (Nd:YAG) and Neodymium–Yttrium Ortho-Vanadate (Nd:YVO_4) (1064 nm) and frequency doubled Nd:YAG and Nd:YVO_4 diode lasers (532 nm) can be used as light source in Raman spectrophotometers [35]. Notch filters, edge pass filters, or band pass filters are used to isolate a single laser beam and to filter the elastic scattered radiation at the wavelength corresponding to the laser line (Rayleigh scattering).

Through the beam splitter in the Michelson interferometer, half of the Raman signal is transmitted to a fixed mirror and half is reflected to a moving mirror. Because of the optical path difference caused by the moving mirror, the two beams reflected from the two mirrors undergo constructive and destructive interference, determined by the travel range of the moving mirror. Finally, the signal with all frequencies is registered on the detector simultaneously and the interferogram can be obtained from the detector. Multichannel CCD detectors are used with laser wavelengths of less than 1 μm while single element low band-gap semiconductor such as Germanium (Ge) or Indium–Gallium–Arsenic (InGaAs) detectors are used with laser wavelengths of greater than 1 μm [35].

3.4 ELECTRONIC SPECTROSCOPY

Electronic spectroscopy is used to study the transition of electrons between different electronic states which fall in the visible and ultraviolet regions of the electromagnetic spectrum. UV-Visible region of the spectrum is conventionally divided into three sub-domains termed as near UV (185-400 nm), visible (400-700 nm) and very near infrared (700-1100 nm). Most commercial spectrophotometers cover the spectral range of 185-900 nm. The origin of absorption in this domain is the interaction of photons from a source with ions or molecules of the sample. When a molecule absorbs a photon from the UV-Visible region, the corresponding energy is captured by one (or several) of its outermost electrons. UV-Visible spectrometers correlate the data over the required range and generate spectrum of the compound under analysis as a graph representing the transmittance (or the absorbance) as a function of wavelength along the abscissa, given in nanometers. In UV-Visible

spectroscopy, the transmittance (T) is a measure of the attenuation of a beam of monochromatic light based upon the comparison between the intensities of the transmitted light (I) and the incident light (I_0) according to whether the sample is placed or not, in the optical pathway between the source and the detector. When plotted into a spectrum as wavelength against absorbance, the absorbance is defined by using the Beer-Lambert law:

$$A = -\log_{10}\left(\frac{I}{I_0}\right) \tag{3.17}$$

where, 'A' is the measured absorbance, 'I_0' is the intensity of the incident light at a given wavelength and 'I' is the transmitted intensity. Most spectrometers display absorbance on the vertical axis and the commonly observed range is from 0 (100% transmittance) to 2 (1% transmittance).

3.4.1 Instrumentation

The schematic diagram of UV spectrophotometer is shown in Fig. 3.5. The basic parts of a spectrophotometer are a light source, a holder for the sample, a diffraction grating in a monochromator or a prism to separate the different wavelengths of light and a detector. The radiation source is often a Tungsten filament (300-2500 nm), a deuterium arc lamp, which is continuous over the ultraviolet region (190-400 nm), Xenon arc lamp, which is continuous from 160-2000 nm; or more recently, light emitting diodes (LED) for the visible wavelengths. The detector is typically a photomultiplier tube, a photodiode, a photodiode array or a charge coupled device (CCD). Single photodiode detectors and photomultiplier tubes are used with scanning monochromator, which filter the light so that only light of a

single wavelength reaches the detector at one time. The scanning monochromator moves the diffraction grating to "step-through" each wavelength so that its intensity may be measured as a function of wavelength. Fixed monochromator are used with CCDs and photodiode arrays.

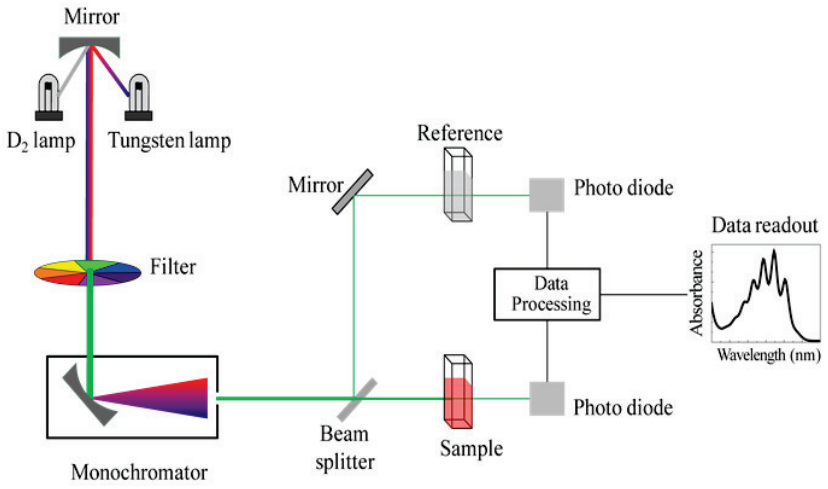

Fig. 3.5 Schematic diagram of UV-Visible spectrometer

3.4.1.1 Radiation source

The radiation sources used often are Deuterium arc lamp, Hydrogen lamp, Xenon discharge lamp, Mercury arc lamp, Tungsten lamp, Mercury vapour lamp and Carbonone lamp. While selecting the radiation source, it is important that the power of the source does not change abruptly over its wavelength range. Also they should have a stable, high intensity output that covers a wide range of wavelengths [36].

3.4.1.2 Filter

It is used to provide a narrow band of radiation. It may be glass filters, organic dye filters, interference filters or Fabry Perrot etalons [37].

3.4.1.3 Monochromator

The monochromator contains an entrance slit, a dispersion device, a collimating lens such as a prism or grating, a focusing lens and an exit slit. The entrance slit allows the polychromatic radiation from the source to enter in to the monochromator. The beam is collimated and strikes the dispersing device at an angle and splits into component wavelengths. Through the exit slit, radiation of a particular wavelength leaves the monochromator. This can be done by moving through the dispersive device or the exit slit. Further this light is split into two beams before it reaches the sample [38].

3.4.1.4 Sample and reference cells

One of the two divided beam is passed through the sample solution and the other is passed through the reference solution. The sample and the reference solution are contained in cells. The cells should be transparent to the wavelength region which is to be recorded. Usually the cells made up of silica or quartz is used. Glass cannot be used for the cells as it absorbs light in the UV region.

3.4.1.5 Detectors

The detector converts the electromagnetic radiation into alternating current. A good detector should have high spectral sensitivity, good wavelength response,

fast response time and high signal to noise ratio. The commonly used detectors are photodiodes, charged coupled device (CCD) and photomultiplier tubes [37]. The samples and reference beam are measured using two detectors which are photodiodes. One of the detectors receives beam from sample cell and the second detector receives beam from the reference. The intensity of radiation from the reference cell is stronger than the beam of sample cell. This results in the generation of pulsating or alternating currents in the detectors. Since this current is of low intensity it is amplified using an amplifier. A computer connected to it stores all the data generated and produces the spectrum of the desired compound [38].

The UV-Visible spectrum is a plot of Absorbance (A) versus wavelength (λ). The wavelength at which the sample absorbs the maximum amount of light is known as λ_{max}.

3.5 QUANTUM CHEMICAL COMPUTATION

Computational chemistry is a branch of chemistry that uses computer simulation to solve chemical problems. It uses methods of theoretical chemistry, in corporate into efficient computer programs, to calculate the structures and properties of molecules and solids. Rather than analytical method Quantum chemical computation acts as an efficient tool in solving many body problems. While computational results normally complement the information obtained by chemical experiments, in some cases it can predict the unobserved chemical phenomena such as structure absolute and relative (interaction) energies, electronic charge density distributions, dipole and multipole moments, vibrational frequencies, reactivity, or other spectroscopic quantities, and cross sections for collision with other particles.

The methods used cover both static and dynamic situations. In all cases, the computer time and other resources (such as memory and disk space) increase rapidly with the size of the system being studied. That system can be one molecule, a group of molecules, or a solid. Computational chemistry methods range from very approximate to highly accurate; the latter are usually feasible for small systems only. *ab initio* methods are based entirely on quantum mechanics and basic physical constants. Other methods are called empirical or semi-empirical because they use additional empirical parameters. Both *ab initio* and semi-empirical approaches involve approximations.

3.5.1 Ab Initio Quantum Chemical Methods

The term *ab initio* is Latin word for "from the beginning and indicates that this calculation is based on fundamental principles. *Ab initio* computations are the computations that are derived directly from theoretical principles with no inclusion of experimental data. This is an approximate quantum chemical calculation. The most common type of Ab initio calculations is called Hartee-Fock (HF) calculations. In HF calculations, the primary approximation is the central field approximation. Here the Columbic electron - electron repulsion is taken into account by integrating the repulsive term. It gives the average effect of the repulsion. This is a variational calculation, which implies that the approximate energies calculated are all equal to or greater than that of the exact energy. Though the accuracy of the calculation depends on the size of the basis set used, due to the central field approximation, the energies from HF calculations are always equal or greater than the exact and tend to a limiting value called Hartree-Fock limit [39].

The second approximation in HF calculations is due to the fact that the wave function must be described by some mathematical function, which is known exactly for only a few one-electron systems. The combinations of either Slater type orbital (STO), exp (-ar) or Gaussian type orbital (GTO), exp ($-ar^2$) are mostly used basis functions. Because of this approximation, most HF calculation gives a computed energy greater than the Hartree-Fock limit. The exact set of basis functions used in either STO-3G or 6-311++G**. In general, *ab initio* calculations give very good qualitative results. Accuracy of the quantitative results is inversely proportional to the size of the molecule [39].

3.5.2 Semi-Empirical Method

Semi empirical methods are zero dimensional just like force field methods and they provide a method for calculating the electronic wave function. Semi-empirical methods such as AM1, PM3 have a minimum basis (lacking polarization and diffuse function), electron correlation is only included implicitly by the parameters and no polarizability data have been used for deriving the parameter. The most common approach for parameterization of modern semi empirical methods aims at reproducing experimental data or the result of correlated *ab initio* or DFT calculations. The availability of large reference data bases for thermo chemistry and molecular geometries helps in the development of modern semi empirical methods. Advantages of semi empirical method are a substantial reduction in the required computation time and an increase in ability to execute calculations for large molecule [40].

3.5.3 Density Functional Theory

DFT is a computational quantum mechanical modeling method used in physics, chemistry and material science to investigate electronic structure of many-body systems, in particular atoms, molecules and the condensed phases. The basis for DFT is the proof by Hohenberg and Kohn that the ground state electronic energy is determined completely by the electron density ρ. The goal of DFT methods is to design functional connecting the electron density with the energy. Early attempts at designing DFT models tried to express all the energy components as a functional of the electron density but these methods had poor performance, and wave function - based methods were consequently preferred. The success of modern DFT methods are based on the suggestion by Kohn and Sham in 1965 that the electron kinetic energy should be calculated from an auxiliary set of orbital's used for representing the electron density. DFT is conceptually and computationally very similar to Hartree- Fock theory, but provides much better results and has consequently become a very popular method. The main problem of DFT is the inability to improve the results, and to describe certain important features such as Van der Waals interactions [40].

Although DFT has its conceptual roots in the Thomas Fermi model, DFT was put on a firm theoretical footing by Hohenberg-Kohn theorems. The first theorem demonstrates the existence of a one-to-one mapping between the ground state electric density and the ground state wave function of many particle system. The second theorem proves that the ground state density minimizes the total electronic energy of the system. The most common implementation of DFT is through the Kohn-Sham method. Within the frame work of Kohn-Sham DFT, the

intractable many body problem of interacting electron in a static external potential is reduced to a tractable problem of non- interacting electrons moving in an effective potential.

3.5.3.1 Kohn-Sham Theory

The foundation for the use of DFT methods in computational chemistry is the introduction of orbital's, as suggested by Kohn and Sham. The KS model is closely related to the HF method, sharing identical formulas for the kinetic, electron-nuclear and Coulomb electron-electron energies. The Hamiltonian operator is given by

$$H_\lambda = T + V_{ext}(\lambda) + \lambda V_{ee} \tag{3.18}$$

The external potential operator V_{ext} is equal to V_{ee} for λ=1, but for intermediate λ values it is assumed that V_{ext} (λ) is adjusted such that the same density is obtained for λ=1, for λ=0 and for all intermediates λ values. For λ=0, the exact solution to Schrodinger equation is given as

$$T_s = \sum_{i=1}^{Nelec} \langle \phi_i | -1/2\nabla^2 | \phi_i \rangle \tag{3.19}$$

The subscript S denotes that the kinetic energy is calculated from Slater determinant. The exact kinetic energy can be calculated from the natural orbital's arising from the exact density matrix.

$$T[\rho_{exact}] = \sum_{i=1}^{\alpha} n_i \langle \phi_i^{NO} | -1/2\nabla^2 | \phi_i^{NO} \rangle \tag{3.20}$$

$$\rho_{exact} = \sum_{i=1}^{\alpha} n_i \left| \phi^{NO}{}_j \right|^2 \tag{3.21}$$

$$N_{elec} = \sum_{i-1}^{\alpha} n_i \tag{3.22}$$

Since the exact density matrix is not known an approximate density can be written as

$$\rho_{approx} = \sum_{i=1}^{N_{elec}} \left|\phi_i\right|^2 \tag{3.23}$$

The general DFT expression can be written as

$$E_{DFT}[\rho] = T_s[\rho] + E_{ne}[\rho] + J[\rho] + E_{xc}[\rho] \tag{3.24}$$

By equating E_{DFT} to the exact energy, this expression defines E_{xc}, and the E_{ne} and J potential energy terms.

$$E_{xc}[\rho] = (T[\rho] - T_S[\rho]) + (E_{ee}[\rho] - J[\rho]) \tag{3.25}$$

Kohn-Sham theory is much less sensitive to inaccuracies in the functional than orbital –free theory. While orbital free theory is a true density functional theory, Khon-Sham methods are independent particle models, analogous to Hartree- Fock theory.

3.5.3.2 Basis Sets

A basis set is a mathematical description of the orbitals within a system used to perform theoretical calculations. The basis sets impose constraints on each electron to a particular region of space as large basis sets impose fewer constraints to approximate exact molecular orbitals. Standard basis sets for electronic structure calculations use linear combinations of basic functions to form the orbitals. An individual molecular orbital is defined as:

$$\phi_i = \sum_{\mu=1}^{n} C_{\mu i} \chi_\mu \tag{3.26}$$

Where 'φ_i' is the i^{th} molecular orbital, '$C_{\mu i}$' are the coefficients of linear combination and 'χ_μ' is the μ^{th} atomic orbital and n is the number of atomic orbitals. There are two types of basis functions commonly used in electronic structure calculations: Slater type orbitals (STO) and Gaussian type orbitals (GTO). Slater type orbitals have the functional form

$$\chi_{\xi,n,l,m}(r,\theta,\varphi) = NY_{l,m}(\theta,\varphi) r^{n-1} e^{-\xi r} \tag{3.27}$$

Where N is a normalization constant and $Y_{l,m}$ are spherical harmonic functions. STOs are primarily used for atomic and diatomic systems where high accuracy is required and in semi-empirical methods where all three and four centre integrals are neglected. They can also be used with density functional methods. Gaussian type orbitals can be written in terms of polar or Cartesian coordinates as:

$$\chi_{\xi,n,l,m}(r,\theta,\varphi) = NY_{l,m}(\theta,\varphi) r^{2n-2-l} e^{-\xi^2} \tag{3.28}$$

$$\chi_{\xi,l_x,l_y,l_z}(x,y,z) = N x^{l_x} y^{l_y} z^{l_z} e^{-\xi^2} \tag{3.29}$$

The sum l_x,l_y and l_z determines the type of orbitals.

3.5.3.3 Classification of basis sets

The most common minimal basis set is STO-nG, where n is an integer. This n value represents the number of Gaussian primitive functions comprising a single basis function.

(i) Commonly used minimal basis sets of this type are: STO-3G, STO-4G, STO-6G, and STO-3 G*-Polarized version of STO-3G.

(ii) Split valence basis sets, such as 3-21G and 6-31G have two (or more) sizes of basis function for each valence orbital and allow orbitals to change size, but not to change shape. Triple split valence basis sets like 6-311G, use three sizes of contracted functions for each orbital type.

(iii) Polarized basis sets, by adding orbitals with angular momentum beyond what is required for the description of each atom in the ground state, allow orbitals to change the shape. The 6-31G (d) basis set also known as 6-31 G* which contains d-functions added to the heavy atoms, is becoming very popular for calculations involving up to medium sized systems.

(iv) Diffuse functions are large-sized versions of s-and p-type functions. They allow orbital to occupy a large region of space. Basis sets with diffuse functions are important for systems where electrons are relatively far from the nucleus: molecules with lone pairs, anions and other systems with significant negative charge, systems in their excited states, systems with low ionization potentials, and so on. The 6-31+G(d) is the 6-31G(d) basis set with diffuse functions added to heavy atoms.

3.6 NATURAL BOND ORBITAL ANALYSIS

Natural bond orbital (NBO) analysis is an effective technique, in which the electronic wave function is interpreted as a set of occupied Lewis-type orbital's, paired with a set of formally unoccupied non-Lewis type orbitals. NBO method was developed by Weinhold and coworkers which is becoming a powerful and popular method for the study of bonding concepts [41]. NBO analysis can also be employed to identify and substantiate the possible intra and intermolecular interactions that

would form the hydrogen bonded network [42]. Existence of intermolecular O-H…O hydrogen bonds which are due to the interaction between the lone pair (LP) of oxygen with the antibonding orbital (σ*) has been confirmed by the results of NBO analysis. The electronic interactions with these orbital's, the deviation from the Lewis electronic structure, and the delocalization effects can be interpreted as charge transfer between the filled Lewis orbital's (donors) and the theoretically empty non-Lewis orbital's (acceptors) [43]. The second –order Fock matrix was used to evaluate the donor-acceptor interactions in the NBO basis. For each donor NBO (i) and acceptor NBO (j), the stabilization energy E is associated with i to j delocalization which is given by the following equation:

$$E(2) = \Delta E_{ij} = \frac{q_i (F_{ij})^2}{\varepsilon_i - \varepsilon_j} \tag{3.30}$$

Where q_i is the ith donor orbital occupancy, ε_i, ε_j are diagonal elements (orbital energies) and F_{ij} is off diagonal elements associated with NBO Fock matrix.

If the value of E (2) is large, interaction between the electron donors and electron acceptors is intensive. Thus the donating tendency from electron donors to electron acceptors and the extent of conjugation of the whole system will be more [44].

3.7 NORMAL COORDINATE ANALYSIS

Detailed description of vibrational modes can be studied by means of normal coordinate analysis. Normal coordinate analysis (NCA) is a procedure for calculating the vibrational frequencies which relates the observed frequencies of preferably the harmonic infrared and Raman frequencies to the force constants,

equilibrium geometry and the atomic masses of the oscillating system. Normal coordinate analysis has proven useful in assigning vibrational spectra but its predictive ability depends on having reliable intramolecular force constants.

NCA begins with formulation of Hessian matrix, which holds the second partial derivatives of the potential V with respect to displacement of the atoms in Cartesian coordinates. This process of forming an appropriate potential model and calculating the second derivatives of the potential with respect to the cartesian coordinates yields the force constant matrix, which is often, spares. Normal coordinates are obtained from the solution of the eigen value problem,

$$\left|GF - E\lambda\right| = 0 \tag{3.31}$$

Where G is the matrix representation for kinetic energy, F is the matrix representation for the potential energy (second order partial derivatives of the potential energy in equilibrium that form a square-symmetric matrix of order 3N called the force constant matrix or Hessian) and E is the unit matrix. The eigen value (λ) and eigen vectors of this matrix characterize the vibrational frequencies and displacement patterns of each atom. The use of NCA has been limited to relatively small systems. This is partly due to the lack of efficient numerical algorithms for large scale NCA calculation. Also, the extreme sensitivity of the Hessian matrix to even small changes in structure in a state of mechanical equilibrium, results in imaginary vibrational frequencies and the mixing of modes. Further, NCA requires the diagonalization of very large sparse matrices which demands excessive computational effort.

NCA is commonly employed nowadays as an aid in the interpretation of the vibrational spectra of large molecules. In order to get meaningful results, knowledge

of vibrational force field is necessary. Since the number of force constants grows quadratically with the number of atoms, one has to employ many approximations in the calculation of harmonic force field even for moderately large molecules. To overcome this difficulty, one can determine a force field even for a set of related molecules using the overlay method introduced by Snyder and Schachtschneider in the 1960's [45]. Gwinn developed a program for normal coordinate analysis using mass weighted Cartesian coordinates, which eliminates the redundancy problems arising when internal valence coordinates are used as in Wilson's GF-method [46, 47]. MOLVIB is a FORTRAN program for the calculation of harmonic force fields and vibrational modes of molecules with a maximum of 50 atoms. All calculations are performed in terms of mass weighted Cartesian coordinated instead of internal coordinates as in the conventional GF-method. This overcomes the problem with redundant coordinates. The force field is refined by a modified least square fit of observed normal frequencies as described by Sundius [48].

In MOLVIB, three methods are available for the scale factor calculations. In two of these methods, the non diagonal terms in the potential energy will depend non-linearly on the scale factors as,

$$2V = \sum_i S_i f_{ij} q_i q_j + \sum_i \sum_{j \neq 1} \sqrt{S_i} S_j f_{ij} q_i q_j \qquad (3.32)$$

The factor $\sqrt{S_i} S_j$ that occurs in front of the non-diagonal force constant has to be repeated. The frequency fit usually converges in four or five iterations, and often just a few repetitions are necessary. The initial values for the scale factors are set to 1. Frequently the force constants are not defined in terms of Cartesian coordinates, but are rather expressed in internal coordinates, which are more natural from a chemical

point of view, because they can be directly related to molecular bond lengths and bond angles [49]. The relation between the internal coordinate's s_i and the Cartesian displacement coordinates x_i can be expressed in matrix form as

$$S=Bx \tag{3.33}$$

Usually the matrix B is rectangular, because the number of internal coordinates is seldom identical to the number of Cartesian displacement coordinates. This means that it is impossible to invert the transformation matrix. The normal modes can be conveniently characterized with matrix, which gives the transformation from the normal to internal coordinates. With the aid of B-matrix and the eigenvector matrix U the L matrix can be calculated according to the formula

$$L = BM^{-1/2}U \tag{3.34}$$

To facilitate the interpretation of the normal modes, the columns are sometimes normalized to 1.

$$P_{ij} = \left(L_{ji}\right)^2 F_{jj} / \lambda_i . \tag{3.35}$$

The P.E.D., which gives the fractional contribution of the diagonal matrix elements of F to the normal modes, is used for a complete symmetry classification of the normal vibrations of the molecule.

3.8 FRONTIER MOLECULAR ORBITAL ANALYSIS (FMO ANALYSIS)

The effect of highest molecular orbital and lowest unoccupied molecular orbital on reaction mechanism, led to Frontier Molecular orbital theory (FMO Theory). Fukui realized good approximation foe reactivity could be found by looking at the frontier orbital (HOMO/LUMO) [50]. This was based on three main observations of molecular orbital theory as two moleculess interact.

1. The occupied orbital of different molecules repel each other
2. Positive charges of one molecule attract the negative charges of other.
3. The occupied orbital of one molecule and the unoccupied orbital of the other interact with each other causing attraction.

FMO theory simplifies reactivity to interactions between the HOMO of one species and the LUMO of the other. Analysis on the HOMO-LUMO energy gap gives knowledge about the intermolecular charge transfer. The information about the molecular stability can also be derived from the investigation of HOMO-LUMO gap. The higher the energy of a HOMO, the easier the electrons can be removed from it. The lower the energy of a LUMO, the easier the electron can be transferred to it [51].

3.9 NATURAL POPULATION ANALYSIS

Natural population analysis was developed by Reed *et al* to solve the problems due to Mulliken populations. The natural analysis is an alternative to conventional Mulliken population analysis, and seems to exhibit improved numerical stability and describe the electron distribution in compounds of high ionic character. Natural populations, n_i (A) are the occupancies of the natural atomic orbital. These rigorously satisfy the Pauli's exclusion principle: $0 < n_i (A) < 2$. The population of an atom n (A) is the sum of natural populations $n (A) = \sum n_i (A)$. A distinguished feature of the NPA method is that it largely resolves the basis set dependence problem encountered in the Mulliken population analysis method. An extension of NPA is natural population analysis which partitions the NPA charges into core orbital, bonding orbital, lone pairs, and Rydberg states. The natural

populations are summarized as an effective valence electron configuration (Natural Electron Configuration) for each atom. In the molecular environment the occupancies of the atomic orbital are non-integer and the effective atomic configuration can be comparable to idealize atomic states in promoted configurations. The analysis of the electron density distribution in a molecular system is based on the orthonormal natural atomic orbitals. Natural populations are the occupancies of the natural atomic orbitals. The population of an atom is the sum of natural populations.

3.10 HIRSFELD SURFACE ANALYSIS

The Hirshfeld surface of a molecule in a crystal is constructed by partitioning space in the crystal into regions where the electron distribution of a sum of spherical atoms for the molecule (the pro molecule) dominates the corresponding sum over the crystal (the pro crystal) [52]. Following Hirsfeld, [53] we define a molecular weight function w(r):

$$w(r) = \frac{\rho_{promolecule}(r)}{\rho_{procrystal}(r)} \tag{3.36}$$

$$w(r) = \frac{\sum_{A \in molecule} \rho_A(r)}{\sum_{A \in crystal} \rho_A(r)} \tag{3.37}$$

ρ_A (r) is a spherically averaged atomic electron density centered on nucleus A, and the pro molecule and pro crystal are sums over the atoms belonging to the molecule and to the crystal respectively. This method is increasingly popular in a discussion of intermolecular interactions in a molecular crystals, and is based on the calculation of the pro molecular electron density both crystal and in gas phase.

3.10.1 2-D Finger Print Plots

The 2-D finger print plot of the Hirsfeld surface represents a truly novel method for summarizing the complex information contained in a molecular crystal structure into a single, unique full color plot, which provides a 'Finger print' of the intermolecular interactions in the crystal. Derived from the Hirsfeld surface, these 2-D finger print plots provide a visual summary of the frequency of each combination of d_e and d_i across the surface of the molecule, so they not only indicate which intermolecular interactions are present, but also the relative area of the surface corresponding to each kind of interaction. Directions and strength of intermolecular interactions within the crystals were mapped onto HS using the descriptor d_{norm} (normalized contact distance). The value of d_{norm} is negative or positive when intermolecular contacts are shorter/longer than υdw separations [54]. Besides crystal explorer program has been used to obtain 2-D finger print plots (FPs), which are generated, based on the d_e and d_i distances. 2-D finger print plots are derived from the HS by plotting the fraction of points on the surface as a function of the pair (d_i, d_e). Each point on the standard 2-D graph represents a bin formed by discrete intervals of d_i and d_e (0.01x0.01Å), and the points are colored as a function of the fraction of surface points in that bin, with a range from blue (few points) through green (moderate fraction) [55].

3.11 MOLECULAR DOCKING

Molecular docking is a method to predict the preferred orientation of one molecule to a second when bound to each other to form a stable complex. Softwares are used to predict or stimulate the possible reaction (and interactions) between the

molecules based on their 3-D structures. Using software, the interactions can be viewed and analyzed to understand and answers some biological important questions regarding certain chemical or biological reaction. This method can therefore be used not only to predict possible binders or inhibitors, but also to predict how strong the association between the molecules (called the binding affinity) can be. Prediction of the binding affinity will be useful for synthesized compounds where the affinity of the desired compound towards a certain target (protein/enzyme) can be predicted. There are two main types of docking in practice

Molecule-Protein (Ligand-Protein docking)

Protein-Protein docking

Protein-Protein docking involves two protein molecules simulated by the computer or computer program to bind or interact with one another. However protein-protein interactions are basically rigid compared to the Ligand-Protein docking. By stimulating certain protein-protein interactions in a specific biological reaction, the insight at the molecular level on a certain mechanism can be got.

3.11.1 In Silico Docking

Computational docking was performed to analysis the binding pattern of the taken compounds. 3-D representations of the compounds were retrieved from ChemSpider (ID 4503582) and 3-D X-ray crystallography structure of androgen receptor retrieved from protein databank (PDB ID: 3rll). AutoDock Tools 1.5.4 was then used to assign hydrogen, Gasteiger charges and rotatable bonds to ligand [56]. Polar hydrogen were added to receptor pdb file (protein), water molecules and co crystallized ligand were deleted. Active site of the receptor was defined in a way that

it included residues of the active site within the grid dimensions of 30Å*30Å*30Å. The most popular algorithm, Lamarckian Genetic Algorithm (LGA) available in Autodock Vina was employed for docking [57]. The active site of the ligand–protein interaction was visualized using Discover Studio Visualizer 4.0.

3.12 QUANTUM THEORY OF ATOMS IN MOLECULES

Quantum Theory of Atoms in Molecules (QTAIM) is suitable for the investigation of molecules interaction with their direct environment. The Laplacian surface of the electron density is a functional useful for characterization of the bonds pattern but also for probing reactive sites in a molecule indicating the degree of susceptibility to electrophilic or nucleophilic attack. However, the QTAIM based formalism reveals the nature and strength of the interactions between neighboring atoms. The QTAIM calculations allowed getting the values of electron density ρ and the corresponding Laplacian $\nabla \rho$ for all BCPs of covalent bonds.

The electron density distribution $\rho(r)$ is treated as a scalar field and examined by the analysis of the accompanying gradient $\nabla \rho(r)$ vector field. Depending on the nature of the stationary points (maxima, saddle points, or minima in the electron density), it describes core-, bond-, ring-, and cage critical points and denoted as nuclear attractor critical point (NACP), local maximum of electron density, bond critical point (BCP), minimum in the direction of the nucleus but it is a maximum in another main direction, ring critical point (RCP), minimum in two principal axes, and cage critical point (CCP), local minimum of electron density, respectively. Bader classifies covalent and ionic interactions as shared and closed shell interactions basis of electron density and Laplace of density. The surface of

Laplacian of the electron density, which delivers better description of the reactive sites in molecule than electron density itself, was determined.

Analysis of the chemical bonding in terms of quantum theory of atoms in molecules based on the electron charge density ρ and on its topology is the most important approach to the three dimensional 3-D analysis of molecular wave functions, and the concept of delocalization index [58,59]. CPs are then designated by the pair (r, s) and the four types of rank 3 CPs in 3-D are denoted as (3, 3), (3,-1), (3, +1) and (3, +3). The (3,-3) CPs are local maxima, the (3, +3) CPs local minima, and the (3,-1) or the (3, +1) CPs are saddle points in one or two dimensions. According to Bader' topological theory of atoms in molecules, occurrence of (3,-1) critical point between any atoms in a system under equilibrium is necessary and sufficient for the chemical bonding between these atoms, and this point in this case is referred to as a bond critical point (BCP). The potential electron energy densities at critical point are correlated with the intra/inter molecular hydrogen bond by the expression $E_{IHB} = 1/2V$ (BCP) [60].

3.13 QUANTITATIVE STRUCTURE- ACTIVITY RELATIONSHIPS

QSAR is used for quantitative correlation of physicochemical attributes of drug like molecules ligand [61]. It has transformed into an extensively used tool, significantly contributing to the drug discovery process. In QSAR modeling, the predictors consist of physico-chemical properties or theoretical molecular descriptors of chemicals; the QSAR response-variable could be a biological activity of the chemicals. QSAR model summarizes a relationship between chemical structures and biological activity in a chemical. QSAR methods can be generally

defined as the applications of mathematical and statistical methods to the problem finding empirical relationship of the form

$$P_i = k\ (D_1, D_2 \ldots\ldots D_n) \qquad (3.38)$$

Where, P_i is the biological activity or any other property of interest of the molecule. D_1, D_2, …D_n is the calculated structural property of the compound and k is some empirically established mathematical transformation that should be applied to descriptors to calculate property value for all the molecules. QSAR studies are aimed at developing models for quantitative and graded response data, respectively. Regression –based approaches are employed when the response data of chemicals are entirely numerical, while qualitative or semi quantitative chemical responses are modeled using classification techniques.

The most commonly used regression –based approaches are

Multiple linear regressions (MLR)

Partial least squares (PLS)

Some classification based approaches are

Linear discriminant analysis (LDA)

Cluster analysis

Multiple linear regressions are one of the most commonly used methods for constructing QSAR models. The generalized expression of an MLR equation will be like the following:

$$Y = a_0 + a_1 \times X_1 + a_2 \times X_2 + a_3 \times X_3 + ...a_n \times X_N \qquad (3.39)$$

In the above expression, Y is the response or dependent variable, X_1, X_2, …X_n are descriptors present in the model with the corresponding regression coefficients a_1, a_2, …a_n respectively, and a_0 is the constant term of the model. Each

regression coefficient should be significant at $p<0.05$ which can be checked from a't' test. The descriptor present in MLR model should not be much inter-correlated. The multiple correlation coefficient (R) can be defined as:

$$R^2 = 1 - \frac{\sum \left(Y_{obs} - Y_{calc}\right)^2}{\sum \left(Y_{obs} - \overline{Y_{obs}}\right)^2} \tag{3.40}$$

In the above equation, Y_{obs} stands for the observed response value, while Y_{calc} is the model-derived calculated response and $\overline{Y_{obs}}$ is the average of the observed response values. For ideal model, the sum of squared residuals being 0, the value of R^2 is 1. As the value of R^2 deviates from 1, the fitting quality of the model deteriorates.