Module 2 Spectroscopic techniques

Lecture 3 Basics of Spectroscopy

Spectroscopy deals with the study of interaction of electromagnetic radiation with matter. Electromagnetic radiation is a simple harmonic wave of electric and magnetic fields fluctuating orthogonal to each other (Figure 3.1A).

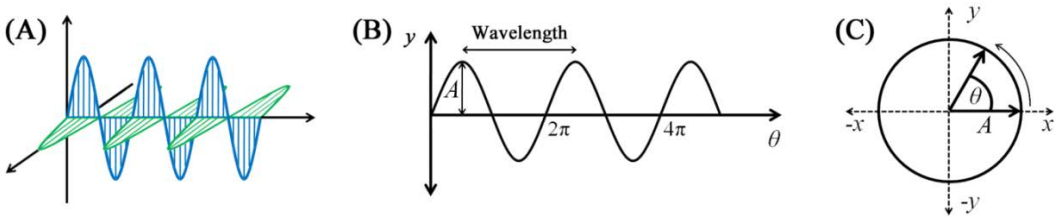


Figure 3.1: An electromagnetic wave showing orthogonal electric and magnetic components (A); a sine wave (B); and uniform circular motion representation of the sine function (C).

A simple harmonic function can be represented by a sine wave (Figure 3.1B):

$$y = A \sin \theta \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (3.1)$$

Sine wave is a periodic function and can be described in terms of the circular motion (Figure 3.1C). The value of *y* at any point is simply the projection of vector *A* on the y-axis, which is nothing but *A sinθ*. Equation (1) can therefore be written in terms of angular velocity, ω.

$$y = A \sin(\omega t) \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (3.2)$$

$$y = A \sin(2\pi \nu t) \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (3.3)$$

$$y = A \sin(2\pi \nu \frac{z}{c}) \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (3.4)$$

where, *z* = displacement in time *t* and *c* is the velocity of the electromagnetic wave

If the wave completes ν cycles/s and the wave is travelling with a velocity *c* metres/ sec, then the wavelength of the wave must be $\frac{c}{\nu}$ metres.

$$y = A \sin(\frac{2\pi z}{\lambda}) \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (3.5)$$

Energy of electromagnetic radiation:

Energy of an electromagnetic radiation is given by

$$E = h\nu = h\frac{c}{\lambda} \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (3.6)$$

where *h* is Planck's constant and has a value of 6.626×10^{-34} $m^2 \cdot kg \cdot s^{-1}$. Based on the energy, electromagnetic radiation has been divided into different regions. The region of electromagnetic spectrum human beings can see, for example, is called visible region or visible spectrum. The visible region constitutes a very small portion of the electromagnetic spectrum and corresponds to the wavelengths of ~400 – 780 nm (Figure 3.2). The energy of the visible spectrum therefore ranges from $\sim 2.5 \times 10^{-19}$ to $\sim 5 \times 10^{-19}$ Joules. It is not convenient to write such small values of energy; the energies are therefore written in terms of electronvolts (eV). One electronvolt equals 1.602×10^{-19} Joules. Therefore, the energy range of the visible spectrum is ~1.6 – 3.1 eV. Spectroscopists, however, prefer to use wavelength (λ) or frequency (ν) or wavenumber ($\bar{\nu}$) instead of energy.

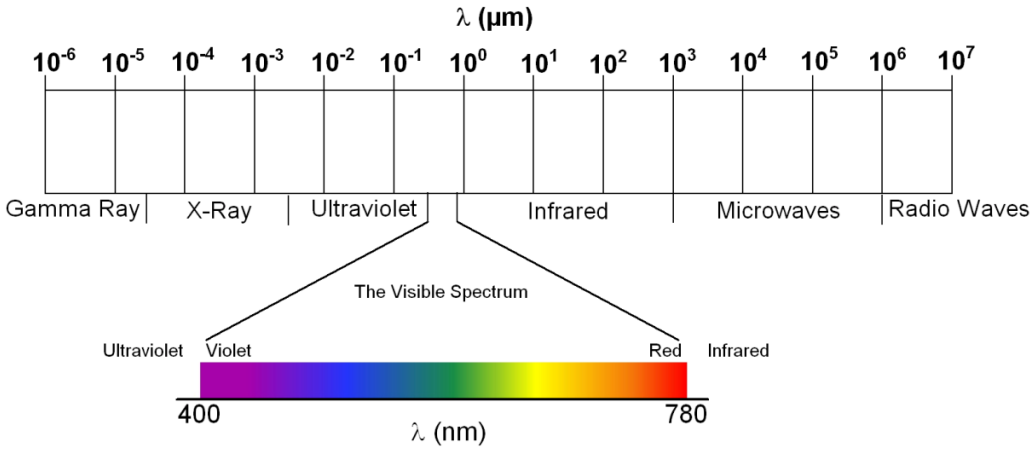


Figure 3.2 Electromagnetic spectrum

Quantization of energy:

As put forward by Max Planck while studying the problem of Blackbody radiation in early 1900s, atoms and molecules can absorb or emit the energy in discrete packets, called quanta (singular: quantum). The quantum for electromagnetic energy is called a photon which has the energy given by equation 3.6. A molecule can possess energies

in different forms such as vibrational energy, rotational energy, electronic energy, etc. Introduction to the structure of an atom in a General Chemistry course mentions about the electrons residing in different orbits/orbitals surrounding the nucleus, typically the first exposure to the discrete electronic energy levels of atoms. In much the same way, rotational and vibrational energy levels of molecules are also discrete. A molecule can jump from one energy level to another by absorbing or emitting a photon of energy that separate the two energy levels (Figure 3.3).

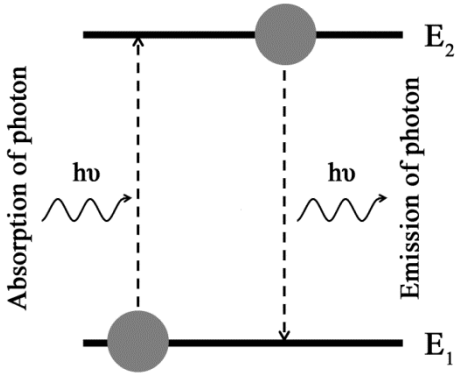


Figure 3.3 Transitions of a molecule between energy levels, E_1 and E_2 by absorbing/emitting the electromagnetic radiation

Electromagnetic spectrum and the atomic/molecular processes:

Molecules undergo processes like rotation, vibration, electronic transitions, and nuclear transitions. The energies underlying these processes correspond to different regions in the electromagnetic spectrum (Figure 3.4):

i. Radiofrequency waves: Radiofrequency region has very low energies that correspond to the energy differences in the nuclear and electron spin states. These frequencies, therefore, find applications in nuclear magnetic resonance and electron paramagnetic resonance spectroscopy.

ii. Microwaves: Microwaves have energies between those of radiofrequency waves and infrared waves and find applications in rotational spectroscopy and electron paramagnetic resonance spectroscopy.

iii. Infrared radiation: The energies associated with molecular vibrations fall in the infrared region of electromagnetic spectrum. Infrared spectroscopy is therefore also known as vibrational spectroscopy and is a very useful technique for functional group identification in organic compounds.

iv. UV/Visible region: UV and visible regions are involved in the electronic transitions in the molecules. The spectroscopic methods using UV or visible light therefore come under 'Electronic spectroscopy'.

v. X-ray radiation: X-rays are high energy electromagnetic radiation and causes transitions in the internal electrons of the molecules.

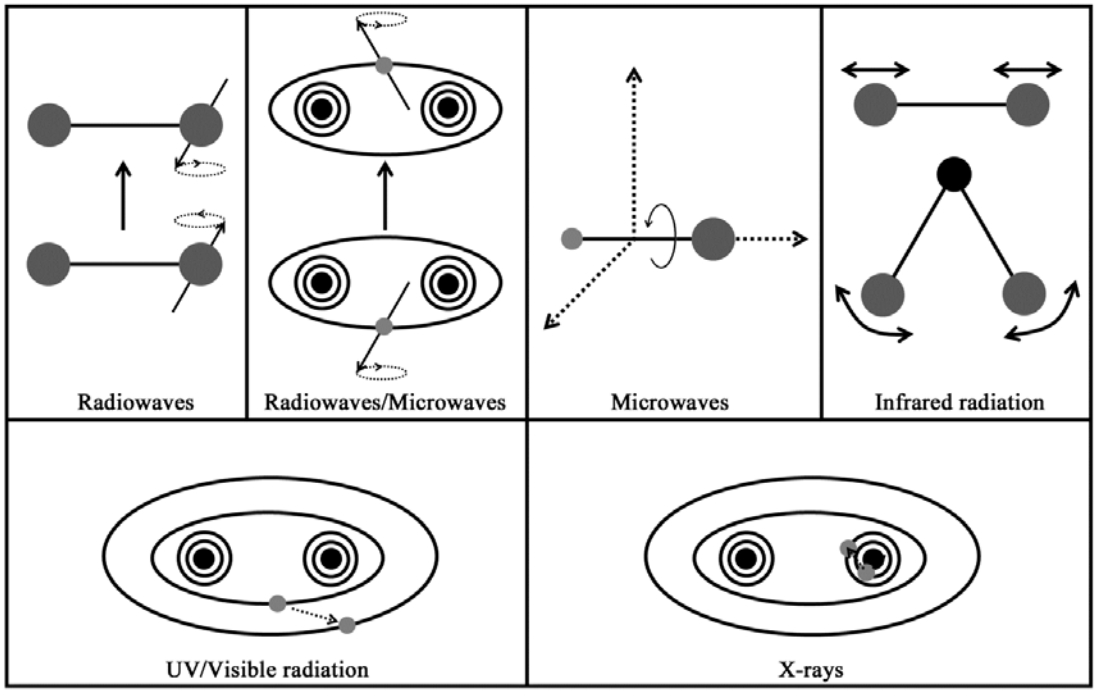


Figure 3.4 The range of atomic/molecular processes the electromagnetic radiation is involved in.

Mechanisms of interaction of electromagnetic radiation with matter:

In order to interact with the electromagnetic radiation, the molecules must have some electric or magnetic effect that could be influenced by the electric or magnetic components of the radiation.

i. In NMR spectroscopy, for example, the nuclear spins have magnetic dipoles aligned with or against a huge magnetic field. Interaction with radiofreqency of appropriate energy results in the change in these dipoles.

ii. Rotations of a molecule having a net electric dipole moment, such as water will cause changes in the directions of the dipole and therefore in the electrical properties (Figure 3.5A and B). Figure 3.5B shows the changes in the y-component of the dipole moment due to rotation of water molecule.

iii. Vibrations of molecules can result in changes in electric dipoles that could interact with the electrical component of the electromagnetic radiation (Figure 3.5C).

iv. Electronic transitions take place from one orbital to another. Owing to the differences in the geometry, size, and the spatial organization of the different orbitals, an electronic transition causes change in the dipole moment of the molecule (Figure 3.5D).

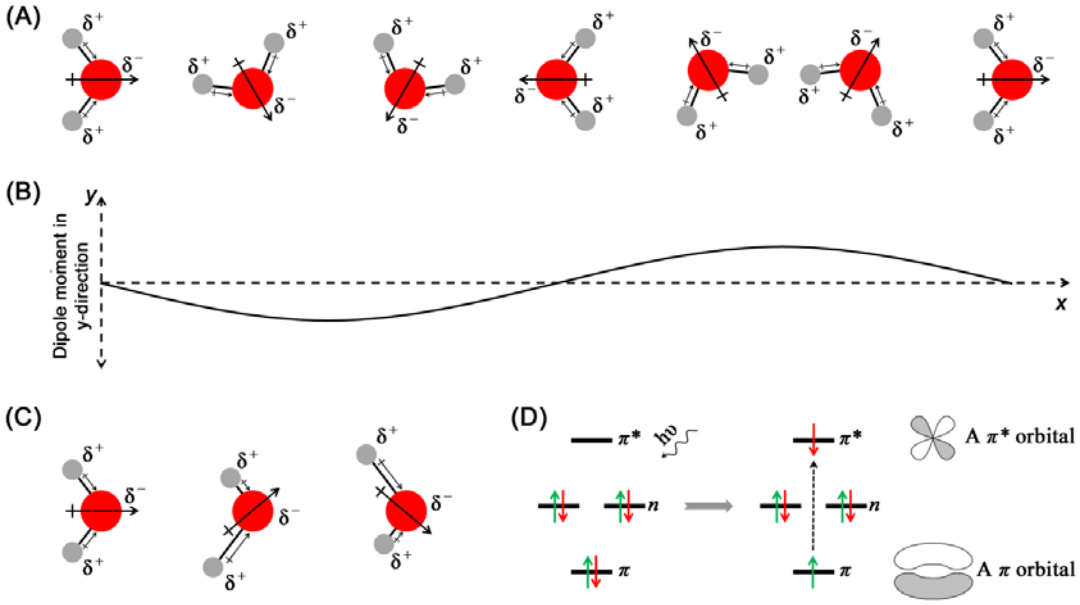


Figure 3.5 Panel A shows the rotation of a water molecule around its centre of mass (A). The change in the dipole moment as a result of rotation in plotted in panel B. Panel C shows the change in dipole moment of water due to asymmetric stretching vibrations of O—H bond. Panel D shows an electronic transition from π to π^* orbital and the geometry of the two orbitals.

The above examples suggest that a change in either electric or magnetic dipole moment in a molecule is required for the absorption or emission of the electromagnetic radiation.

THINK TANK??

Will there be a change in the dipole moment if there is symmetric stretching of O—H bond in water molecule?

Absorption peaks and line widths:

Absorption of radiation is the first step in any spectroscopic experiment. Absorption spectra are routinely recorded for the electronic, rotational, and vibrational spectroscopy. It is therefore important to see how an absorption spectrum looks like.

As we have already seen, a transition between states takes place if the energy provided by the electromagnetic radiation equals the energy gap between the two states *i.e.* $\Delta E = h\upsilon = \frac{hc}{\lambda}$. This implies that the molecule precisely absorbs the radiation of wavelength, λ and ideally a sharp absorption line should appear at this wavelength (Figure 3.6A). In practice, however, the absorption lines are not sharp but appear as fairly broad peaks (Figure 3.6B) for the following reasons:

i. Instrumental factors: The slits that allow the incident light to impinge on the sample and the emerging light to the detector have finite widths. Consider that the transition occurs at wavelength, λ_t. When the wavelength is changed to $\lambda_t + \Delta\lambda$ or $\lambda_t - \Delta\lambda$, the finite slit width allows the radiation of wavelength, λ_t to pass through the slits and a finite absorbance is observed at these wavelengths. The absorption peaks are therefore symmetrical to the line at $\lambda = \lambda_t$.

ii. Sample factors: Molecules in a liquid or gaseous sample are in motion and keep colliding with each other. Collisions influence the vibrational and rotational motions of the molecules thereby causing broadening. Two atoms/molecules coming in close proximity will perturb the electronic energies, at least those of the outermost electrons resulting in broadening of electronic spectra. Motion of molecules undergoing transition also causes shift in absorption frequencies, known as Doppler broadening.

iii. Intrinsic broadening: Intrinsic or natural broadening arises from the Heisenberg's uncertainty principle which states that the shorter the lifetime of a state, the more uncertain is its energy. Molecular transitions have finite lifetimes, therefore their energy is not exact. If Δt is the lifetime of a molecule in an excited state, the uncertainty in the energy of the states is given by:

$$\Delta E \times \Delta t \geq \frac{h}{4\pi} \qquad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (3.7)$$

$$\Delta E \times \Delta t \geq \frac{\hbar}{2} \qquad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (3.8)$$

where, $\hbar = \frac{h}{2\pi}$

Two more features worth noticing in the Figure 3.6B are the fluctuations in the baseline and the baseline itself, which is not horizontal. The small fluctuations in the baseline are referred to as noise. Noise is the manifestation of the random weak

signals generated by the instrument electronics. To identify the sample peaks clear of the noise, the intensity of the sample peaks has to be at least 3-4 times higher than the noise. A better signal-to-noise ratio is obtained by recorded more than one spectra and averaging; the noise being random gets cancelled out. Instrumental factors are responsible for the non-horizontal baseline observed in Figure 3.6B: The light sources used in the instruments emit radiations of different intensities at different wavelengths and usually the detector sensitivity is also wavelength-dependent. A reasonable horizontal baseline for the samples can easily be obtained by subtracting the spectrum obtained from the solvent the sample is dissolved in.

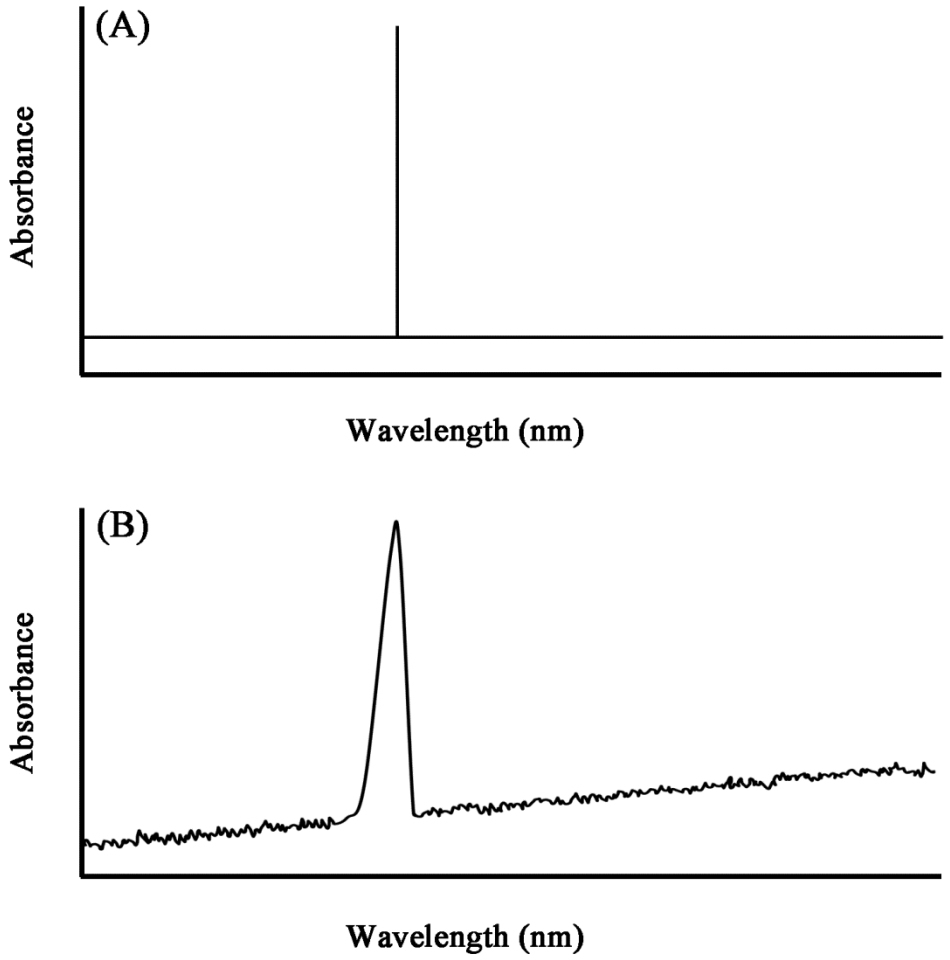


Figure 3.6 An idealized spectrum for a single wavelength transition (A) and an experimentally obtained spectrum (B)

Other features of the spectroscopy and the spectra obtained will be discussed as and when they arise in the following lectures.

Lecture4 UV/Visible Absorption Spectroscopy-I

We see a lot of colorful things around us. What exactly is the color and what make the things exhibit these colors? We know that the color we see is the visible region of the electromagnetic spectrum. We also know that matter can absorb the electromagnetic radiation of different energy (or wavelengths). The region of electromagnetic energy that is not absorbed is simply reflected back or getstransmitted through the matter. The colored compounds are colored because they absorb the visible light. The color that is perceived is called the complement color to the absorbed wavelength and is represented by a color wheel (Figure 4.1).

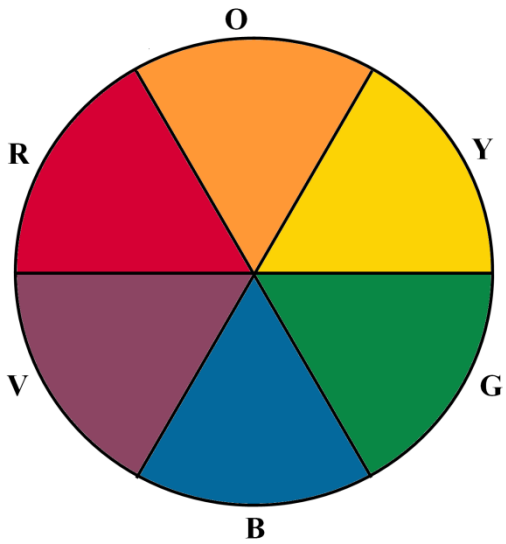


Figure 4.1 A simplified color wheel showing complementary colors. Green is interesting as it can arise from the absorption of radiation to either end of the visible spectrum.

Absorption of ultraviolet (UV) and visible radiation is one of the most routinely used analytical tools in life sciences research. The simplest application of UV/Visible radiation is to quantify the amount of a substance present in a solution.UV region of electromagnetic radiation encompasses the wavelengths ranging from ~10 nm – ~400 nm while visible region encompasses the wavelengths from ~400 nm – ~780 nm. For the sake of convenience in discussing the observations, UV region is loosely divided into near UV (wavelength region nearer to the visible region, λ ~ 250 nm – 400 nm), far UV region (wavelength region farther to the visible region, λ ~ 190 nm – 250 nm) and vacuum UV region (λ< 190 nm). The wavelength ranges defined for these regions are not strict and people use slightly different ranges to define these regions. We shall, however, stick to the wavelengths defined here. As has been discussed in the previous lecture, the absorption of UV and visible light is through the transition of an electron in the molecule from lower to a higher energy molecular orbital. The various electronic transitions observed in organic compound are shown in Figure 4.2.

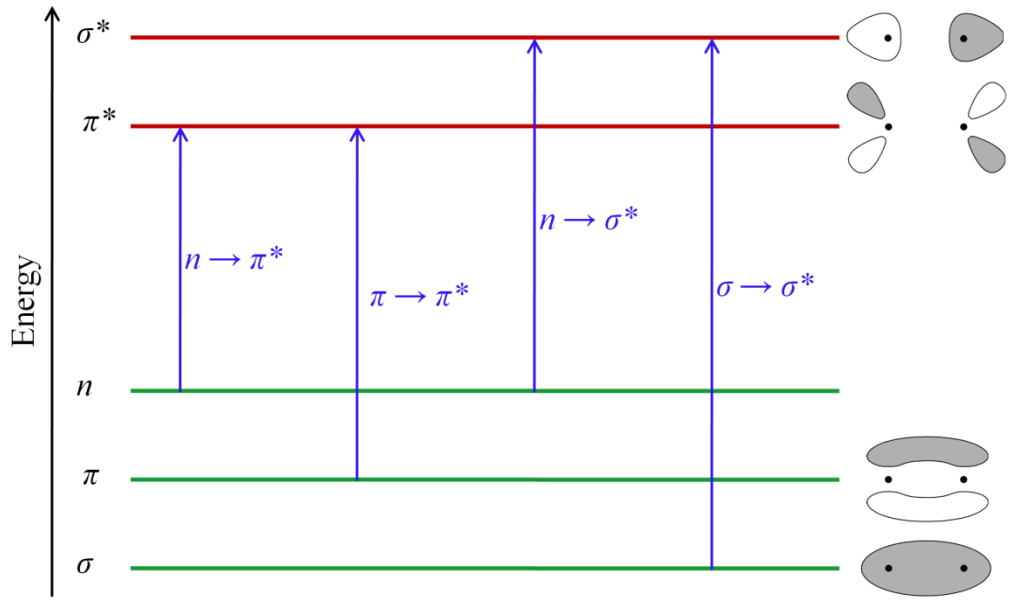


Figure 4.2 Schematic diagram showing energy levels of different orbitals and possible absorption transitions

As shown in figure 4.2, $\sigma \rightarrow \sigma^*$ transition is a high energy process and therefore lies in the vacuum UV region. Alkanes, wherein only $\sigma \rightarrow \sigma^*$ transition is possible show absorption bands ~150 nm wavelength. Alkenes haveπ and π^* orbitals and can show several transition; the lowest-energy transition, $\pi \rightarrow \pi^*$ gives an absorption band ~170-190 nm for non-conjugated alkenes (effects of conjugation on electronic transitions are discussed later). The presence of nonbonding electrons in a molecule

further expands the number of possible transitions. The entire molecule, however, is not generally involved in the absorption of the radiation in a given wavelength range. In an aliphatic ketone, for example, the absorption band around 185 nm arises due to the $\pi \rightarrow \pi^*$ transition in the carbonyl group. Atoms that comprise the molecular orbitals involved in the electronic transitions constitute the molecular moiety that is directly involved in the transition. Such a group of atoms is called a **chromophore**. A structural modification in a chromophore is generally accompanied by changes in the absorption properties.

Instrumentation:

Figure 4.3A shows a schematic diagram of a single-beam spectrophotometer. The light enters the instrument through an entrance slit, is collimated and focused on to the dispersing element, typically a diffraction grating. The light of desired wavelength is selected simply by rotating the monochromator and impinged on the sample. The intensity of the radiation transmitted through the sample is measured and converted to absorbance or transmittance (discussed later). Double beam spectrophotometers overcome certain limitations of the single beam spectrophotometers and are therefore preferred over them. A double beam spectrophotometer has two light beams, one of which passes through the sample while other passes through a reference cell (Figure 4.3B). This allows more reproducible measurements as any fluctuation in the light source or instrument electronics appears in both reference and the sample and therefore can easily be removed from the sample spectrum by subtracting the reference spectrum. Modern instruments can perform this subtraction automatically. The most commonly used detectors in the UV/Visible spectrophotometers are the photomultiplier tubes (PMT). Modern instruments also use photodiodes as the detection systems. These diodes are inexpensive and can be arranged in an array so that each diode absorbs a narrow band of the spectrum. Simultaneous recording at multiple wavelengths allows recording of the entire spectrum at once. The monochromator in these spectrophotometers is placed after the sample so that the sample is exposed to the entire spectrum of the incident radiation and the transmitted radiation is dispersed into its components.

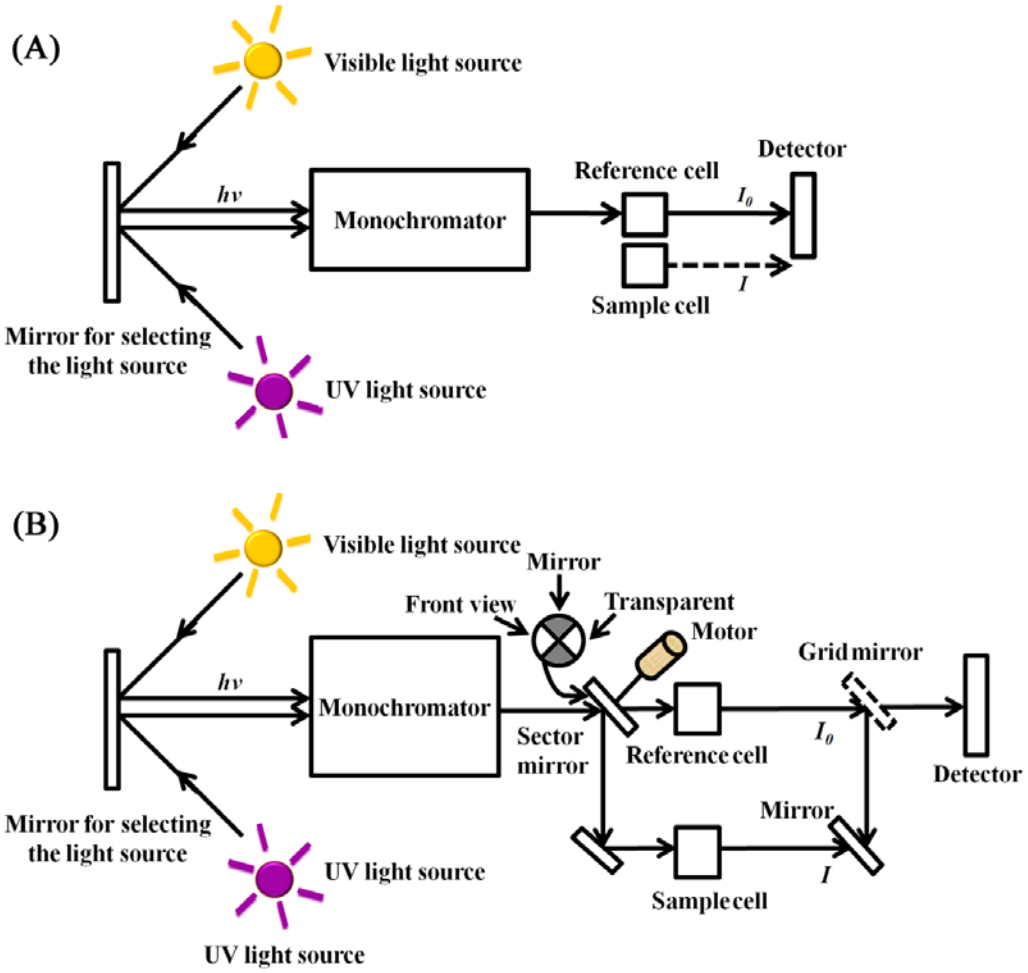


Figure 4.3 Schematic diagram showing a single beam (A) and a double beam (B) spectrophotometer

Beer-Lambert Law

It is quite intuitive that a higher concentration of the absorbing species in a sample would lead to higher absorption of the light. Furthermore, the higher thickness of the sample should result in higher absorption. Consider a cell (also called cuvette) of length, *l*, containing a solution of an absorbing molecule. The absorbing species in the sample can be represented by discs of cross-sectional area, σ. Now, let us consider a slab of infinitesimal thickness, *dx* and area, A (Figure 4.4). If an incident radiation of the resonance frequency (the frequency that causes maximum transition) having intensity I_o enters the sample cell, its intensity decreases as it penetrates the sample. Let us suppose that the intensity of the radiation before entering the infinitesimal slab is I_x.

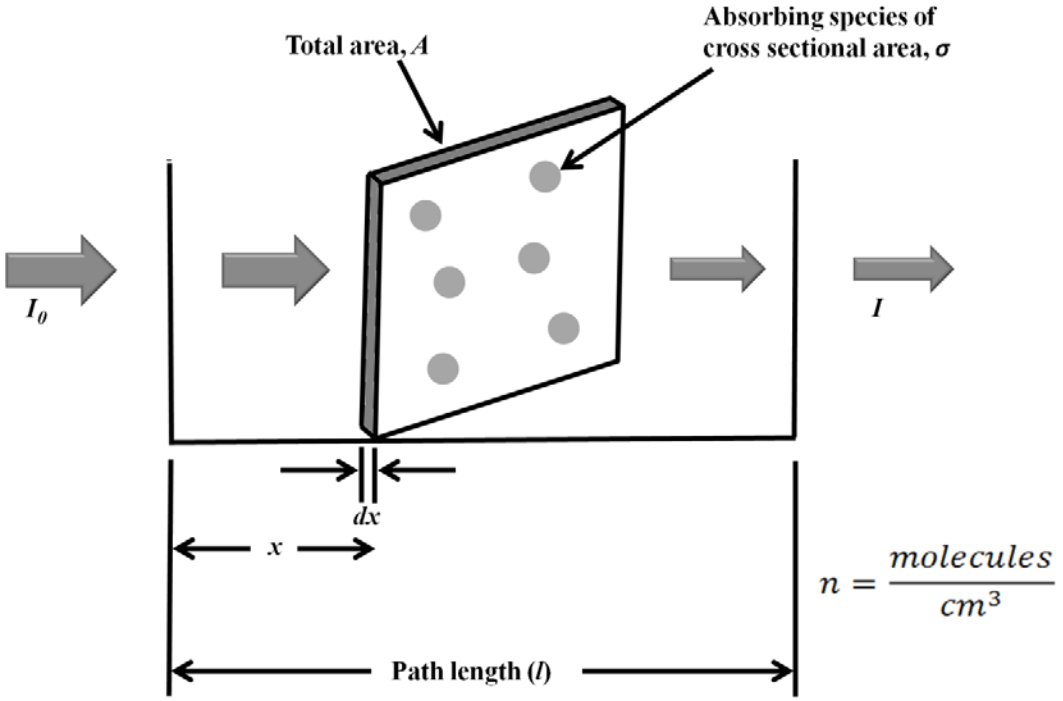


Figure 4.4A diagrammatic representation of light absorption by sample molecules in an infinitesimal thin slab within the sample

If the concentration of the absorbing molecules = $n\ \frac{molecules}{cm^3}$, the fraction of the area occupied by the molecules in the slab $= \frac{\sigma \times n \times Volume\ of\ slab}{A} = \frac{\sigma \times n \times A \times dx}{A} = \sigma \times n \times dx$

Therefore, the fraction of the photons $(\frac{dI}{I_x})$ absorbed is proportional to $\sigma \times n \times dx$. Assuming the probability of absorption if a photon strikes the molecule to be unity,

$$\frac{dI}{I_x} = -\sigma \times n \times dx \qquad \cdots\cdots\cdots\cdots\cdots\cdots (4.1)$$

The negative sign represents a decrease in intensity

Integrating equation 4.1 from $x = 0$ to $x = l$

$$\ln I|_{I_o}^{I} = -\sigma \times n \times x|_0^l \qquad \cdots\cdots\cdots\cdots\cdots\cdots (4.2)$$

$$\ln I - \ln I_o = -\sigma \times n \times l \qquad \cdots\cdots\cdots\cdots\cdots\cdots (4.3)$$

$$-\ln \frac{I}{I_o} = \sigma \times n \times l \qquad \cdots\cdots\cdots\cdots\cdots\cdots (4.4)$$

Now, the molar concentration of the molecules, *c* can be given by:

$$c\left(\frac{\text{moles}}{\text{litre}}\right) = n\left(\frac{\text{molecules}}{\text{cm}^3}\right) \times \frac{1}{6.022 \times 10^{23}}\left(\frac{\text{mole}}{\text{molecules}}\right) \times 1000\left(\frac{\text{cm}^3}{\text{litre}}\right)$$

Substituting for *n* in equation 4.4 and converting natural logarithm, *ln* into $\log_{10}$ gives:

$$-2.303 \times \log\frac{I}{I_o} = \sigma \times \{c \times 6.022 \times 10^{20}\} \times l \quad \cdots\cdots\cdots\cdots\cdots\cdots \text{(4.5)}$$

$$-\log\frac{I}{I_o} = \sigma \times c \times \left(\frac{6.022 \times 10^{20}}{2.303}\right) \times l \quad \cdots\cdots\cdots\cdots\cdots\cdots \text{(4.6)}$$

$-\log\frac{I}{I_o}$ is defined as the absorbance and $\sigma \times \left(\frac{6.022 \times 10^{20}}{2.303}\right)$ is defined as the molar absorption coefficient, denoted by the Greek alphabet, ε. Therefore, equation 4.6 can be written as:

$$\text{Absorbance, } A = \varepsilon cl \quad \cdots\cdots\cdots\cdots\cdots\cdots \text{(4.7)}$$

This equality showing linear relationship between absorbance and the concentration of the absorbing molecule (or chromophore, to be precise) is known as the Beer-Lambert law or Beer's law.

Transmittance is another way of describing the absorption of light. Transmittance (*T*) is simply the ratio of the intensity of the radiation transmitted through the sample to that of the incident radiation. Transmittance is generally represented as percentage transmittance (*%T*):

$$\%T = \frac{I}{I_o} \times 100$$

As is clear from the definition of absorbance and transmittance, both are dimensionless quantities. Absorbance and transmittance are therefore represented in arbitrary units (AU). The quantity of interest in an absorption spectrum is the molar absorption coefficient, ε which varies with wavelength (Figure 4.5). The wavelength at which highest molar absorption coefficient (ε_{max}) is observed is represented as λ_{max}. Area of cross-section of the absorbing species puts an upper limit to the molar absorption coefficient.

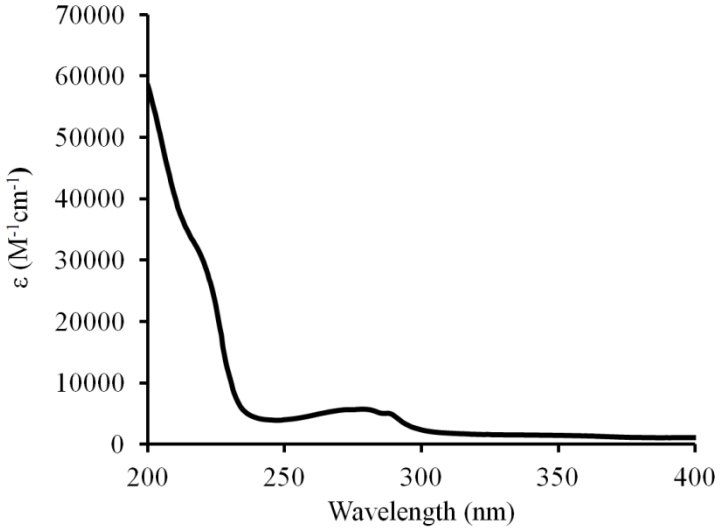


Figure 4.5 An absorption spectrum of N-acetyl-tryptophanamide

Deviations from Beer-Lambert law:

Beer-Lambert law can be used to determine the ε values of a compound by recording its absorption spectra at known concentrations. Alternatively, knowledge of ε enables the user to calculate the concentration of a compound in a given solution. It is, however, not uncommon to observe deviations from the Beer-Lambert law. Three major reasons that are responsible for the breakdown of linear relationship between absorbance and the concentration of the absorbing molecule are:

i. *High sample concentration*: The Beer-Lambert law generally holds good only for dilute solutions. At higher concentrations, the molecules come in close proximity thereby influencing their electronic properties. Although introduced as a constant at a particular wavelength for a compound, ε depends on the concentration of the compound and therefore results in deviation from linearity. At lower concentrations, however, ε can practically be assumed to be a constant.

ii. *Chemical reactions*: If a molecule undergoes a chemical reaction and the spectroscopic properties of the reacted and unreacted molecules differ, a deviation from Beer-Lambert law is observed. Change in the color of the pH indicator dyes is a classical example of this phenomenon.

iii. *Instrumental factors*: As ε is a function of wavelength, Beer-Lambert law holds good only for monochromatic light. Use of polychromatic radiation will result in deviation for linearity between absorbance and concentration.

For practical purposes, the samples giving absorbance values between 0.05 – 0.5 are considered highly reliable. At lower concentrations, the signal to noise ratio is small while at higher concentrations, absorbance values underestimate the concentration of the compound as increase in absorbance no longer matches the increase in concentration. If the absorbance values are higher, a sample can be diluted or a sample cell with smaller path length can be used; usually dilution of sample is preferred.

In the following lecture, we shall discuss the various factors that influence the absorption spectra of molecules and look at the applications of UV/Visible absorption spectroscopy for studying the biomolecules.

Lecture 5 UV/Visible Absorption Spectroscopy-II

In the previous lecture, we studied that UV/Visible radiation is absorbed by the molecules through transition of electrons in the chromophore from low energy molecular orbitals to higher energy molecular orbitals. We are interested in the transitions that lie in the far UV, near UV, and visible regions of the electromagnetic spectrum. The molecules that absorb in these regions invariably have unsaturated bonds. Plants are green due to unsaturated organic compounds, called chlorophylls. A highly unsaturated alkene, lycopene, imparts red color to the tomatoes (Figure 5.1).

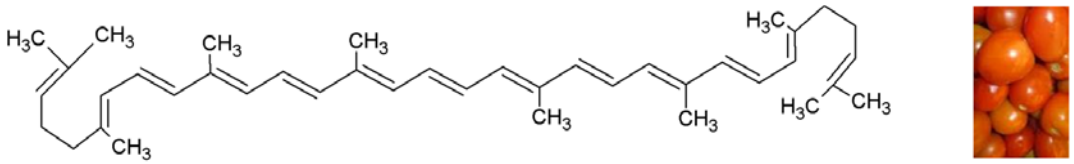


Figure 5.1 Structure of lycopene, the pigment that imparts red color to the tomatoes

As can be seen from its structure, lycopene is a highly conjugated alkene. As compared to the simple non-conjugated alkenes that typically absorb in vacuum UV region, absorption spectrum of lycopene is hugely shifted towards higher wavelengths (or lower energy). There can be factors that could shift the absorption spectra to smaller wavelengths or can increase/decrease the absorption intensity. Before understanding how conjugation causes shift in the absorption spectra, let us look at some important terms that are used to refer to the shifts in absorption spectra (Figure 5.2):

Bathochromic shift: Shift of the absorption spectrum towards longer wavelength

Hypsochromic shift: Shift of the absorption spectrum towards smaller wavelength

Hyperchromic shift: An increase in the absorption intensity

Hypochromic shift: A decrease in the absorption intensity

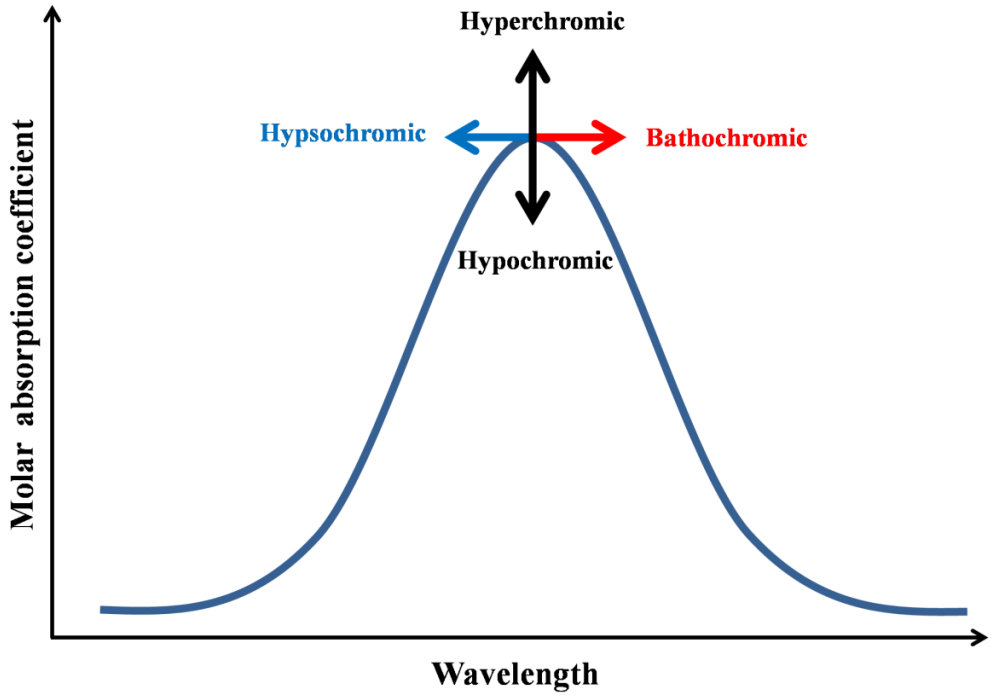


Figure 5.2 Terminology for shifts in absorption spectra

Conjugation: Conjugation brings about a bathochromic shift in the absorption bands. The higher the extent of conjugation, the more is the bathochromic shift. Such shift in absorption spectra can easily be explained using molecular orbital theory. Figure 5.3 shows the molecular orbitals drawn for ethylene; 1,3-butadiene; and 1,3,5-hexatriene on a qualitatively same energy scale for comparing their energies. As is clear from the figure, the energy differences between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) decreases as the conjugation increases. This provides an explanation as to why an electronic transition is possible at lower energy (higher wavelength) as the conjugation increases.

Energy levels of conjugated alkenes' molecular orbitals: The energy levels of the orbitals increase as the number of vertical nodes increase. The lowest energy π orbital has no nodes while the highest energy π^* orbital has $n–1$ nodes where n is the number of p–orbitals combined.

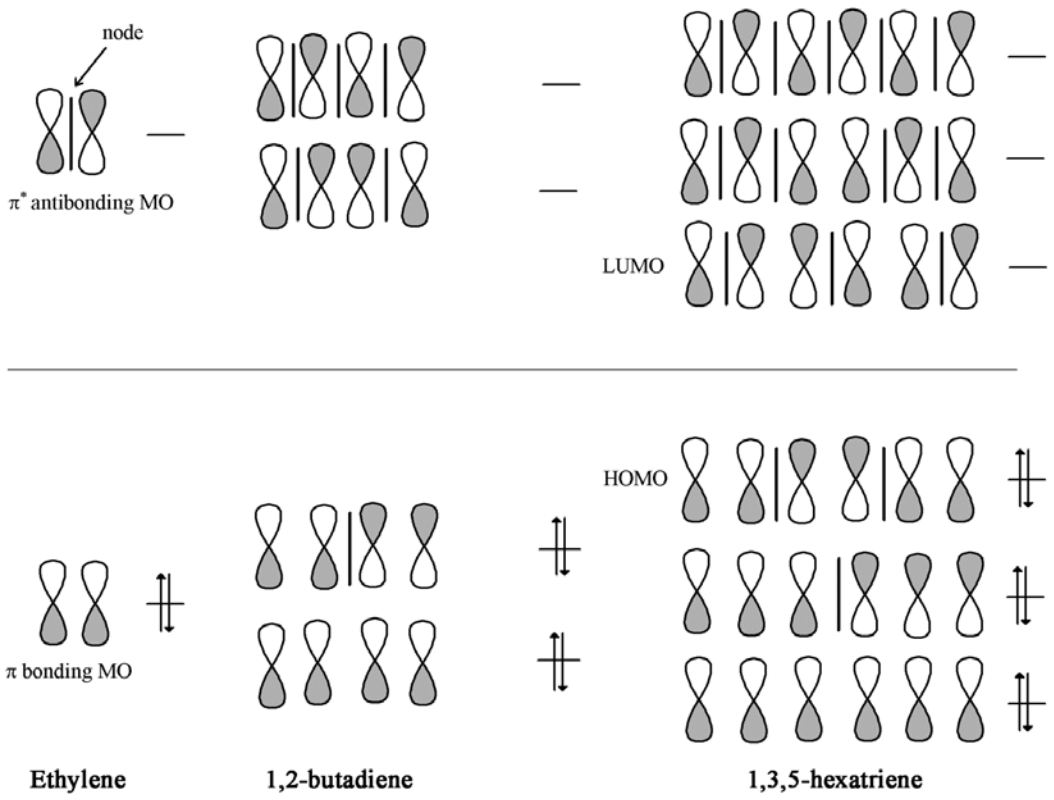


Figure 5.3 Molecular orbitals of ethylene; 1,3-butadiene; and 1,3,5-hexatriene. Notice the decrease in the energy gap of HOMO and LUMO as the conjugation increases.

Auxochrome: Auxochromes are the chemical groups that result in a bathochromic shift when attached to a chromophore. The strongest auxochromes like –OH, $–NH_2$, –OR, etc. possess nonbonding electrons. They exhibit bathochromism by extending conjugation through resonance.

The auxochrome modified chromophore is a new chromophore in real sense. The term auxochrome is therefore rarely used these days, and the entire group (basic chromophore + auxochrome) can be considered as a chromophore different from the basic chromophore. Alkyl groups also result in the bathochromic shifts in the absorption spectra of alkenes. Alkyl groups do not have non-bonded electrons, and the effect is brought about by another type of interaction called *hyperconjugation*.

Solvents: The solvents used in any spectroscopic method should ideally be transparent (non-absorbing) to the electromagnetic radiation being used. Table 5.1 shows the wavelength cutoffs (the lowest working wavelength) of some of the solvents used in UV/visible spectroscopy.

Table 5.1 Solvents commonly used in UV/visible spectroscopy

Solvent	**Wavelength cutoff**
Water	190 nm
Acetonitrile	190 nm
Cyclohexane	195 nm
Methanol	205 nm
95% ethanol	205 nm

Water, the solvent of biological systems, thankfully is transparent to the UV/visible region of interest *i.e.* the regions above $\lambda > 190$ nm. Solvents also play important role on the absorption spectra of molecules. Spectrum of a compound recorded in one solvent can look significantly different in intensity, wavelength of absorption, or both from that recorded in another. This is not something unexpected because energies of different electronic states will depend on their interaction with solvents. Polarity of solvents is an important factor in causing shifts in the absorption spectra. Conjugated dienes and aromatic hydrocarbons are little affected by the changes in solvent polarity. α,β-unsaturated carbonyl compounds are fairly sensitive to the solvent polarity. The two electronic transitions $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ respond differently to the changes in polarity. Polar solvents stabilize all the three molecular orbitals (*n, π,* and π^*), albeit to different extents (Figure 5.4). The non-bonding orbitals are stabilized most, followed by π^*. This results in a bathochromic shift in the $\pi \rightarrow \pi^*$ absorption band while a hypsochromic shift in $n \rightarrow \pi^*$ absorption band. Shift to different extents of the two bands will result in the different shape of the overall absorption spectrum.

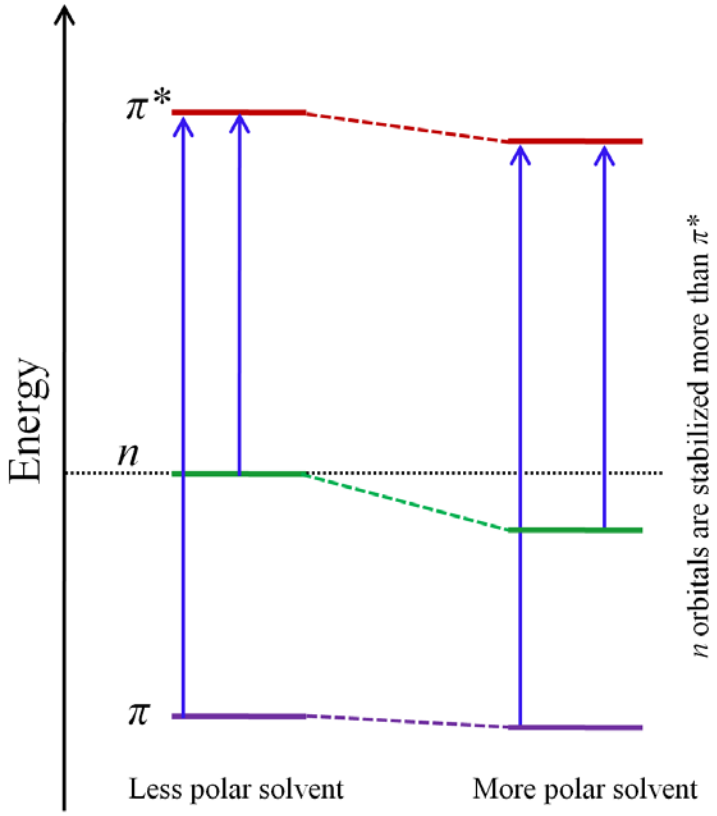


Figure 5.4 Differential stabilization of molecular orbitals in polar solvents

Biological chromophores

Amino acids and proteins: Among the 20 amino acids that constitute the proteins, tryptophan, tyrosine, and phenylalanine absorb in the near UV region. All the three amino acids show structured absorption spectra. The absorption by phenylalanine is weak with an ε_{max} of ~200 $M^{-1}cm^{-1}$ at ~250 nm. Molar absorption coefficients of ~1400 $M^{-1}cm^{-1}$ at 274 nm and ~5700 $M^{-1}cm^{-1}$ at 280 nm are observed for tyrosine and tryptophan, respectively. Disulfide linkages, formed through oxidation of cysteine resides, also contribute to the absorption of proteins in near UV region with a weak ε_{max} of ~300 $M^{-1}cm^{-1}$ around 250-270 nm. The absorption spectra of proteins are therefore largely dominated by Tyr and Trp in the near UV region. In the far UV region, peptide bond emerges as the most important chromophore in the proteins. The peptide bond displays a weak $n \rightarrow \pi^*$ transition ($\varepsilon_{max} \approx 100$ $M^{-1}cm^{-1}$) between 210-230 nm, the exact band position determined by the H-bonding interactions the peptide backbone is involved in. A strong $\pi \rightarrow \pi^*$ transition ($\varepsilon_{max} \approx 7000$ $M^{-1}cm^{-1}$) is observed around 190 nm. Side chains of Asp, Glu, Asn, Gln, Arg, His also contribute

to the absorbance in the far UV region. Figure 5.5 shows an absorption spectrum of a peptide

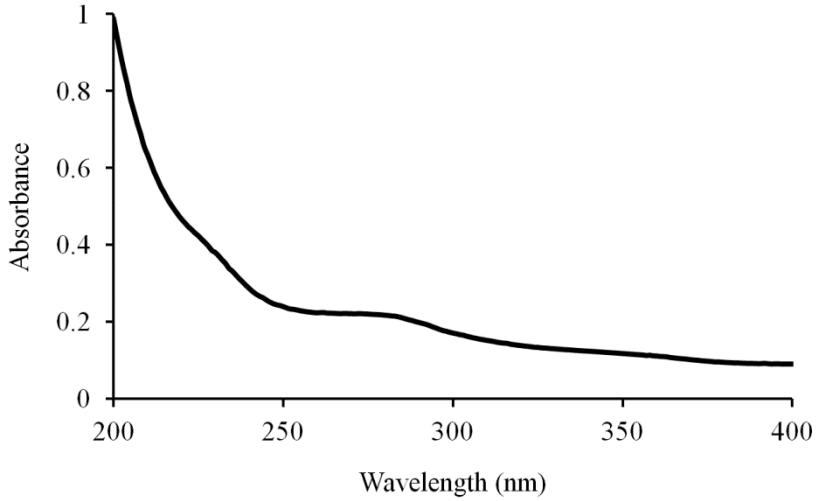


Figure 5.5 Absorption spectrum of a peptide. The absorption band ~280 nm is due to aromatic residues. Absorption band in the far UV region arises due to peptide bond electronic transitions.

Nucleic acids: Nucleic acids absorb very strongly in the far and near UV region of the electromagnetic spectrum. The absorption is largely due to the nitrogenous bases. The transitions in the nucleic acid bases are quite complex and many $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ transitions are expected to contribute to their absorption spectra. A 260 nm wavelength radiation is routinely used to estimate the concentration of nucleic acids. Though the molar absorption coefficients vary for the nucleotides at 260 nm, the average ε_{max} can be taken as ~10^4 $M^{-1}cm^{-1}$. It is important to mention that nucleotides show hyperchromicity when exposed to aqueous environment. The absorbance of the free nucleotides is higher than that of single stranded nucleic acid which is higher than that of the double stranded nucleic acid (assuming equal amount of the nucleotides present in all three).

Other chromophores: Nucleotides like NADH, NADPH, FMN, and FAD; porphyrins such as heme, chlorophylls and other plant pigments; retinal (light sensing molecule); vitamins; and a variety of unsaturated compounds constitute chromophores in the UV and visible region.

Having studied the principles of the UV/visible absorption spectroscopy and various factors that influence the electronic transitions, we can now have a look at its applications, especially the applications for analyzing the biological samples.

Applications:

i. *Determination of molar absorption coefficient*: From Beer-Lambert law, $A = \varepsilon cl$. It is therefore straightforward to calculate the molar absorption coefficient of a compound if the concentration of compound is accurately determined.

ii. *Quantification of compounds*: This is perhaps the most common application of a UV/visible spectrophotometer in a bioanalytical laboratory. If the molar absorption coefficient at a wavelength is known for the compound, the concentration can easily be estimated using Beer-Lambert law. The compounds can still be quantified if their molar absorption coefficients are not known. Estimation of total protein concentration in a given solution is an important example of this. As the given solution is a mixture of many different proteins, the ε is not available. There are, however, dyes that specifically bind to the proteins producing colored complex. The color produced will be proportional to the amount of the protein present in the solution. Performing the experiment under identical conditions using known concentrations of a protein gives a standard graph between absorbance of the dye and the amount of protein. This standard graph is then used to estimate the concentration of the given protein sample.

iii. *Quality control*: A given organic compound such as a drug can be studied for its purity. Comparison of spectrum with the standard drug will detect the impurities, if any. UV/Visible absorption is often used to detect the nucleic acid contamination in the protein preparations. Aromatic amino acids as well as the nucleotides show absorption band in the near UV region and there is a considerable overlap in the absorption spectra of aromatic amino acids and the nucleotides. A nucleic acid contamination in a protein, however, can be determined by measuring

$\frac{A_{260}}{A_{280}}$ ratio is not useful in detecting protein contaminations in DNA preparations. This is because of the large difference in molar absorption coefficients of these molecules. To cause an appreciable change in the $\frac{A_{260}}{A_{280}}$ ratio, there should a large amount of protein present.

absorbances at 260 and 280 nm. A typical nucleic acid containing all four bases shows an absorption band centered ~260 nm while a protein having aromatic amino acids shows absorption band centered ~280 nm. It is possible to determine the purity of protein preparations by recording absorbances at both 260 and 280 nm. A ratio of the absorbance at 260 nm to that at 280 nm *i.e.* $\frac{A_{260}}{A_{280}}$ is a measure of the purity.

iv. *Chemical kinetics*: UV/visible spectroscopy can be used to monitor the rate of chemical reactions if one of the reactants or products absorbs in a region where no other reactant or product absorbs significantly.

v. *Detectors in liquid chromatography instruments*: UV/visible detectors are perhaps the most common detectors present in liquid chromatography systems. Modern instruments use photodiode array detectors that can detect the molecules absorbing in different spectral regions (Figure 5.6).

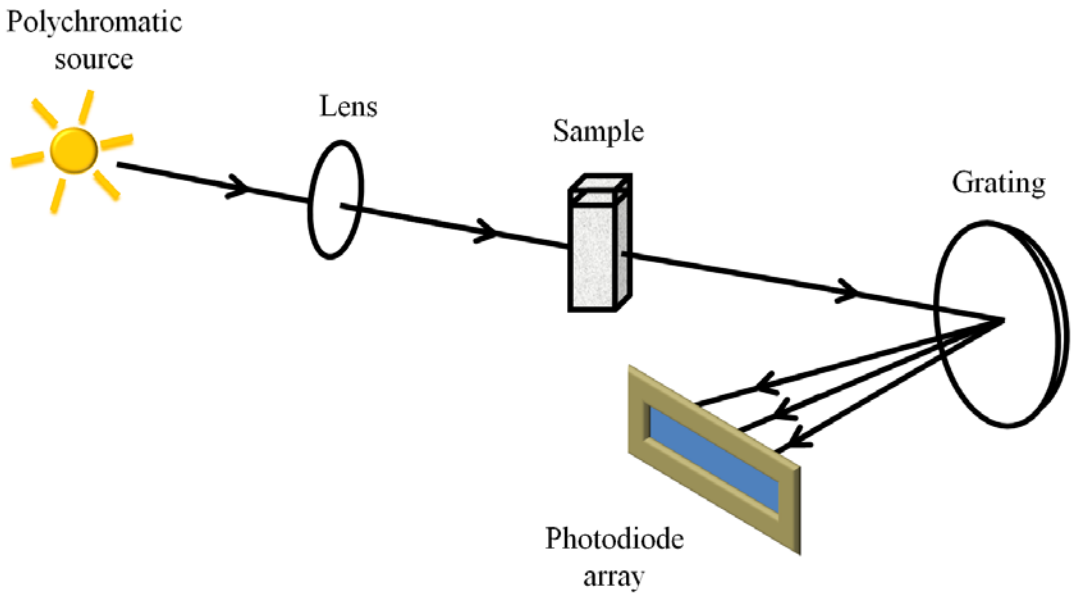


Figure 5.6 Diagram of a photodiode array detector

vi. *Determination of melting temperature of DNA*: A double stranded DNA molecule can be denatured into the single strands by heating it. Melting temperature, T_m is the temperature at which 50% of the DNA gets denatured into single strands. Denaturation of DNA is accompanied by hyperchromic shift in the absorption spectra in the near UV region. A melting curve (plot between temperature and absorbance at 260 nm) is plotted and T_m is determined (Figure 5.7).

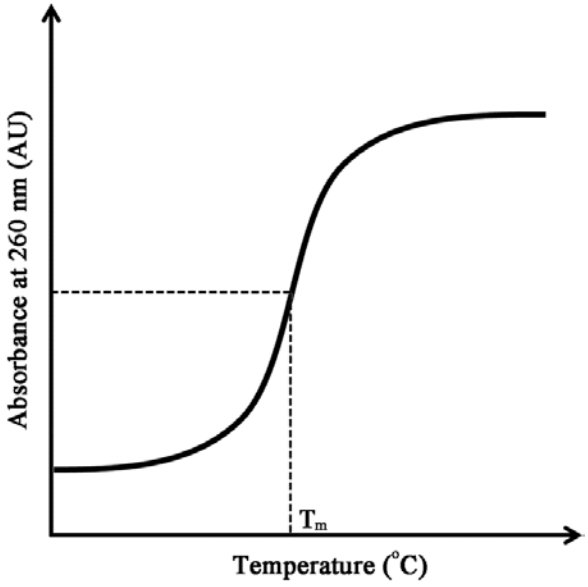


Figure 5.7 Thermal denaturation of a DNA sample; a plot of absorbance at 260 nm against the temperature allows determination of the melting temperature (T_m).

vii. *Microbial growth kinetics*: A UV/visible spectrophotometer is routinely used to monitor the growth of microorganisms. *The underlying principle behind this, however, is not absorbance but scattering*. As the number of microbial cells increase in a culture, they cause more scattering in light. The detector therefore receives less amount of radiation, recording this as absorbance. To distinguish this from actual absorbance, the observed value is referred to as the optical density.

QUIZ

Q1: The molar absorption coefficient of tyrosine in water is 1280 $M^{-1}cm^{-1}$ at 280 nm. Calculate the concentration of a tyrosine solution in water if the absorbance of the solution is 0.34 in a 1 cm path length cell.

Ans: Given:

λ = 280 nm $\quad \varepsilon_{280\ nm}$ = 1280 $M^{-1}cm^{-1}$ $\quad l$ = 1 cm $\quad A$ = 0.34

From Beer-Lambert Law:

$$A = \varepsilon c l$$

$$c = \frac{A}{\varepsilon l} = \frac{0.32}{1280 M^{-1}cm^{-1} \times 1\ cm} = 0.00025\ M = 250\ \mu M$$

Q2: Calculate the concentration of a tryptophan solution that gives an absorbance of 0.25 at 280 nm in a 1 mm path length cell (Given $\varepsilon_{280\ nm}$ = 5690 $M^{-1}cm^{-1}$).

Ans: The concentration of the given sample can be estimated using Beer-Lambert law:

$$A = \varepsilon c l$$

$$c = \frac{A}{\varepsilon l}$$

$$c = \frac{0.25}{5690\ M^{-1}cm^{-1} \times 1\ mm}$$

$$c = \frac{0.25}{5690\ M^{-1}cm^{-1} \times 0.1\ mm}$$

$$c = 4.39 \times 10^{-5}\ M = 43.9\ \mu M$$

Q3: Concentration of a pure compound in solution can easily be determined by taking absorbance at any wavelength in a given spectral region if ε at these wavelengths is known. Why then absorbance is generally recorded at λ_{max}?

Ans: This is done for the following reasons:

a) At λ_{max}, the ε value is maximum, therefore reliable absorbance *i.e.* A between 0.05 – 0.5 can be obtained at lower concentrations of the compound.

b) At λ_{max}, the slope of the absorption spectrum, $\frac{dA}{d\lambda}$ or $\frac{d\varepsilon}{d\lambda}$, is zero. This ensures that for a given bandwidth of the incident radiation, the ε is relatively constant in this region as compared to the regions of non-zero slopes. If ε is not constant, the linearity of the Beer Lambert law is compromised.

Lecture 6 Fluorescence Spectroscopy-I

This lecture is a very concise review of the phenomenon of fluorescence and the associated processes. Let us move a step forward from the absorption of the UV/visible radiation. What happens to the electrons that absorb UV/visible light and occupy the high energy molecular orbitals? In a UV/visible absorption experiment, the samples continue absorbing light. This means that the higher energy molecular orbitals never get saturated. This further implies that after excitation, the molecules somehow get rid of the excess energy and return back to the ground state. The electrons can return back to the ground state in different ways such as releasing the excess energy through collisions or through emitting a photon. In fluorescence, the molecules return back to the ground state by emitting a photon. The molecules that show fluorescence are usually referred to as the *fluorophores*. Various electronic and molecular processes that occur following excitation are usually represented on a Jablonski diagram as shown in Figure 6.1.

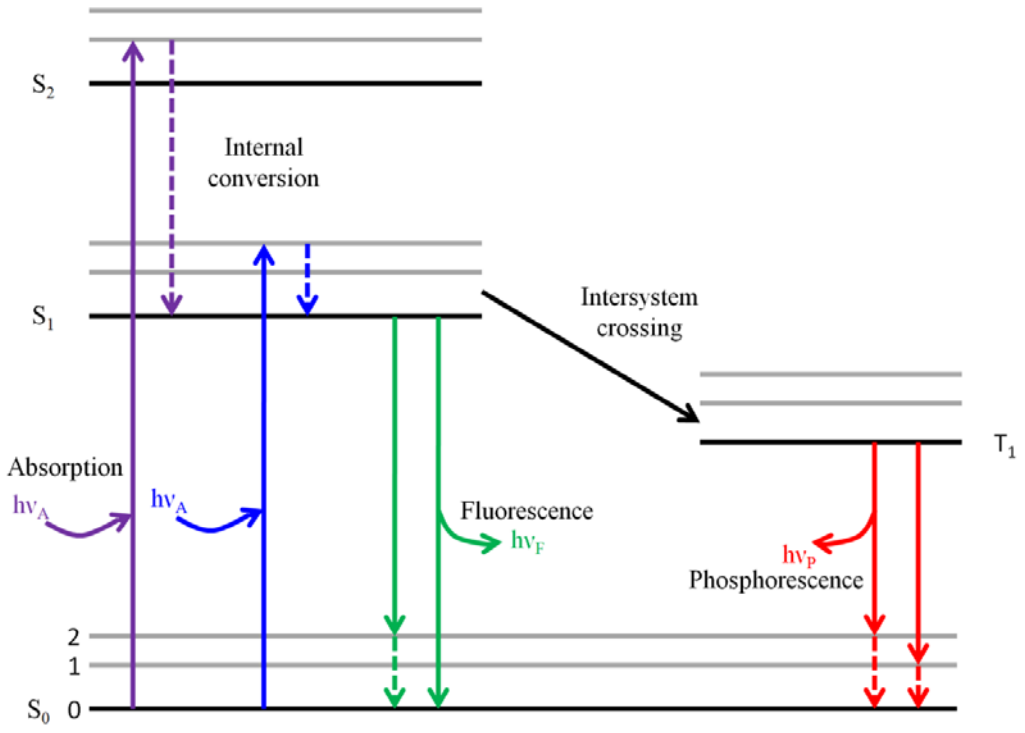


Figure 6.1 Jablonski diagram showing various processes following absorption of light by the fluorophore

S_0, S_1, and S_2 represent the singlet electronic states while the numbers 0, 1, 2 represent the vibrational energy levels associated with the electronic states. T_1 depicts the first triplet electronic state. Let us go through the processes shown in Figure 6.1:

Absorbance: S_0 state with 0^{th} vibrational level is the state of lowest energy and therefore, the highest populated state. Absorption of a photon of resonant frequency usually results in the population of S_1 or S_2 electronic states; but usually a higher vibrational state. Transition of electrons from low energy molecular orbital to a high energy molecular orbital through absorption of light is a femtosecond (10^{-15} s) phenomenon. The electronic transition, therefore, is too quick to allow any significant displacement of the nuclei during transition.

Internal conversion: Apart from few exceptions, the excited fluorophores rapidly relax to the lowest vibrational state of S_1 through non-radiative processes. Non-radiative electronic transition from higher energy singlet states to S_1 is termed as *internal conversion* while relaxation of a fluorophore from a higher vibrational level of S_1 to the lowest vibration state is termed as *vibrational relaxation*. The terms 'internal conversion' and 'vibrational relaxation', however, are often interchangeably used. The timescale of internal conversion/vibrational relaxation is of the order of 10^{-12} seconds.

Fluorescence: Fluorescence lifetimes are of the order of 10^{-8} seconds, implying that the internal conversion is mostly complete before fluorescence is observed. Therefore, fluorescence emission is the outcome of fluorophore returning back to the S_0 state through $S_1 \rightarrow S_0$ transition emitting a photon. This also explains why emission spectra are usually independent of the excitation wavelength, also known as Kasha's rule (However, there are exceptions wherein fluorescence is observed from $S_2 \rightarrow S_1$ transition). The $S_1 \rightarrow S_0$ transition, like $S_0 \rightarrow S_1$ transition, typically results in the population of higher energy vibrational states. The molecules then return back to the lowest vibrational state through vibrational relaxation.

Intersystem crossing: Intersystem crossing referes to an isoenergetic non-radiative transition between electronic states of different multiplicities. It is possible that a molecule in a vibrational state of S_1 can move to the isoenergetic vibrational state of T_1. The molecule then relaxes back to the lowest vibrational state of the triplet state.

Phosphorescence: The molecule in the triplet state, T_1, can return back to the S_0 state emitting a photon. This process is known as phosphorescence and has time scales of several orders of magnitudes higher than that of fluorescence (10^{-3} – 10 s).

Characteristics of fluorescence:

Figure 6.2 shows absorption and fluorescence emission spectrum of a hypothetical fluorophore. The important characteristics of the fluorescence emission can be briefly summarized as follows:

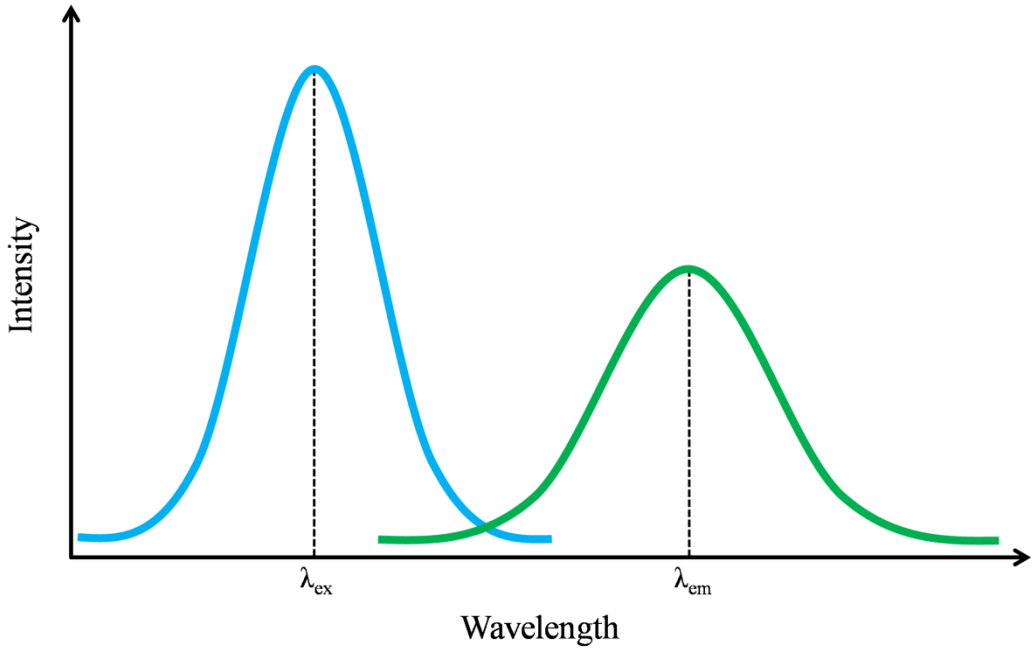


Figure 6.2 Absorption and fluorescence emission spectrum of a hypothetical fluorophore

Stokes shift: A fluorescence emission spectrum is always shifted towards longer wavelengths with respect to the absorption spectrum. This shift is known as *Stokes shift* and is expected as excited molecules lose energy through processes like internal conversion and vibrational relaxation. The emitted radiation is therefore expected to be of lower energy *i.e.* higher wavelength.

Kasha's rule: As fluorescence emission is observed from $S_1 \rightarrow S_0$ transtions (except a few exceptions), fluorescence absorption spectrum is independent of the excitation wavelength.

Franck-Condon principle: The Franck-Condon principle states that the positions of the nuclei do not change during electronic transitions. The transitions are said to be vertical. This implies that if the probability of $0^{th} \rightarrow 2^{nd}$ vibrational transition during $S_0 \rightarrow S_1$ transition is highest, the $2^{nd} \rightarrow 0^{th}$ transition will be most probable in the reciprocal transition (Figure 6.3).

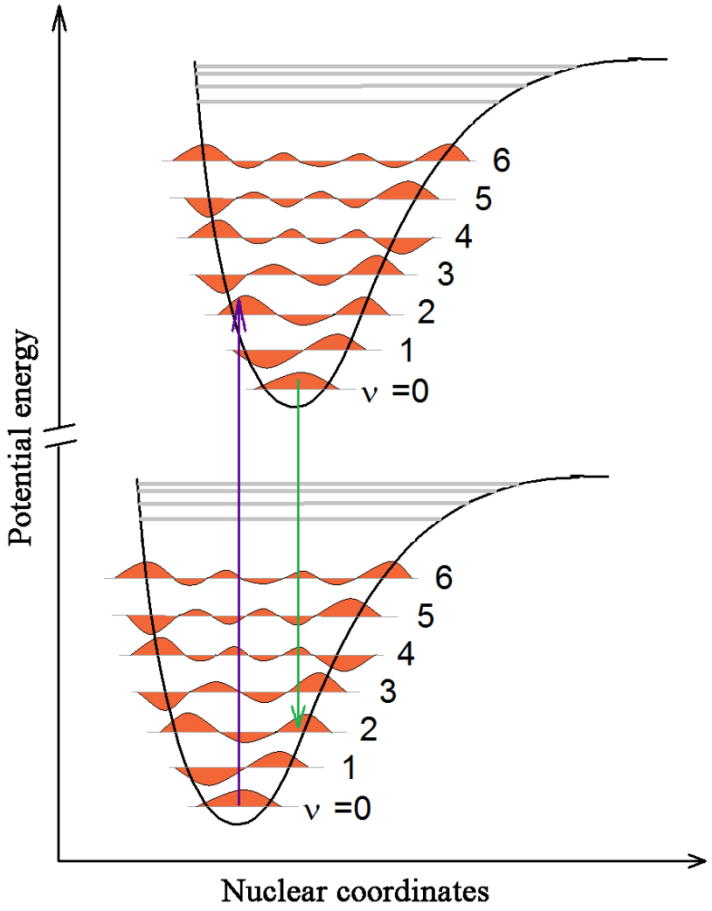


Figure 6.3 Potential energy diagrams showing the Franck-Condon principle

This results in an emission spectrum that is a mirror image of the $S_0 \rightarrow S_1$ transition in terms of the shape. There are several exceptions to the mirror image rule that arise largely due to the excited state reactions of the molecule.

Quantum yield: As has been mentioned earlier, an excited molecule can come back to the ground state through non-radiative pathways.

Fluorescence quantum,

$$Q = \frac{number\ of\ photons\ emitted}{number\ of\ photons\ absorbed} = \frac{\Gamma}{\Gamma + k_{nr}}$$

where,

Γ is the rate of radiative process *i.e.* fluorescence

k_{nr} is the rate of all the non-radiative processes bringing molecule to the S_0 state

Fluorescence lifetime: Lifetime of a fluorophore is defined as the average time it spends in the excited state before returning to the S_0 state. It is therefore the reciprocal of the rate of processes de-exciting the molecule.

$$\text{Fluorescence lifetime, } \tau = \frac{1}{\Gamma + k_{nr}}$$

Fluorescence quenching, resonance energy transfer and anisotropy

Fluorescence spectroscopy comprises of experiments exploiting various different phenomena related to it. Discussion of all these experiments is beyond the scope of this course, but we shall have a quick look at a few important phenomena related to fluorescence.

Fluorescence quenching: A decrease in fluorescence intensity is referred to as quenching. A molecule that quenches the fluorescence of a fluorophore is called a quencher. A quencher can be either a collisional quencher or a static quencher. A collisional quencher brings about decrease in fluorescence intensity by de-exciting the excited fluorophore through collisions. Addition of another non-radiative process to the system leads to lower quantum yield. A static quencher forms a non-fluorescent complex with the fluorophore. It effectively leads to a decrease in the concentration of the fluorophore thereby decreasing the fluorescence emission intensity.

Resonance energy transfer: Resonance energy transfer (RET), also known as fluorescence resonance energy transfer (FRET) is an excited state phenomeneon wherein energy is transferred from a donor molecule (*D*) to an acceptor molecule (*A*). The prerequisite for the energy transfer is that there should be an overlap between the emission spectrum of the *D* and the absorption spectrum of the *A* (Figure 6.4).

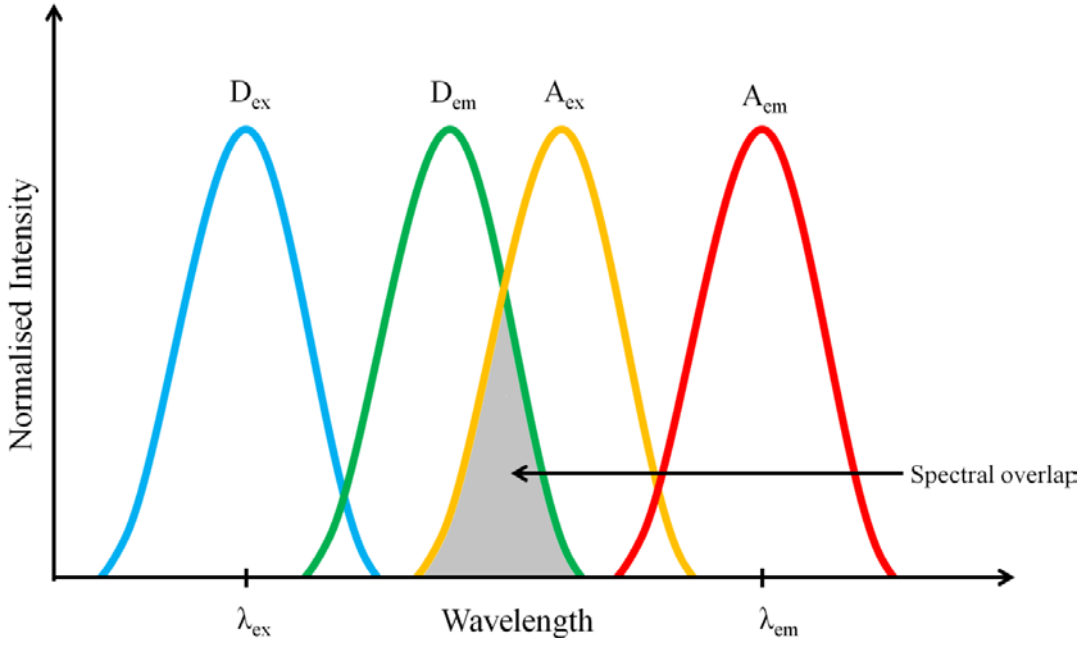


Figure 6.4 Diagrammatic representation of spectral overlap between donor's emission and acceptor's absorption spectrum.

The efficiency of energy transfer depends upon

i. the distance between *D* and *A*
ii. the relative orientation of the transition dipoles of *D* and *A*
iii. the extent of the overlap between *D*'s emission spectrum and *A*'s absorption spectrum

$$\text{Efficiency of energy transfer } E = \frac{R_0^6}{R_0^6 + r^6}$$

where,

r is the distance between *D* and *A*.

R_0 (also called the Förster distance) is the distance (*r*) between *D* and *A* at which the efficiency of energy transfer is 50%, and is characteristic of a *D*-*A* FRET pair.

Resonance energy transfer can be used to determine the distances between *D* and *A*, and is therefore also termed as molecular ruler.

Fluorescence anisotropy: The radiation emitted by a sample following excitation with polarized light can be polarized. Polarization is measured in terms of anisotropy. Zero anisotropy implies isotropic/non-polarized radiation while non-zero anisotropy implies some degree of polarization. Figure 6.5 shows how fluorescence anisotropic measurements are made.

Transition dipole moment: The transition dipole moment represents the transient dipole moment generated from the charge displacement during a transition. The transition dipole moments are defined vector quantities for the transitions of a particular molecule.

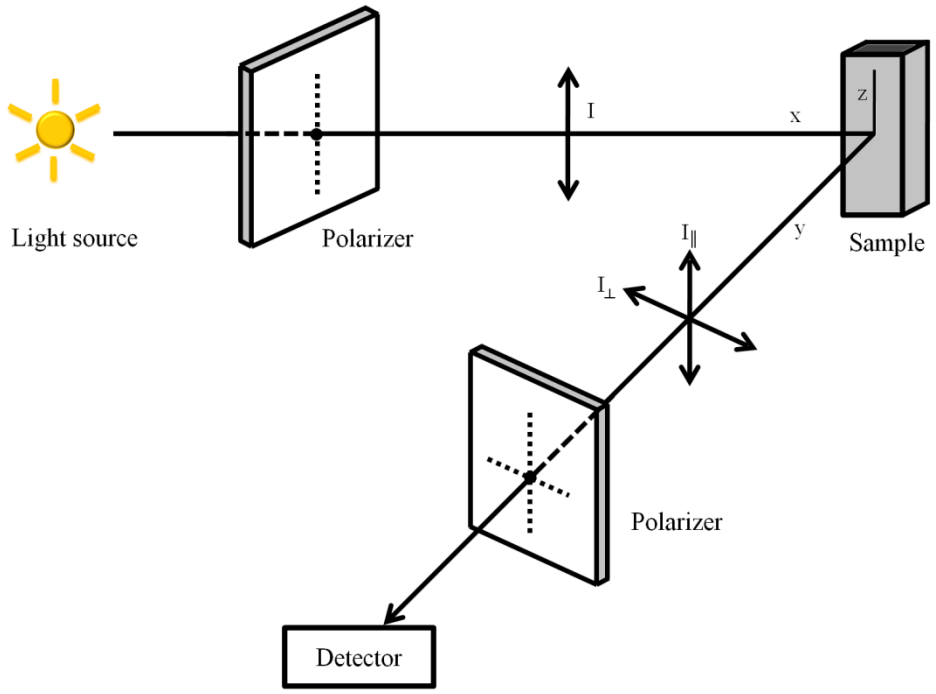


Figure 6.5 A schematic diagram showing the measurement of fluorescence anisotropy

The sample is excited with the linearly polarized light and emission is recorded at 90°. A polarizer is placed before the detector that allows intensity measurement of the light polarized parallel ($I_{\parallel}$) and perpendicular ($I_{\perp}$) to the direction of excitation radiation. The anisotropy (r) is given by

$$r = \frac{I_{\perp} - I_{\parallel}}{I_{\perp} + 2I_{\parallel}}$$

Molecular tumbling before emission changes the orientation of the transition dipole moment, resulting in the loss of polarization (Figure 6.6). As rotational diffusion of the molecules depends on their sizes, fluorescence anisotropy can be used to measure the diffusion coefficient and therefore the sizes of the molecules.

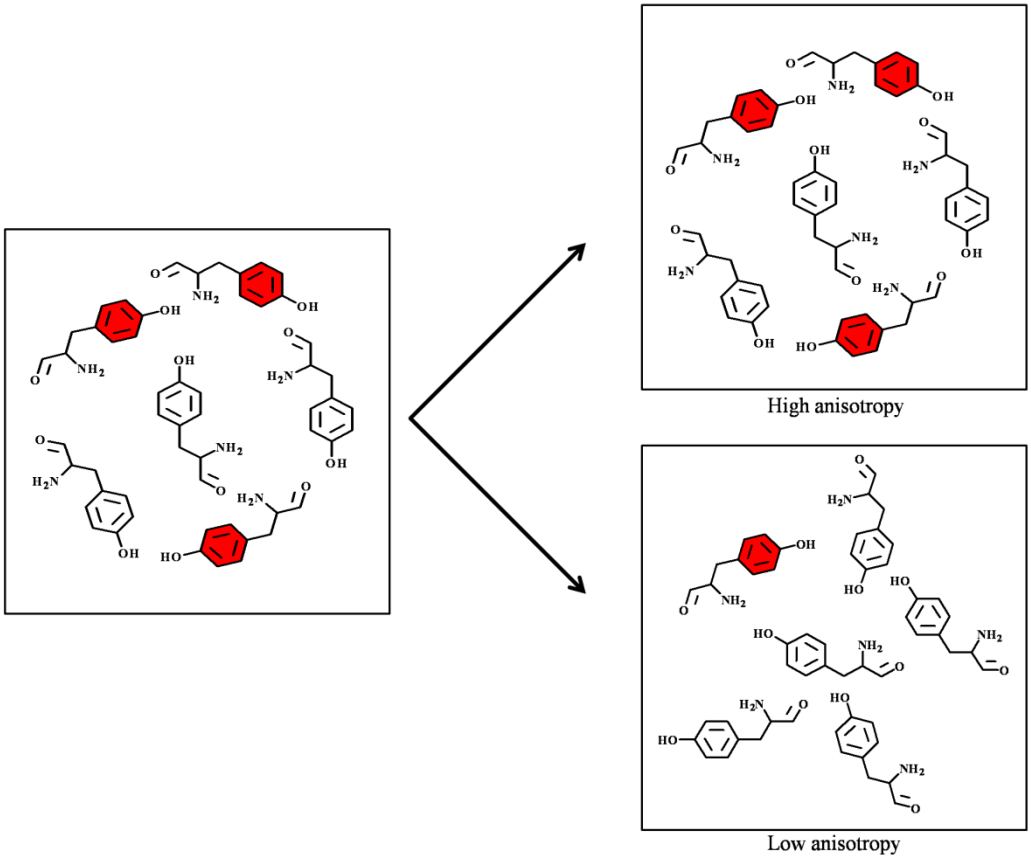


Figure 6.6 Depolarization of radiation as a result of molecular tumbling

We shall, in the next lecture, discuss the biological fluorophores and the applications of fluorescence in understanding the biomolecules.

Lecture 7 Fluorescence Spectroscopy-II

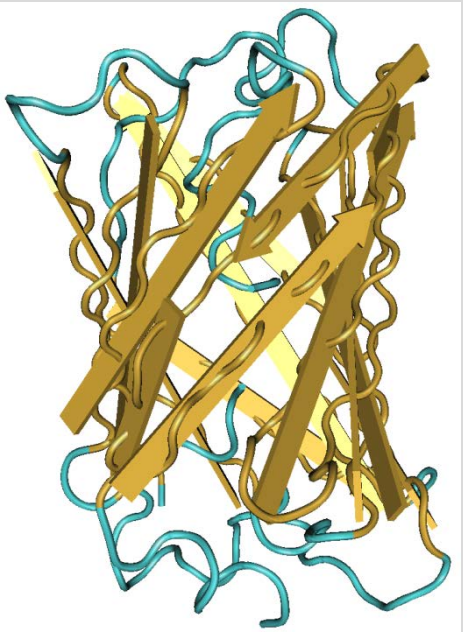
Biological fluorophores

Amino acids: Aromatic amino acids tryptophan (Trp), tyrosine (Tyr), and phenylalanine (Phe) are perhaps the most important intrinsic biological fluorophores. Proteins harboring these amino acids become intrinsically fluorescent.

Proteins: Proteins are fluorescent due to the presence of aromatic amino acids that fluoresce in the near UV region. Certain proteins, however, do fluoresce in the visible region. Green fluorescent protein (GFP), for example, fluoresces in the green region of the electromagnetic spectrum. The discovery of green fluorescent protein has revolutionized the area of cell biology research. It is therefore important to see what green fluorescent protein is and why it fluoresces in the visible region (See Box 1).

Box 7.1: Green Fluorescent Protein (GFP)

Green fluorescent protein, abbreviated as GFP was discovered by Shimomura and coworkers in 1962. The protein was isolated from the jellyfish, *Aequorea victoria*, that glows in the dark. GFP is a 238 amino acid long protein that folds into an 11-stranded β-barrel structure wherein an α-helix passes through the barrel.



PDB ID: 1EMA

The fluorophore of the GFP, p-hydroxybenylideneimidazolinone is formed by the residues 65-67 (Ser-Tyr-Gly) and is present in the α-helix passing through the barrel.

The excitation spectrum of GFP exhibits a strong absorption band at 395 nm and a weak band at 475 nm. Emission is observed at ~504 nm *i.e.* in the green region. GFP is an excellent fluorophore with a molar absorption coefficient of ~30000 $M^{-1}cm^{-1}$ at 395 nm and fluorescence quantum yield of 0.79. GFP has been engineered through extensive mutations to remove the undesirable properties that could affect its use as a potential fluorophore. For example, a Ser65 → Thr65 mutant has improved quantum yield and its major excitation band shifted to 490 nm. GFP has the tendency to form oligomers, seriously questioning its use as a fluorescent probe. The aggregation tendency has also been removed through extensive mutations. GFP can easily be tagged to a protein by expressing the fusing gene (GFP gene fused with the gene expressing the desired protein). The GFP then acts as a reporter for all the processes the linked protein is involved in. Several color variants of GFP have been generated through modifications in the residues that constitute the fluorophore. Development of the GFP variants with varying excitation and emission characteristics has made it possible to label the proteins differentially. This is a huge breakthrough and allows easy monitoring of the biological processes using fluorescence microscopy as discussed in lectures 15 and 20.

Nucleotides: Nicotinamide adenine dinucleotide in its reduced form, NADH and the flavin adenine dinucleotide in its oxidized form, FAD are fluorescent in the visible region of the electromagnetic spectrum. It is not necessary for all the biomolecules to have an intrinsic fluorophore to perform fluorescence experiments. Fluorescent groups can be covalently incorporated into the molecules making them fluorescent with desirable fluorophore. Such externally incorporated fluorophores are called extrinsic fluorophores.

Applications of fluorescence

Protein folding: High sensitivity of tryptophan fluorescence to the polarity of solvent makes it an interesting intrinsic fluorescent probe for studying protein folding. In the proteins having Phe and Tyr, Trp can be selectively excited at 295 nm. In water and other aqueous solutions, tryptophan fluoresces with an emission maximum, λ_{max} around 350 nm. A tryptophan present in the hydrophobic environment usually displays a blue shift in the emission spectrum and an increase in quantum yield. Due to the hydrophobic nature of the indole side chain, tryptophans are usually buried inside the core of the proteins. The folding can therefore be studied by monitoring the Trp fluorescence as protein folds burying the water-exposed Trp residues inside the protein.

Peptide-lipid interactions: Interaction of the peptides having Trp residues with lipid bilayers can easily be studied using fluorescence spectroscopy. Interaction of the peptide with lipids brings the tryptophan in relatively hydrophobic environment causing a blue shift in emission spectrum (Figure 7.1).

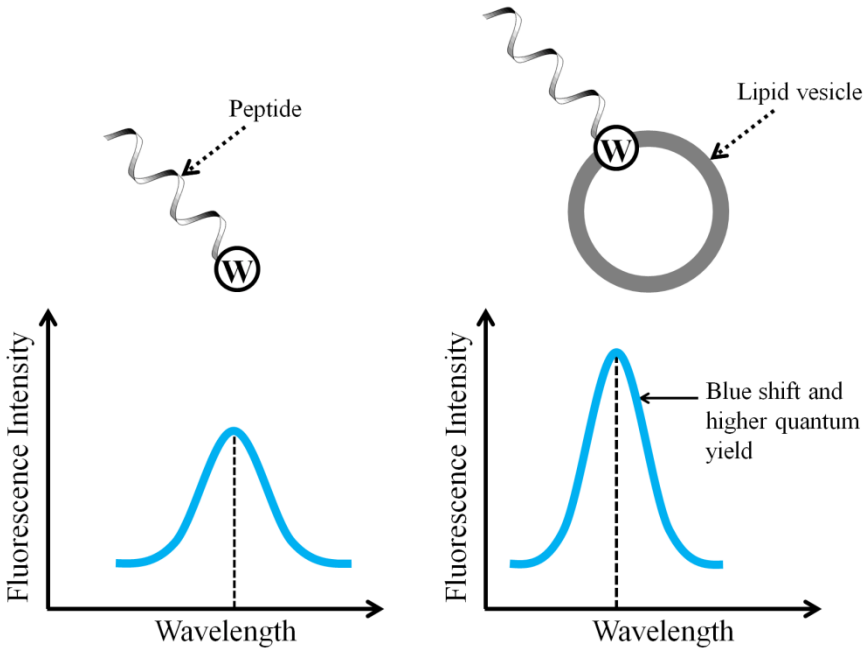


Figure 7.1 Spectral changes in tryptophan fluorescence upon binding to lipid bilayers

Binding studies: Binding of small fluorescent molecules to the biomacromolecules can be studied using fluorescence anisotropy. Binding of the fluorophore to a macromolecule will reduce its tumbling (increase its rotational correlation time) thereby resulting in higher fluorescence anisotropy.

FRET:

i. The distance between two sites in a biomacromolecule such as a protein can be calculated by labeling these sites with suitable donor-acceptor FRET pair. FRET can also be used to study the intermolecular interaction if the interacting molecules comprise of the fluorophores making a FRET pair.

ii. Interactions of peptides and other molecules with lipid bilayers comprising fluorophore labeled lipids. If the interacting molecule makes a FRET pair with the fluorescent lipid, the distance between them can be calculated providing information about the insertion of the molecule in the lipid bilayer.

iii. FRET has been utilized to study the kinetics of enzymatic reactions. For example, a DNA molecule, tagged with the fluorescence donor at one end and an acceptor at the other end can be used as a substrate to study the restriction endonuclease activity and cleavage reaction kinetics (Figure 7.2). A similar assay can be used to study the proteases using peptides as the substrates.

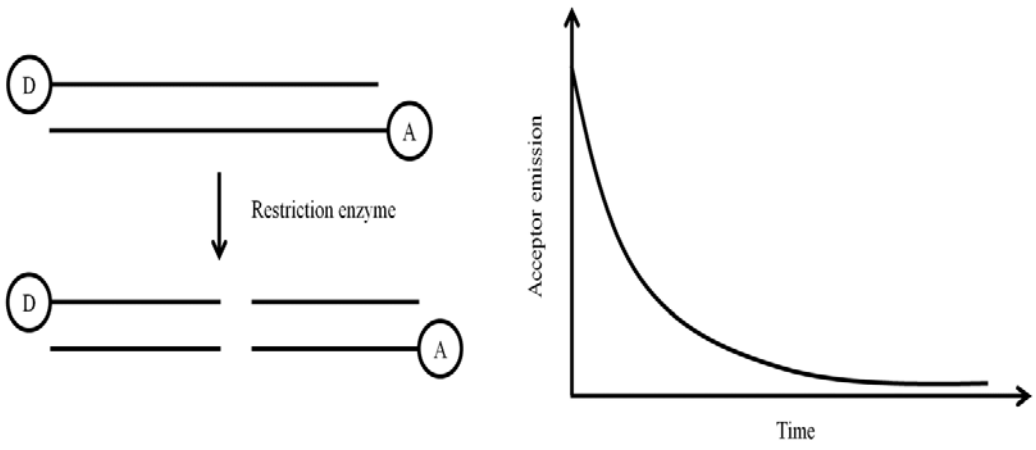


Figure 7.2 Decrease in fluorescence intensity of the acceptor following cleavage of DNA molecules.

Fluorescence quenching:

i. Interaction of a fluorophore with another molecule(s) may provide it protection against a collisional quencher. For example, interaction of a Trp containing peptide with lipid bilayers can be studied using iodide (I^-) as the collisional quencher. The peptide sample in the presence of lipid vesicle is titrated with the potassium iodide (KI) and fluorescence spectra recorded at each quencher concentration. The collisional fluorescence quenching is described by a plot of 'the ratio of quantum yield in the absence of quencher to that in the presence of quencher' against 'the quencher concentration'. Such a plot is known as the Stern-Volmer plot. The Stern-Volmer equation is given by

$$\frac{F_0}{F} = 1 + k_q\tau_0[Q] = 1 + K_{sv}[Q] \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (7.1)$$

where,

F_0 = Fluorescence intensity in the absence of quencher

F = Fluorescence intensity in the presence of quencher

k_q = Bimolecular quenching constant

τ_0 = Fluorescence lifetime in the absence of quencher

$[Q]$ = Quencher concentration

K_{sv} = Stern-Volmer constant

A normalized accessibility factor (NAF) is defined as the ratio of 'the K_{sv} in the presence of the binding partner of the fluorophore' to 'that without the binding partner'.

ii. The fluorescence intensity of a sample increases with an increase in the fluorophore concentration. Beyond certain concentration, however, the fluorescence intensity decreases due to self collisional quenching. This property is often used to study the membranolytic activities of a compound. A fluorescent dye at self-quenching concentrations is trapped inside a lipid vesicle. A membranolytic compound results in the release of the fluorescent dye causing increase in fluorescence emission intensity (Figure 7.3).

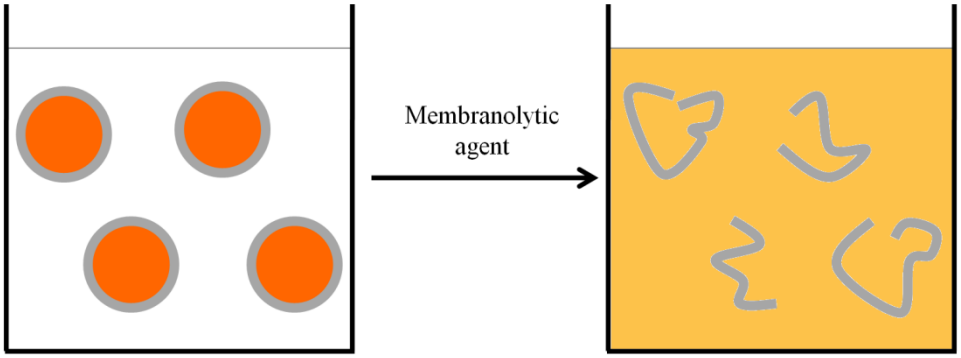


Figure 7.3 Membranolytic activity of a compound monitored through dye release assay. Release of dye from the lipid vesicle diminishes the self-quenching resulting in enhanced fluorescence emission.

iii. Fusion of lipid vesicles can also be studied using the same approach. Vesicles that contain self-quenching concentrations of the fluorescent dye are titrated with the vesicles without fluorophores. A fusion will result in the dilution of fluorophores; the consequent decrease in self-quenching is exhibited as an increase in the fluorescence intensity (Figure 7.4).

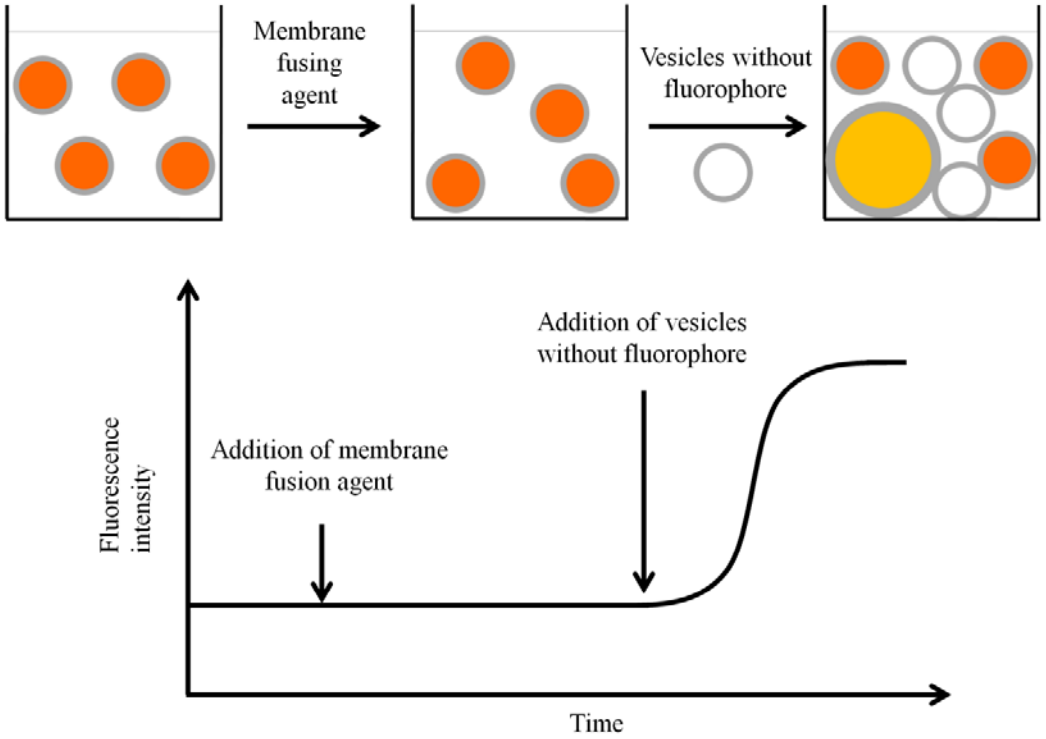
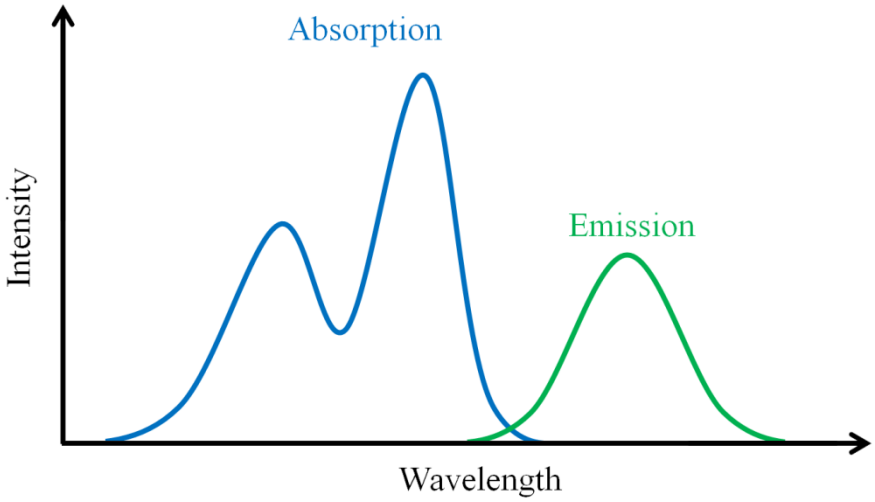


Figure 7.4 Fusion of fluorescent dye-containing lipid vesicles with vesicles without dye results in dilution of dye. The dilution results in lesser self-quenching thereby increasing the fluorescence intensity.

QUIZ

Q1: Shown below are the absorption and emission spectra of a fluorophore. The fluorescence emission for this fluorophore is not the mirror image of the absorption spectrum. How do you explain this?



Ans: The low wavelength absorption band is likely to be arising from $S_0 \rightarrow S_2$ transition. As fluorophore relaxes back to S1 state prior to emission, the fluorescence band is the mirror image of the band arising from $S_0 \rightarrow S_1$ transition, not the entire absorption spectrum.

Q2: If the efficiency of energy transfer between a donor and acceptor is 80%. Calculate the distance between them if the Förster distance between them is 40 nm?

Ans: The efficiency of energy transfer, *E* is given by:

$$E = \frac{R_0^6}{R_0^6 + r^6}$$

Given: E = 80% = 0.8, R_0 = 40 nm

Rearranging the expression for the efficiency of energy transfer

$$E = \frac{1}{1+\left(\frac{r}{R_0}\right)^6}$$

$$0.8 = \frac{1}{1+\left(\frac{r}{40}\right)^6}$$

$$1 + \left(\frac{r}{40}\right)^6 = \frac{1}{0.8} = 1.25$$

$$\left(\frac{r}{40}\right)^6 = 1.25 - 1 = 0.25$$

$$\frac{r}{40} = (0.25)^{\frac{1}{6}} = 0.7937$$

$$r = 0.7937 \times 40 = 31.748 \text{ nm} \approx 31.75 \text{ nm}$$

Lecture 8 Circular Dichroism Spectroscopy-I

Introduction

Before going ahead to see what circular dichroism (abbreviated as CD) means, let us have a quick revisit on the polarized light. Light, as we have discussed in lecture 3 is electromagnetic radiation where electric field and the magnetic field are always perpendicular to each other. From now on, we shall mention only electric field; it is implicit that at all points in time and space, the magnetic field vector is perpendicular to both the electric field vector and the direction of the propagation of light. Unpolarized light is comprised of several electromagnetic waves with their electric field vectors (and therefore magnetic field vectors also) pointing in all possible directions, but perpendicular to the direction of light propagation. If the vectors in all, but one, directions are cut off, the resulting radiation is a plane polarized light as the electric field vector is confined to one plane (Figure 8.1). Looking towards the light source will exhibit electric field fluctuations in one line; the plane polarized light is therefore also referred to as the linearly polarized light.

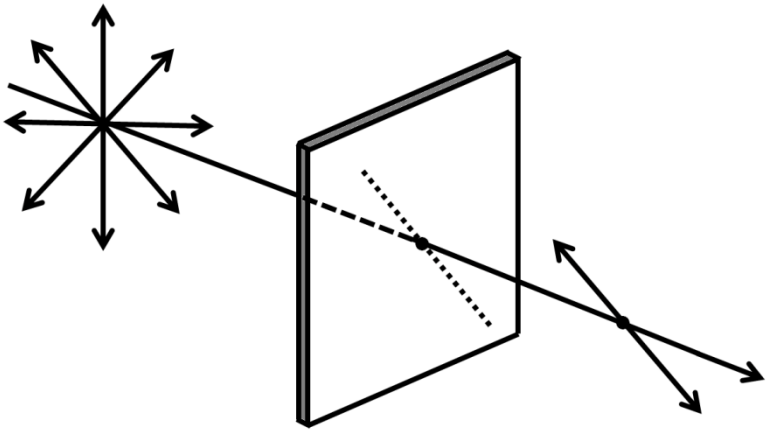


Figure 8.1 Plane polarized light produced by a linear polarizer

Superposition of polarized waves

Two electromagnetic waves can be superposed through vector addition of their electric field vectors. The properties of the resultant waves depend on the wavelength, polarization, and the phase of the superposing waves. In-phase superposition of two waves of same wavelength that are linearly polarized in two perpendicular planes results in a linearly polarized light with its electric field vector oscillating in a plane that is inclined at an angle of 45° to the polarization planes of both the waves (Figure 8.2).

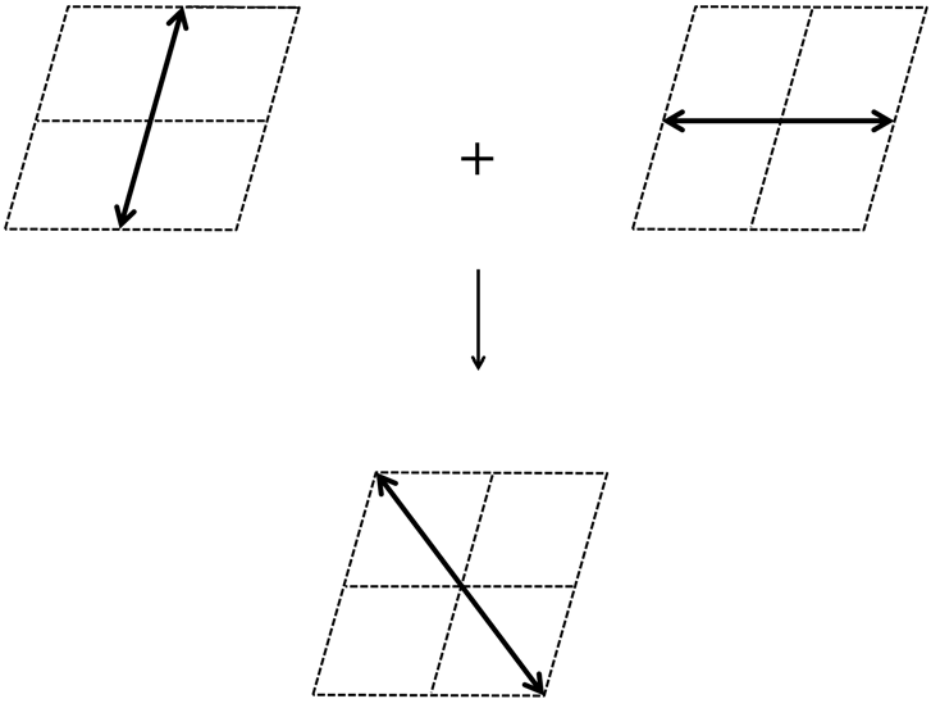


Figure 8.2 Superposition of linearly polarized waves

Let us see what happens when the two plane polarized waves, polarized in two perpendicular planes meet each other out of phase. Suppose the two waves have a phase difference of 90°. As the two waves have same wavelength, a 90° phase difference implies that when one of the wave is at maximum amplitude, the amplitude of the other one is minimum and vice versa. If the amplitudes of the two waves are equal, their superposition with a 90° phase difference results in a wave wherein electric field vector traverses a circular path (Figure 8.3). The electric field of the resultant wave is never zero but a vector of constant length. When looked at the travelling wave from the direction of propagation, the electric field appears to be

rotating in a circle. The resulting light is therefore termed as circularly polarized light (Figure 8.3).

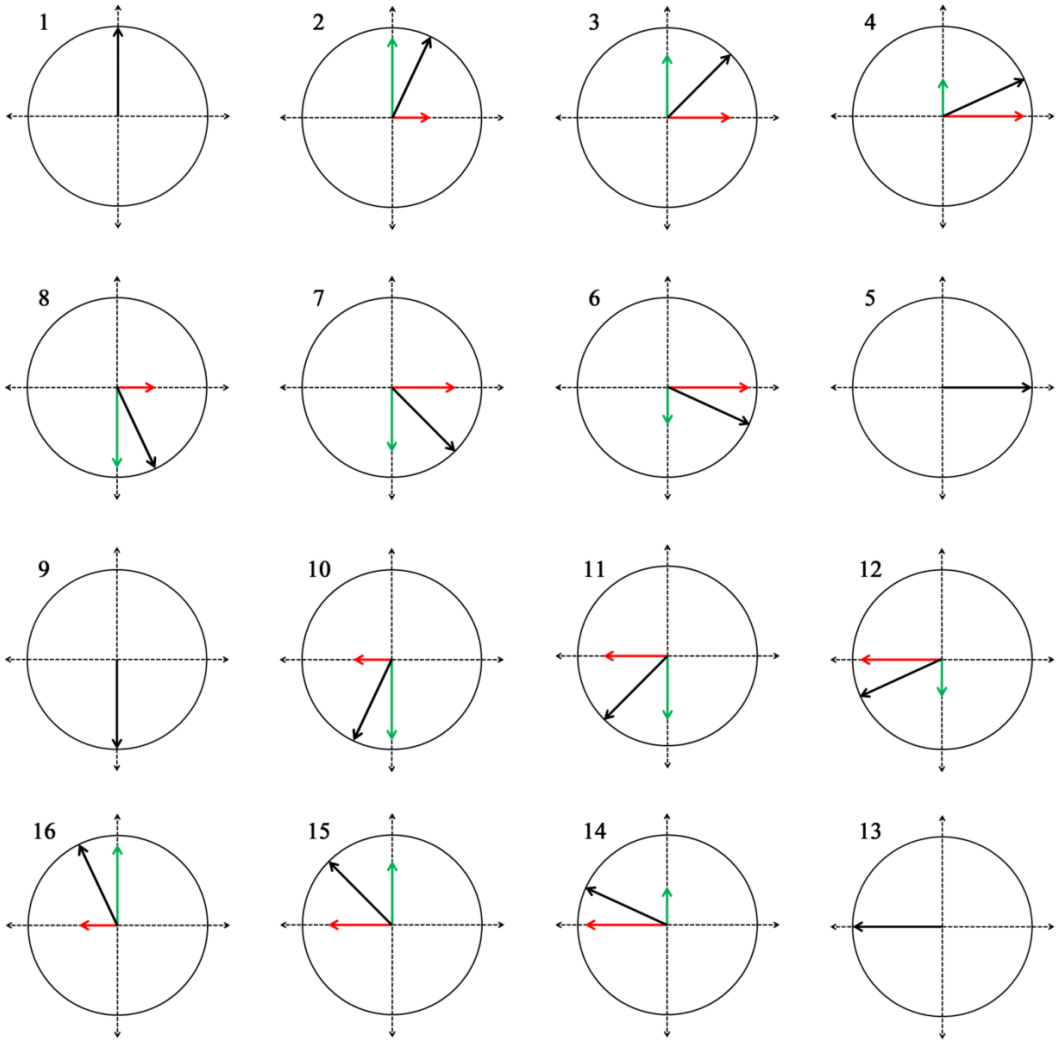


Figure 8.3 Superposition of waves linearly polarized in mutually perpendicular plain and that meet together 90° out of phase.

The direction of rotation depends on phase difference; a -90° phase difference would result in a circularly polarized light where the electric field rotates in opposite direction. When looked towards the light source, the electric field vector of a right circularly polarized wave appears to rotate counterclockwise in space while that of a left circularly polarized wave rotates clockwise. What happens when the right circularly polarized light (RCPL) and the left circularly polarized light (LCPL) superpose? The resultant wave is a linearly polarized wave (Figure 8.4). A linearly polarized light can therefore be considered as being composed of a right circularly polarized light and a left circularly polarized light.

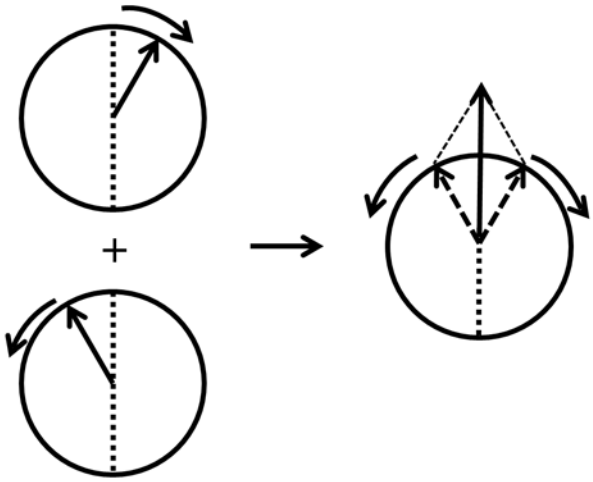


Figure 8.4 Superposition of left and right circularly polarized light resulting in plane polarized light.

Circular Dichroism

Circular dichroism, abbreviated as CD, is a chiroptical spectroscopic method. A chiral molecule or an achiral molecule in asymmetric environment interacts differently with the LCPL and the RCPL. The literal meaning of dichroism is 'two colors'. In chiroptical spectroscopy, dichroism means differential absorption of the lights with different polarizations. Circular dichroism, therefore, refers to the differential absorption of the left and right circularly polarized light and is defined as:

$$CD = \Delta A = A_l - A_r \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots(6.1)$$

where, A_l and A_r are the absorbances for the left and right circularly polarized lights, respectively.

We can therefore say that the molar absorption coefficients for the two lights are different and can write the equation 6.1 can be written as:

$$CD = (\varepsilon_l - \varepsilon_r)cl \quad \cdots\cdots(6.2)$$

$$CD = \Delta\varepsilon cl \quad \cdots\cdots(6.3)$$

The preferential absorption of LCPL over RCPL (or vice versa) results in elliptical polarized light (Figure 8.5).

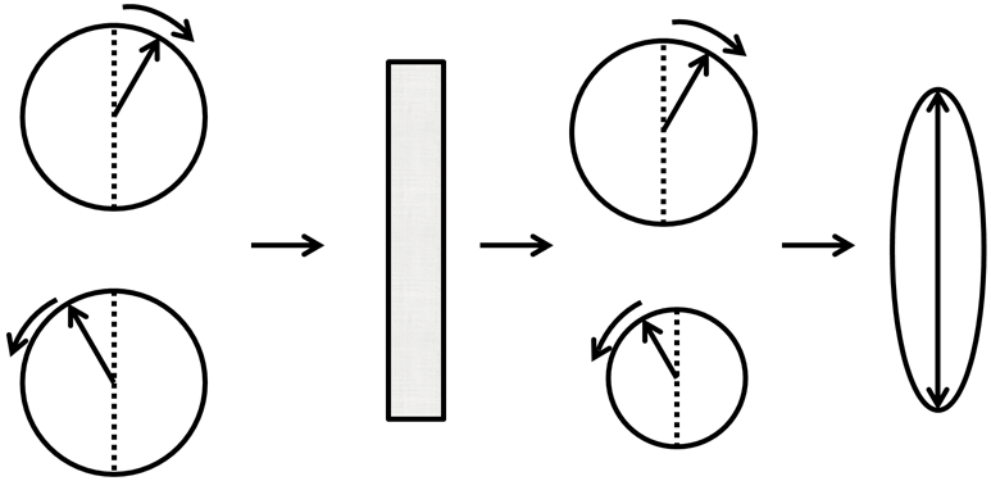


Figure 8.5 Differential absorption of the left and right circularly polarized light resulting in elliptically polarized light. Notice that if one component is completely absorbed, the resultant wave will be circularly polarized.

CD is historically represented in terms of ellipticity (θ) which is the tangent of ratio of minor to major axis of the ellipse. The relationship between CD and θ is give by:

$$\theta \text{ (radians)} = \frac{2.303}{4} \times CD \quad \cdots\cdots(6.4)$$

$$\theta \text{ (degrees)} = \frac{2.303}{4} \times CD \times \frac{180}{\pi} \quad \cdots\cdots(6.5)$$

$$\theta \text{ (degrees)} \approx 33.0 \times CD \quad \cdots\cdots(6.6)$$

A plot between ΔA or $\Delta\varepsilon$ or θ against the wavelength of light represents a CD spectrum. In this lecture, we shall be discussing only electronic CD. That means that we shall be looking at the electromagnetic region that causes electronic transition, which of course is UV/Visible region.

Circular birefringence

If a sample reduces the velocity of the LCPL and RCPL to different extents, the sample is said to be circularly birefringent and the phenomenon circular birefringence. Let us see what happens when the linearly polarized light (having two components, LCPL and RCPL) traverses a circular birefringent medium: the velocities of the two components are reduced to different extents *i.e.* they have different wavelengths in the

sample. After emerging from the samples, the wavelength is restored but two components can be out of phase. This results in the rotation of the polarization axis. If the material is not circularly dichroic, the plane of the linearly polarized light is rotated (Figure 8.6A). If the material is both circularly dichroic and birefringent, the plane polarized light will become elliptically polarized light with the major axis of the ellipse tilted with respect to the polarization axis of the incident polarized light (Figure 8.6B).

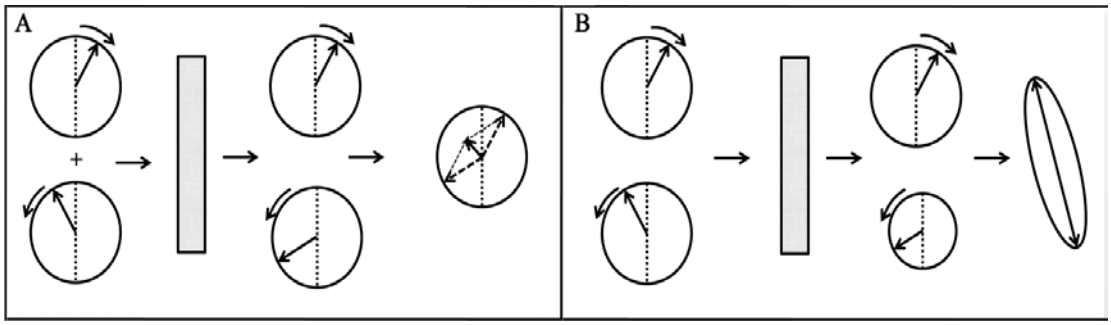


Figure 8.6 A linearly polarized light passing through a circular birefringent but not circular dichroic material (A) and through a material that is both circular birefringent and circular dichroic (B). Circular dichroism results in elliptically polarized light while circular birefringence causes change in the polarization axis.

Instrumentation

As CD is simply the difference in the absorbance of the LCPL and RCPL lights, a CD spectrometer, also known as a CD spectropolarimeter, is basically an absorption spectrophotometer (Figure 8.7). The instrument has a light source, usually a Xenon lamp. The polychromatic light from the source is converted to monochromatic radiation which is further converted to linearly polarized light by a polarizer. The linearly polarized light passes through a photoelastic modulator that alternately converts the linearly polarized light into LCPL and RCPL. The LCPL and the RCPL, therefore pass through the sample alternately and their absorbance gets recorded. Absorbance is recorded at various wavelengths to obtain a CD spectrum.

Photoelastic modulator: A photoelastic material is the one that exhibits birefringenece under mechanical stress. The photoelastic modulator in a CD instrument comprises of a quartz crystal fused to a piezoelectric material. Oscillations in the piezoelectric material drive the quartz crystal to oscillate at the same frequency. The crystal optical axis is at 45° to the linearly polarized light. The crystal retards one component of the light more than the other when compressed. When expanded the velocity of the two components gets reversed. A PEM, therefore gives alternating LCPL and RCPL.

Single wavelength CD values are also important in studying the fast reactions such as protein folding/unfolding (discussed in the next lecture).

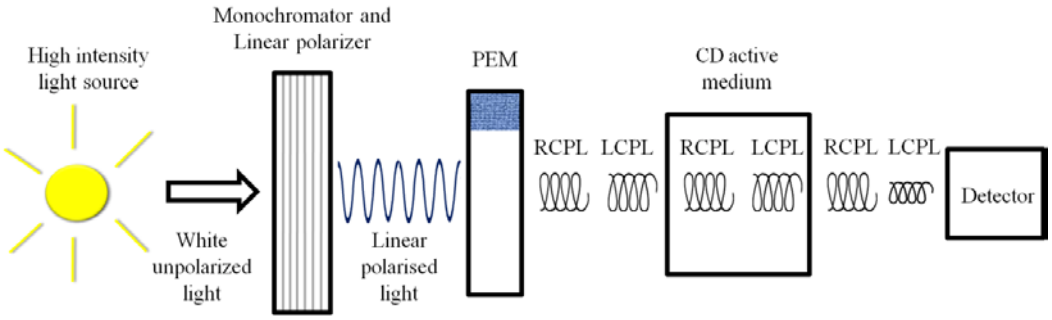


Figure 8.7 Schematic diagram of a CD spectropolarimeter.

Lecture 9 Circular Dichroism Spectroscopy-II

CD of biomolecules

Most biomolecules are chiral and the biomacromolecules are composed of chiral components. Folding of biomacromolecules into higher order structures further imparts them the asymmetry. CD has not been used as much to study other biomolecules probably, as it has been used to study proteins.

CD of proteins

Proteins are usually composed of 20 amino acids, 19 of which (except glycine) are chiral. This chirality also reflects in the higher order structures that the polypeptides adopt; α-helix, for example, is a right handed helix. If a polypeptide adopting α-helical structure is synthesized using D-amino acids, it folds into the left-handed α-helix under identical conditions. The other structural features of a polypeptide backbone include β-sheets, that are comprised of extended polypeptide chains; β-turns, that usually, but not essentially, link the β-strands in an antiparallel β-sheet; and unordered conformation. CD spectra of the proteins contain information about the asymmetric features of the polypeptide backbone. Furthermore, it can provide information about the orientation of the side chains. CD, therefore, is capable of providing information about the structure of proteins which in turn helps understanding their function. The chromophore that provides information about the conformation of the peptide backbone is the peptide bond (Figure 9.1); the spectra are therefore recorded in the far UV region, the region where peptide bond absorbs.

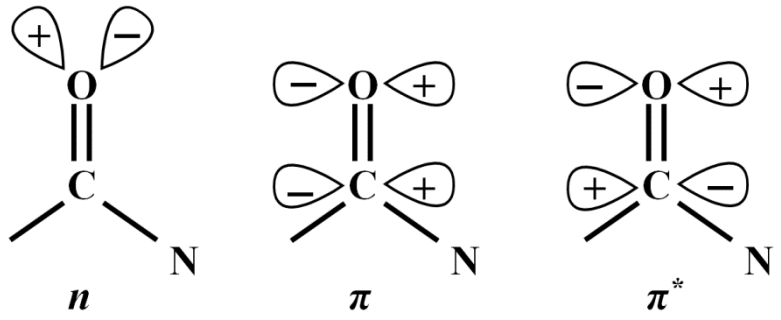


Figure 9.1 The peptide bond showing molecular orbitals involved in electronic transitions

Let us have a look at the CD spectra characteristic of the different structural components of the proteins (Figure 9.2).

- *α-helix:* The right handed α-helix displays two negative absorption bands centered around 222 nm ($n \rightarrow \pi^*$ transition) and 208 nm (a part of the $\pi \rightarrow \pi^*$ transition) and a strong positive band around 192 nm (a part of the $\pi \rightarrow \pi^*$ transition).
- *β-sheet:* β-sheets are characterized by the presence of a negative band centered around 216-218 nm ($n \rightarrow \pi^*$ transition) and a positive band of comparable intensity at around 195 nm ($\pi \rightarrow \pi^*$ transition).
- *β-turn:* A β-turn comprises of a four residue protein motif that causes the polypeptide backbone to take an approximately 180° turn. The CD spectrum for a β-turn is not well defined. A typical β-turn, however, shows a weak negative band around 225 nm ($n \rightarrow \pi^*$ transition), a strong positive band between 200 – 205 nm ($\pi \rightarrow \pi^*$ transition), and a strong negative band ($\pi \rightarrow \pi^*$ transition) between 180 – 190 nm.
- *Random coil:* Random coil or unordered conformation shows a weak positive band around 218 nm ($n \rightarrow \pi^*$ transition) and a strong negative band ($\pi \rightarrow \pi^*$ transition) below 200 nm.

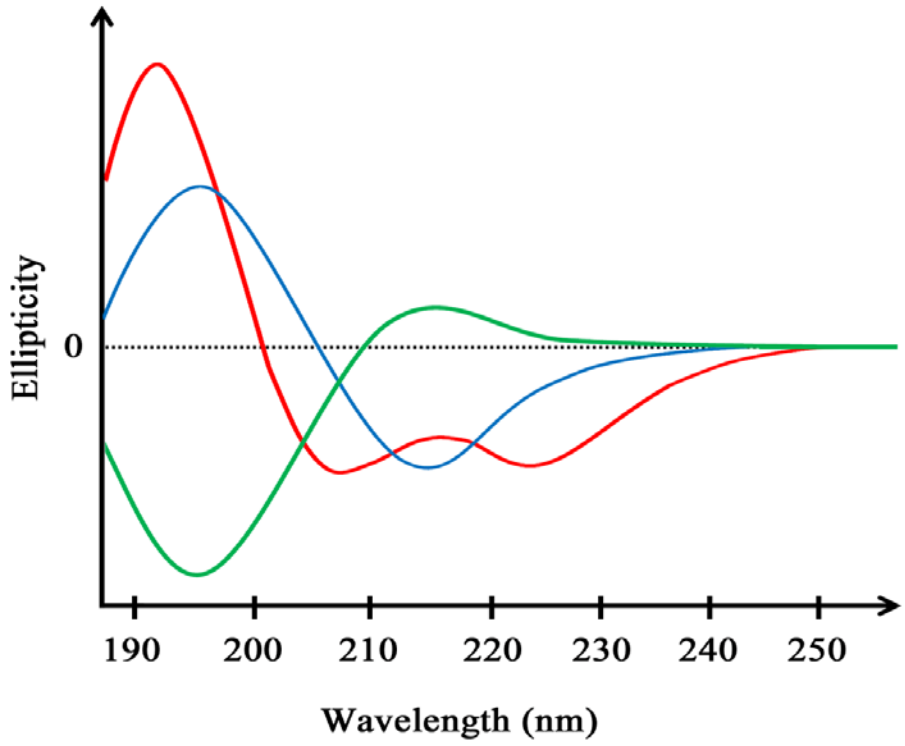


Figure 9.2 Far UV circular dichroism spectra of α-helix (red), β-sheet (blue), and unordered conformation (green)

The CD spectrum of a protein can be written as a linear combination of the spectra of all the structural components:

$$CD\ (protein) = a\ CD(\alpha\text{-helix}) + b\ CD(\beta\text{-sheet}) + c\ CD\ (Random\ coil)$$

As the CD spectra of different structural components are quite distinct, it is possible to estimate the fraction of different structural components in a protein from its CD spectrum. As discussed in lecture 5, proteins also have chromophores that absorb in the near UV region. These include the aromatic amino acids and disulfide linkages. The CD of aromatic amino acids is highly dependent on their environment and therefore near UV CD of proteins can provide the information about the environments these residues reside in as well as their orientations in the structure. As it provides information about the tertiary region, near UV CD is also referred to as tertiary CD in the context of the proteins.

CD of nucleic acids

As mentioned in lecture 5, nitrogenous bases constitute the chromophores of nucleic acids in the near and far UV region. The CD of the stacked bases is larger in magnitude as compared to that of the isolated bases. As the double helical nucleic acids have stacked bases, what we measure essentially is the CD that arises due to coupling of the chromophores. As the stacking geometries are different for different forms of nucleic acids such as B-DNA, Z-DNA, and A DNA; CD can help in determining which DNA form is present in a given sample.

Applications in biomolecular analysis

i. Determination of protein/peptide structure: As has already been discussed earlier, far UV CD spectroscopy provides information about the secondary structural elements in a protein. A mixture of structures can be deconvoluted to obtain the fraction of different structural elements. Furthermore, near UV CD provides information about the tertiary structure of the protein.

ii. Comparison of structures: Mutants of proteins are often required for understanding the functions of the proteins. It, however, needs to be ascertained that the mutation does not cause any significant change in the overall structure of the protein. CD spectroscopy happens to be a fast and extremely reliable tool to compare the conformations of the wild type proteins with their mutants.

iii. Stability of proteins: Stability of the proteins to denaturants or heat can be studied using CD spectroscopy. In such studies CD is usually monitored at a single wavelength, typically around 220 nm. Plotting the change in ellipticity against increasing denaturant concentration/temperature provides the denaturation curve. Figure 9.3 shows the denaturation curves for three related proteins. The denaturation curves suggest that the protein indicated with the blue trace is most stable while the one indicated with red trace the least.

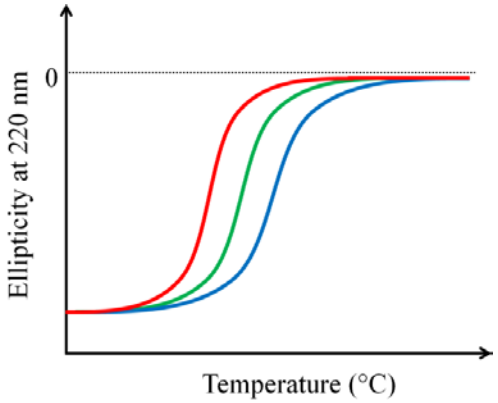


Figure 9.3 Comparison of thermostability of three related proteins. The blue trace represents the most stable protein.

iv. Binding of ligands to proteins: Binding of a ligand to a protein usually does not affect the secondary structural elements significantly. However, such a binding can cause changes in the local tertiary structure. Binding of ligands accompanying such conformational changes can be studied using tertiary CD if the binding region happens to have one or more aromatic residues. Short peptides, on the other hand, can undergo large scale structural changes sometime involving completely switching from one secondary structure to another. Such changes can easily be observed using far UV CD.

v. DNA structure: CD in the 200 – 300 nm region can be used to identify which structural isoform of DNA is present in the given sample. The left-handed helical DNA form, the Z-DNA was indeed identified using CD spectroscopy. The typical CD signatures of the B, Z, and A form of DNA are:

B-DNA: In its most common form *i.e.* B-DNA with ~10.4 bases per turn, a positive band ~275 nm, a crossover ~258 nm, and a negative band at ~240 nm are observed.

Z-DNA: A negative band ~290 nm and a positive band ~260 nm; a crossover between 180-185 nm.

A-DNA: A positive band ~260 nm, a negative band ~210 nm.

vi. Protein folding/unfolding: CD is used for studying the folding and unfolding of proteins. For monitoring the fast reactions such as protein folding, a single wavelength CD is recorded in a stopped flow experiment wherein the protein solution is mixed with a denaturant and CD is recorded as a function of time. Modern instruments take ~1 millisecond time between mixing and recording data allowing the understanding of the folding/unfolding events that occur on milliseconds to seconds timescale. A diagrammatic unfolding experiment is shown in figure 9.4

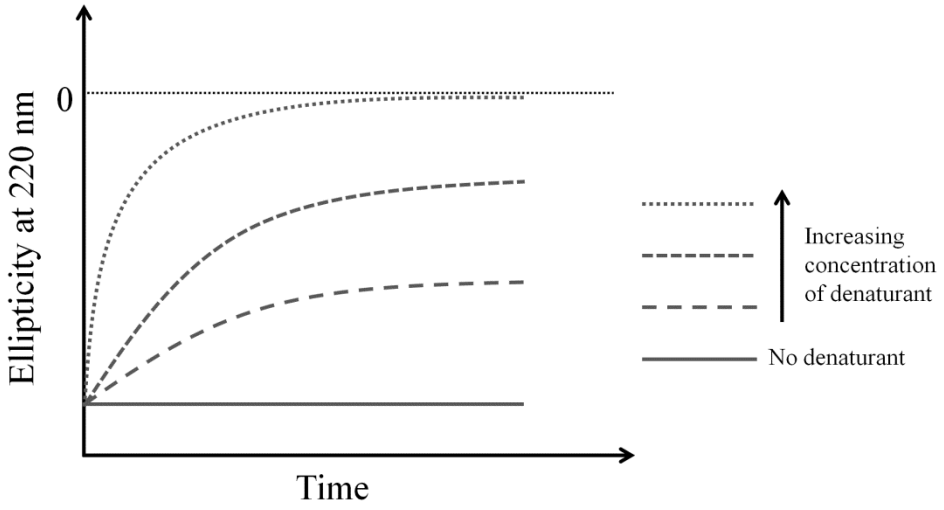


Figure 9.4 A diagram showing the kinetics of unfolding of a hypothetical protein. The protein is unfolded with different concentrations of a denaturant. Protein and denaturant are mixed in a stopped flow apparatus (mixing time typically ~1 ms) and changes in ellipticity are monitored over time.

vii. Molecular self-assembly: Self-assembly into structural and functional superstructures is integral to biomolecules and therefore to living systems. Inspired by the naturally occurring superstructures, short peptides have attracted considerable attention as the monomers for designing superstructures with novel properties and applications in biomedicine. Circular dichroism has been central in elucidating the conformations of the peptides in superstructures as well as the interactions that drive this assembly.

Circular dichroism, therefore, is a powerful tool in studying the conformations of biomolecules as well as the processes these molecules are involved in.

Lecture 10 Infrared Spectroscopy

Introduction

Infrared (IR) region of the electromagnetic spectrum lies between visible and microwave regions and therefore spans the wavelengths from 0.78 – 250 μm. The energies associated with molecular vibrations are smaller than those associated with electronic transitions and fall in the IR region. IR spectroscopy, therefore, is used to probe the vibrations in molecules and is also known as vibrational spectroscopy. Infrared region is usually divided into three regions: near infrared, mid-infrared, and far infrared (Figure 10.1). IR spectroscopists use wavenumbers ($\bar{\nu}$) to represent the IR spectra and we shall be following the same convention. Mid-IR region (λ = 2.5 -25 μm; $\bar{\nu}$ = 4000 – 400 cm^{-1}) is the region of interest for studying molecular vibrations.

Conventions for IR radiation

Wavelength: The wavelength of IR region ranges from ~780 nm – 250000 nm. Writing such big number is avoided by expressing the wavelengths in micrometers (0.78 – 250 μm).

Wavenumber ($\bar{\nu}$): Wavenumber means the number of wavelengths per unit distance. Therefore, 100 cm^{-1} implies there are 100 wavelengths per cm. $\bar{\nu}$ in cm^{-1} is given by:

$$\bar{\nu}\ (cm^{-1}) = \frac{1}{\lambda\ (\mu m)} \times 10^4$$

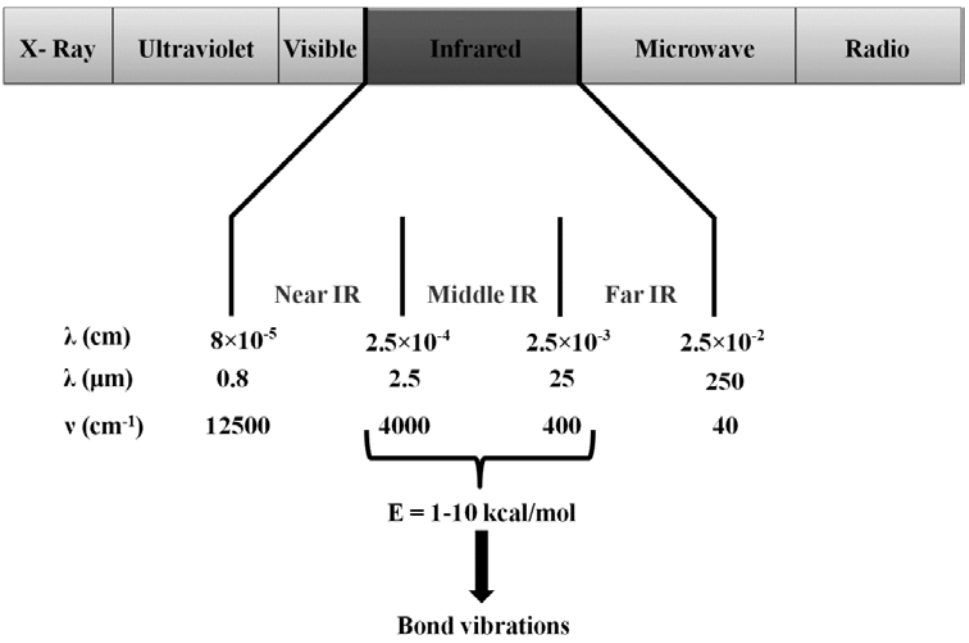


Figure 10.1 Infrared region of the electromagnetic spectrum

Degrees of freedom and molecular vibrations

At non-zero temperatures, *i.e.* temperatures above 0 K, all the atoms in a molecule are in motion. The molecule itself also is in translational and rotation motion. In a three dimensional space, an atom in isolation has 3 degrees of freedom, corresponding to the motion along the three independent coordinate axes. *A molecule composed of N atoms has a total of 3N degrees of freedom* (Figure 10.2).

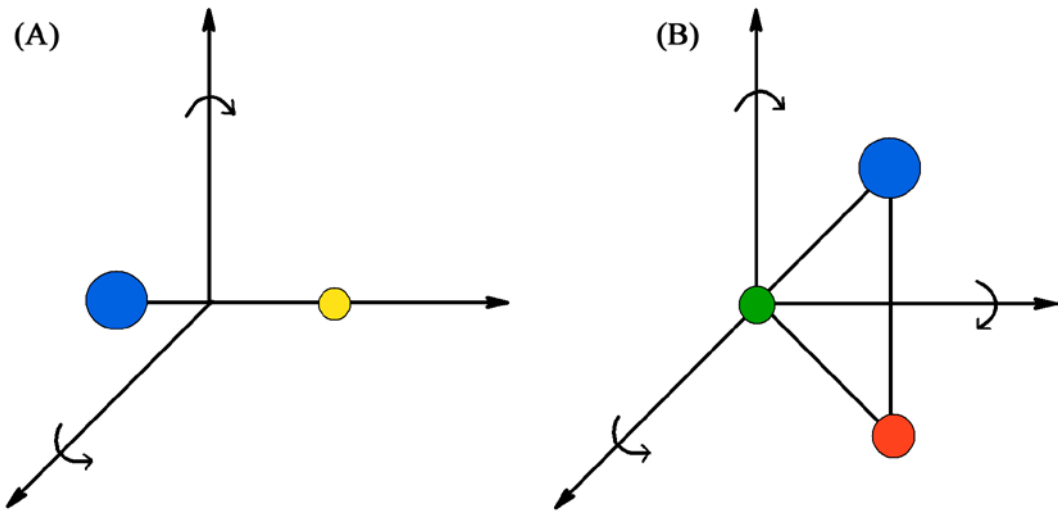


Figure 10.2 Degrees of rotational freedom for a diatomic (A) and a triatomic (B) molecule

For a non-linear molecule, three of these 3*N* degrees of freedom correspond to translational motion, three correspond to rotational motion while rest 3N-6 are the vibrational degrees of freedom. For a linear molecule, there are only two rotational degrees of freedom that correspond to the rotation about the two orthogonal axes perpendicular to the bond (Figure 10.2). A linear molecule, therefore, has 3*N*-5 vibrational degrees of freedom. Let us have a look at the degrees of freedom of a diatomic molecule. A diatomic molecule has a total of 3 × 2 = 6 degrees of freedom. *Three* of these *six* degrees of freedom correspond to translational motion of the molecule; *two* of them define rotational degrees of freedom; while *one* corresponds to the vibration of the atoms along the bond. The 3N-6 vibrational degrees of freedom (3N-5 for linear molecules) represent the true/fundamental modes of vibration of a molecule. The different types of vibrations are shown in Figure 10.3.

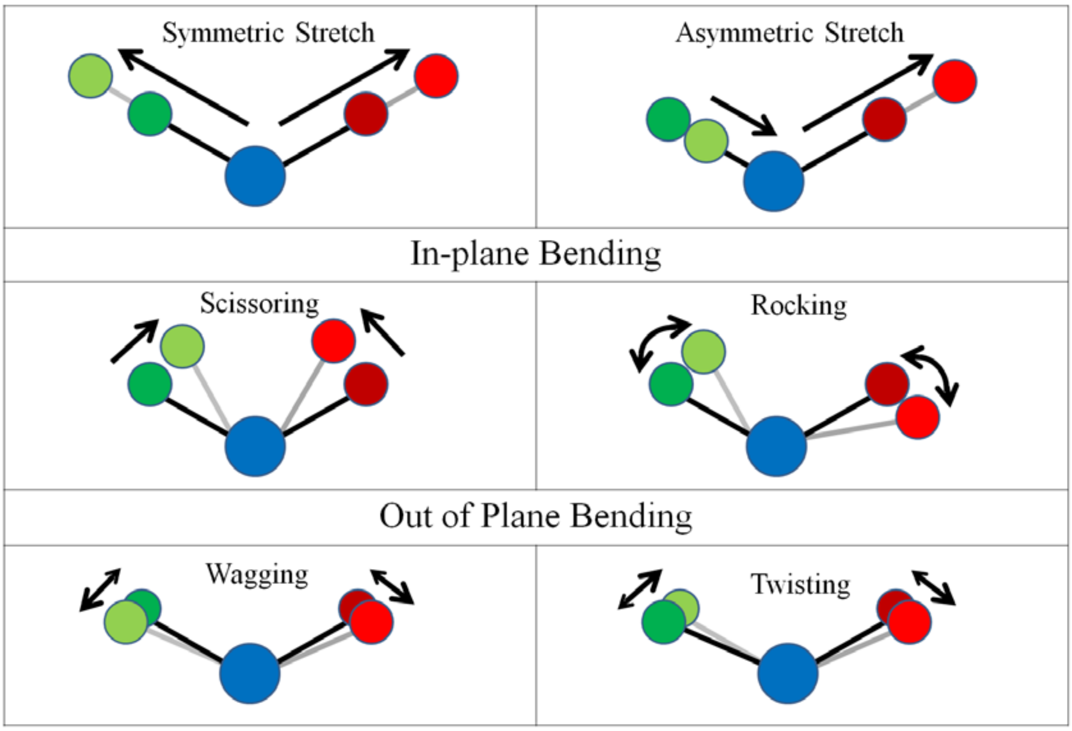


Figure 10.3 Stretching and bending vibrations in molecules

Hooke's law and frequency of vibration

We have seen that the bonds are not static but vibrating in different ways. A vibrating bond can therefore be considered a spring with its ends tethered to two atoms (Figure 10.4).

Figure 10.4 Spring analogy of a bond vibration

If the masses of the atoms are m_1 and m_2, the frequency of stretching vibration of the diatomic molecule can be given by the Hooke's law:

$$\nu = \frac{1}{2\pi}\sqrt{\frac{k}{\mu}} \quad \text{...} \quad (10.1)$$

where, ν is the frequency of vibration, *k* is the spring constant, and μ is the reduced mass *i.e.* $\frac{m_1 m_2}{m_1 + m_2}$

Dividing equation 10.1 by λ gives:

$$\frac{\nu}{\lambda} = \frac{1}{2\pi\lambda}\sqrt{\frac{k}{\mu}} \quad \text{...} \quad (10.2)$$

$$\frac{1}{\lambda} = \frac{1}{2\pi(\lambda\nu)}\sqrt{\frac{k}{\mu}} \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (10.3)$$

$$\bar{\nu} = \frac{1}{2\pi c}\sqrt{\frac{k}{\mu}} \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (10.4)$$

The spring constant, *k* is the measure of the bond strength. The stronger the bond, the higher the *k*, and consequently the higher is the frequency of vibration. This treatment implies that the diatomic molecule is a simple harmonic oscillator. The energy of a quantum harmonic oscillator is given by:

Anharmonic oscillator

Real molecules are anharmonic oscillators. Unlike harmonic oscillator wherein energy levels are equally spaced; energy levels in an anharmonic oscillator are more closely spaced at higher interatomic distances. A treatment for anharmonicity is beyond the scope of our discussion.

$$E = \left(n + \frac{1}{2}\right)h\nu \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (10.5)$$

where, *n* = 0, 1, 2, …… and *h* is the Plancks's constant

Absorption of infrared radiation

A molecular vibration is IR active i.e. it absorbs IR radiation if the vibration results in a change in the dipole moment. A diatomic molecule, that has one mode of vibration, may not absorb an IR radiation if the vibration does not accompany a change in the dipole moment. This is true for all the homonuclear diatomic molecules such as H_2, N_2, O_2, etc. Vibration of carbon monoxide (C=O), on the other hand, causes a change in dipole moment and is therefore IR active. Vibration of a bond involving two atoms that have large electronegativity difference is usually IR active.

An IR active vibration of a particular frequency absorbs the IR radiation of same frequency. Let us calculate the position of absorption band for carbonyl stretching vibration (frequency = 5.1×10^{13} vibrations/second) in acetone.

$$\bar{\nu} = \frac{1}{\lambda} = \frac{\nu}{c}\ \text{cm}^{-1}$$

$$\bar{\nu} = \frac{5.1 \times 10^{13}\ sec^{-1}}{3 \times 10^{10}\ cm/sec} = 1700\ \text{cm}^{-1}$$

Instrumentation

Two types of infrared spectrometers are commercially available: dispersive and Fourier Transform infrared (FTIR) spectrometers.

Dispersive spectrometer: A dispersive spectrometer is very similar in design to a UV/visible spectrophotometer. It has a radiation source, a grating monochromator, and a detector. The IR radiation generated by the source is dispersed into different frequencies by a monochromator. The selected frequencies go through sample and reference cells and the transmitted light is measured by the detector. The infrared sources are usually inert solids that are electrically heated to radiate infrared radiation. The detectors usually are either thermal sensors such as thermocouples and thermistors or the semiconductor materials that conduct following absorption of IR radiation (absorption of photon causes transition of electrons from the valence band to the conduction band).

Fourier Transform Spectrometer: A Fourier transform spectrometer uses an interferometer in place of the monochromator. An interference of polychromatic radiation is generated using an interferometer, usually a Michelson interferometer (Please see Box 10.1). Absorption of any particular wavelength will bring a change in the interferogram which gets detected. An interferogram is a time domain signal and is converted to frequency domain signal though Fourier Transformation.

Dispersive infrared spectrometers are still in use but FTIR spectrometers are slowly taking over. FTIR spectrometers have several advantages over the dispersive ones:

i. Better speed: FTIR spectrometers detect absorption of all the frequencies simultaneously; consequently, they are much faster than the dispersive spectrometers that scan the entire frequency range stepwise.

ii. Better sensitivity: Their speed of data acquisition makes FTIR spectrometers more sensitive. A large number of spectra can be recorded in small time thereby giving an improved signal to noise $\left(\frac{S}{N}\right)$ ratio.

$$\left(\frac{S}{N}\right) \propto \sqrt{n} \qquad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (10.6)$$

where, *n* is the number of independent measurements

iii. More radiation energy: Dispersive spectrometers use slits that result in loss of radiation energy. FTIR spectrometers lack the slits as filtering of radiation is not required; this provides higher radiation energy for recording the absorbance.

iv. Simple design: As dispersion of the radiation and filtering are not required in the FTIR spectrometer, the movable mirror is the only moving part in the spectrometer.

v. Wavelength accuracy: The FTIR spectrometers usually have a He-Ne laser emitting light of 632.8 nm. This serves as an internal calibration for the wavelength and provides an accuracy of 0.01 cm^{-1} or better.

Attenuated Total Reflectance – Fourier Transform Infrared Spectrometer (ATR-FTIR)

An ATR-FTIR works on the principle of total internal reflection and the evanescent field (Figure 10.5). The refractive index of the ATR crystal (usually Zinc selenide, diamond, Germanium) is significantly higher than that of the samples that are to be studied. An IR beam gets refracted at the interface of the ATR material and the sample (Figure 10.5A).

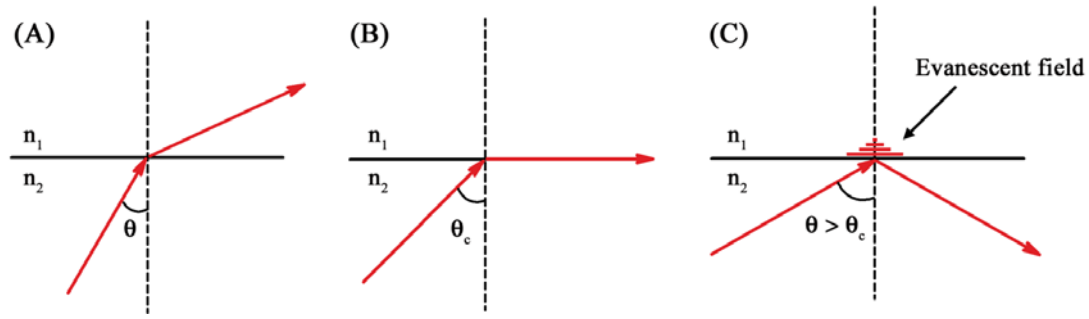


Figure 10.5 Phenomenon of total internal reflection. Notice the exponentially decaying evanescent field in the medium of lower refractive index (n_1)

If the angle of incidence, θ is more than the critical angle, θ_c (Figure 10.5B and C), the beam gets totally internally reflected. Before getting totally reflected, however, an exponentially decaying field penetrates into the medium of low refractive index. This field is called an evanescent field and can interact with the molecules that are in the close proximity of the ATR crystal. If some part of the evanescent field is absorbed by the molecules, the reflected beam will be attenuated (become less intense) by this factor. The reflected beam will therefore be of lesser intensity implying absorption of radiation. The commercially available ATR instruments are FTIR spectrometers. ATR-FTIR allows studying the samples like thin films, powders, pastes by directly placing the sample on the ATR crystal.

Functional group region and fingerprint region

The most common application of IR spectroscopy is perhaps to identify the functional groups. This is possible because different functional groups vibrate at different frequencies allowing their identification. The frequency of vibration, however, depends on additional factors such as delocalization of electrons, H-bonding, and substitutions at the nearby groups. The wavenumbers for some of the bonds are shown in Table 10.1.

Table 10.1 Typical vibrational frequencies of functional groups

Bond	**Molecule**	**Wavenumber (cm^{-1})**
C–O	Alcohols, ethers, esters, carboxylic acids, etc.	1300 – 1000
C=O	Aldehydes, ketones, esters, carboxylic acids	1750 – 1680
C=O	Amides	1680 – 1630
N–H (Stretching)	Amines and amides	3500 – 3100
–N–H (Bending)	Amines and amides	1640 – 1550
O–H	Alcohols	3650 – 3200
C–N	Amines	1350 – 1000
S–H	Mercaptans	2550

The absorption bands in the 4000 – 1500 cm^{-1} region help in the identification of functional groups; this region therefore is also termed the functional group region of the IR spectrum (Figure 10.6). The lower energy portion of the mid-IR region (1500 – 400 cm^{-1}) usually contains a very complicated set of peaks arising due to complex vibrations involving several atoms. This region is unique to a particular compound and therefore is known as the fingerprint region of the IR spectrum. Though it is

difficult to assign the vibrational modes to these peaks, these are useful to identify a compound if the spectrum of the compound is already known.

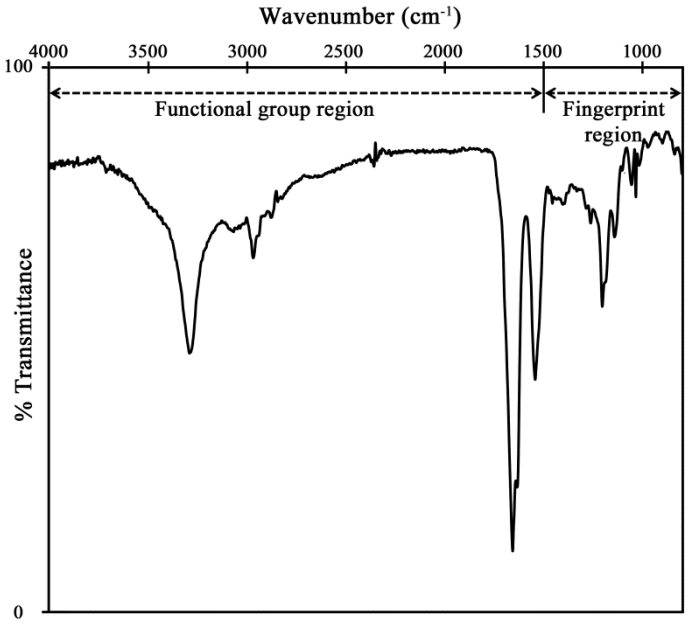


Figure 10.6. A typical IR spectrum showing functional group and fingerprint regions

Applications

i. *Identification of functional groups:* As has already been discussed, IR spectroscopy allows identification of functional groups. Carbonyl (C=O) is an interesting functional group worth discussing. Carbonyl is a double bond (high spring constant, *k*) with very high polarity. Stretching vibration of carbonyl group causes large changes in the dipole moment consequently resulting in a very intense absorption band. Furthermore, the frequency of carbonyl stretching does not differ significantly for aldehydes, ketone, carboxylic acids, and esters (Table 10.1). The large intensity and relatively unchanged frequency of carbonyl stretching allows easy identification of the carbonyl compounds (It is important to note that carbonyl stretching frequency can be much lower for amides and much higher for anhydrides and acid chlorides).

ii. *Identification of compounds:* The fingerprint region of the IR spectrum is unique to each compound. It is possible to identify a compound from its IR spectrum if the spectrum for the compound is already known and available for comparison. This is particularly useful in pharmaceutical research and development. A patented drug, if suspected to be synthesized by another

pharmaceutical company, can easily be identified by comparing the IR spectra in the fingerprint region.

iii. *Presence of impurities:* Comparison of the IR spectra of the given compound with the spectra of pure compound helps in the assessment of its purity. It is important to ascertain the purity of the active molecule and the excipients used in preparing drug formulations.

iv. *Structural transitions in lipids:* Structural lipids are those that are organized in bilayers in biological membranes. Glycerophospholipids constitute the major class of the structural lipids (Figure 10.7). The lipids have several structural phases such as a gel phase with all-*trans* conformation and a liquid crystalline phase where *gauche* conformations are also present. Methylene ($–CH_2–$) stretching vibrations give the most intense absorption band in lipids as expected for a molecule having long hydrocarbon chains.

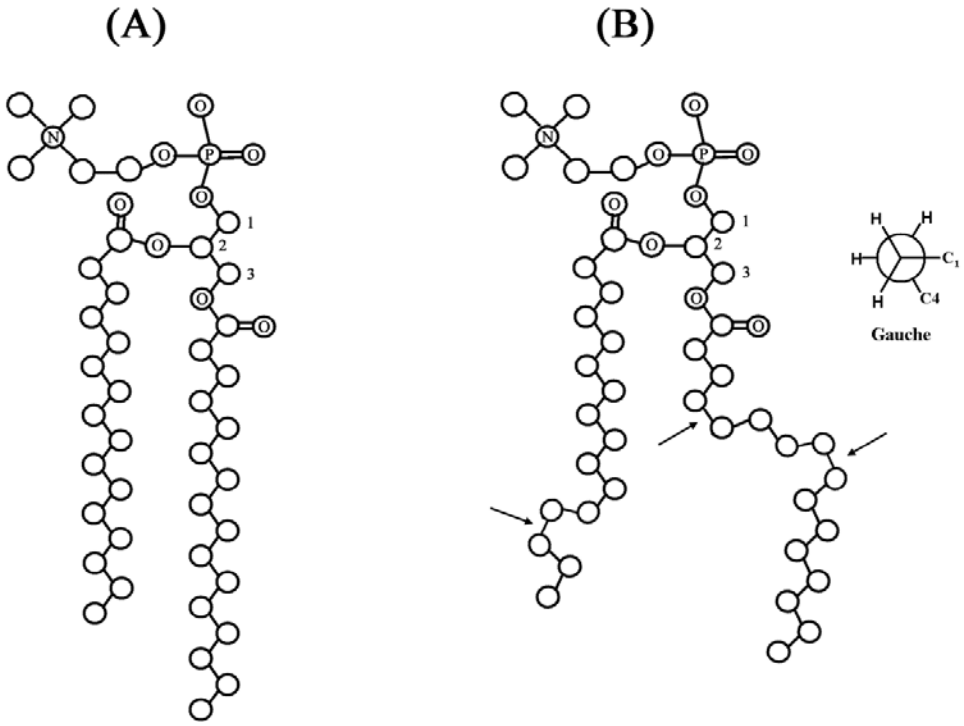


Figure 10.7. Structure of a glycerophospholipid: all-*trans* conformation (A); lipid with *trans* and *gauche* conformations (B), *gauche* conformations are indicated with arrows

Both $–CH_2–$ stretching and bending vibrations are sensitive to the conformations of the lipids and therefore provide information about the transition of lipids between different phases. Vibration modes of the head group and the interfacial region also provide useful information about local acyl chain conformation. Carbonyl stretching vibration (1750 – 1700 cm^{-1}) in the ester bond is sensitive to the conformation of the local acyl chain conformation.

v. *Protein and peptide structure:* Infrared spectroscopy is routinely used to study the structures of proteins and peptides. Like CD spectroscopy, the region of interest in determining the conformation of the polypeptide backbone is the peptide bond. The peptide group results in nine distinct bands, labeled as amide A, B, and I-VII. Amide I is the most useful band in studying the polypeptide backbone conformation. Amide I band (1700 – 1600 cm^{-1}) arises largely due to the carbonyl stretching with small components of C–N stretching and N–H bending. The frequency of carbonyl stretching vibration is sensitive to the H-bonding, and therefore to the conformation of polypeptide backbone. The frequencies of absorption of different secondary structural elements are shown in Table 10.2

Table 10.2 Vibrational frequencies of the secondary structural elements of proteins in H_2O

Structure	**Wavenumber (cm^{-1})**
α-helix	1657 – 1648
β-sheet	1641 – 1623
Unordered	1657 – 1642
Antiparallel β-sheet	1695 – 1675

There is considerable overlap of the bands arising from α-helical and the unordered conformations. It is therefore generally difficult to assign the bands appearing in this region. Recording an IR spectrum in D_2O decreases this overlap to some extent. Dissolution a protein in D_2O results in the exchange of solvent exposed amide protons by deuterium. Hydrogens of the unordered amides are more easily exchanged as compared to those involved in the secondary structures. Despite this, it is not easy to unambiguously assign the bands arising in the 1657- 1648 cm^{-1} region. Circular dichroism and IR spectroscopy therefore complement each other wherein α-helices are easily detected by CD and β-sheets by IR. Like CD, an IR spectrum of a protein can also be deconvoluted to determine the fractions of different secondary structural elements as shown in Figure 10.8.

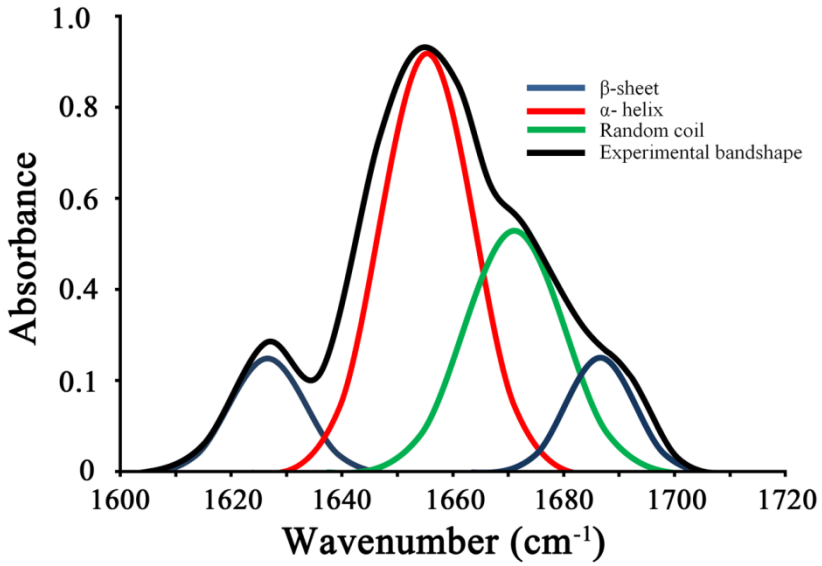
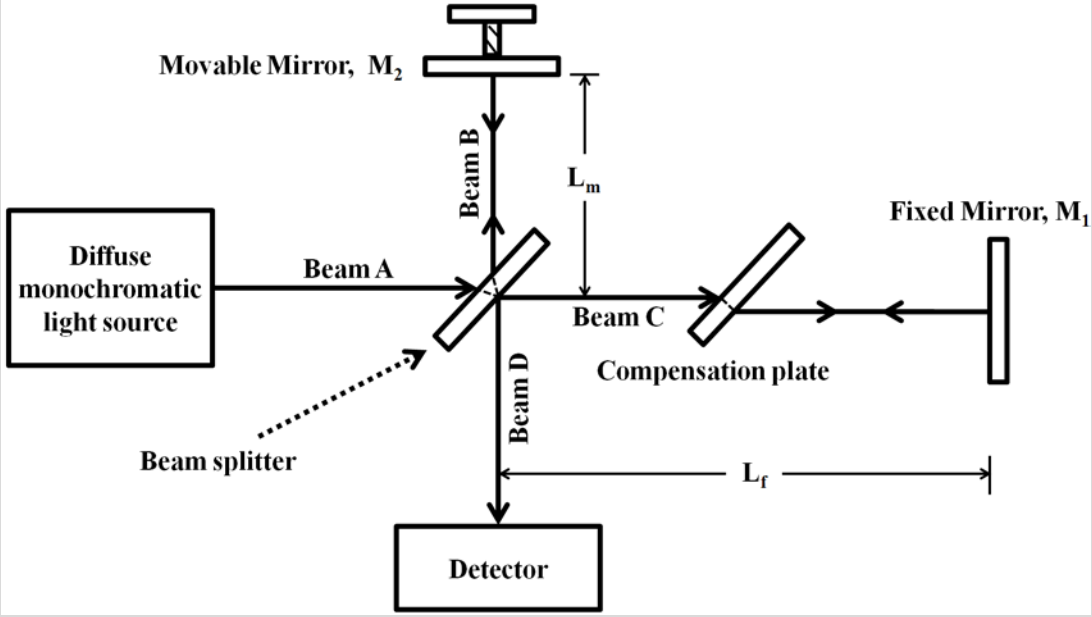


Figure 10.8. Deconvolution of Amide I band of a protein to identify the fractions of different structural elements

Box 10.1: Michelson interferometer

A Michelson interferometer has a radiation source, a collimator, a beam-splitter, a movable mirror, a fixed mirror, a compensator, and a detector.



The radiation coming from the source is collimated and focused on the beam-splitter. 50% of the radiation gets transmitted while 50% gets reflected. The mirrors reflect back the radiation towards the beam splitter that again allows 50% transmission and 50% reflection. This allows the beams, B and C to interfere and give the beam D. As the beam, B travels through the beam-splitter twice while beam C does not travel through it even once, a compensation plate of same material (un-mirrored) and thickness as the beam-splitter is, is placed between the beam-splitter and the fixed mirror. This allows the beams, B and C to travel the equal distance. The motion of the movable mirror, M_2 causes the two beams to travel different distances thereby generating interference. Let us see what happens when a monochromatic radiation is used in the Michelson interferometer. If the beams, B and C travel the equal distance, they are in phase and will interfere constructively. If however, the M_2 moves, say towards the beam-splitter by a distance of $\frac{\lambda}{4}$, the beam B travels a distance of $\frac{\lambda}{2}$ less than that travelled by beam C. This allows a phase difference of 180° resulting in destructive interference. A continuous motion of the mirror M_2, therefore, will generate a sinusoidal signal through interference. The detector therefore detects a time domain signal. If a sample placed before the

detector absorbs this radiation, the intensity of the light goes down. If a polychromatic light is used, the interference pattern generated carries all the wavelengths present in the polychromatic light. Absorption of any wavelength will result in a change in the interfering pattern. The interfering pattern, also known as interferogram is then Fourier transformed to obtain the frequency domain data.

QUIZ

Q1: If the stretching frequency of a hydrogen molecule is 1.2 × 10^{14} vibrations/sec. Calculate the wavenumber where hydrogen molecule absorption band will be observed in an IR spectrum.

Ans: The frequency of hydrogen stretching can be represented in terms of wavenumbers as follows:

$$\bar{\nu} = \frac{1}{\lambda} = \frac{\nu}{c} \text{ cm}^{-1}$$

$$\bar{\nu} = \frac{1.2 \times 10^{14}\ sec^{-1}}{3 \times 10^{10}\ cm/sec} = 4000 \text{ cm}^{-1}$$

Hydrogen, however, is a homodiatomic molecule; the stretching vibration does not cause any change in the dipole moment. Therefore, hydrogen will not absorption the IR of 4000 cm^{-1} and consequently will not appear in an IR spectrum.

Lecture 11 Mass Spectrometry-I

Mass spectrometry (abbreviated as MS) has slowly emerged as a very powerful tool in analyzing the organic molecules including biomolecules. A mass spectrometer separates the molecules based on their mass and charge. The underlying principle is conceptually very simple: a moving charged particle can be deflected by applying electric and magnetic fields. The deflection caused by the electric and magnetic fields depends on the mass and the charge of the particle. Let us see what happens to a charged particle in an electric field (Figure 11.1).

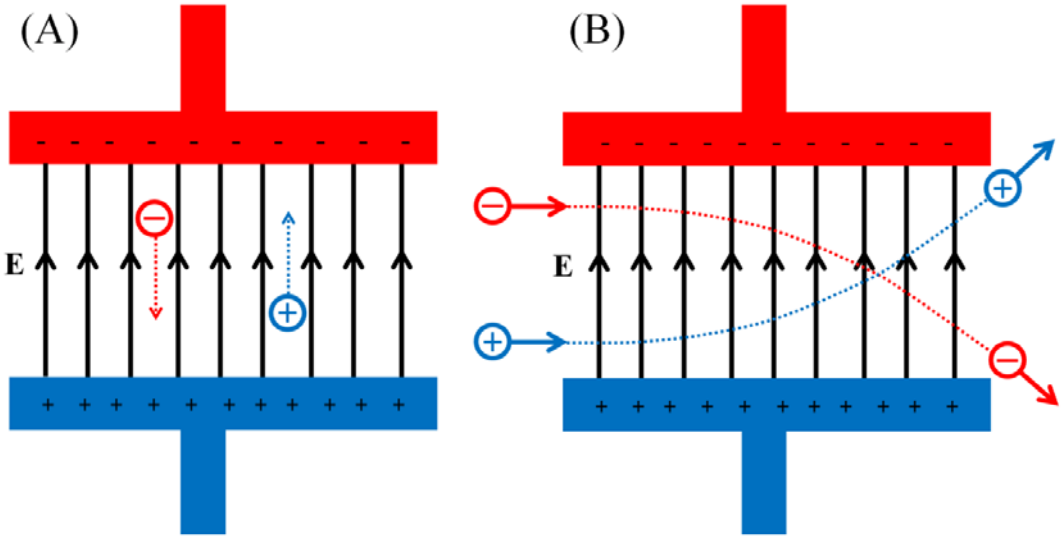


Figure 11.1 Force experienced in an electric field by stationary charged particles (A) and charged particles with uniform motion (B). Panel B represents the situation when the charged particles are in uniform motion with their initial velocity vectors perpendicular to the electric field vector.

The force experienced, F by a particle with charge, q in an electric field, E is given by:

$$F = qE \qquad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (11.1)$$

The force causes the particle to accelerate in the electric field which is given by

$$ma = qE \qquad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (11.2)$$

$$a = \frac{qE}{m} \qquad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (11.3)$$

where, m is the mass of the particle and a is the acceleration under the electrostatic force.

Equation 11.3 shows that the acceleration of the particle depends on the mass to charge ratio, $\frac{m}{q}$. A lighter particle is accelerated more than a heavier particle carrying the same charge. Similarly, a particle with higher charge is accelerated more as compared to the particle of same mass but having lesser charge.

In a magnetic field, a moving charged particle experiences a force, *F* that is given by the Lorentz force law:

$$F = q\ (v \times B) \qquad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (11.4)$$

where, v is the velocity of the moving charged particle and B is the magnetic field strength.

The direction of the force can be determined using the right hand rule; If the fingers represent the magnetic field (*B*) and the thumb represents the velocity (v), then the direction of the force is given by the direction of the palm (Figure 11.2A). As the force is always perpendicular to the velocity, the deflected particle moves in a circular path (Figure 11.2B).

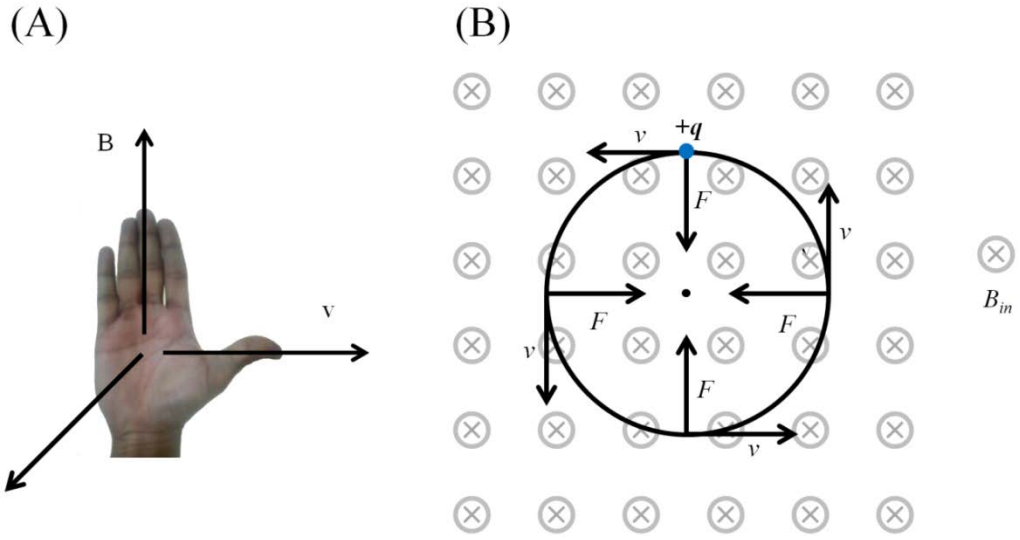


Figure 11.2 The right hand rule showing the direction of the Lorentz force in a magnetic field (A) and the circular path taken by the particle (B).

The Lorentz force therefore provides the particle a centripetal force. Therefore, equation 11.4 can be written as

$$\frac{mv^2}{r} = q\ (v \times B) \qquad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (11.5)$$

where, *r* is the radius of the circular path

Rearranging in terms of *r*:

$$r = \frac{m\,v^2}{qvB\,sin\theta} \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (11.6)$$

$$r = \frac{mv}{qB\,sin\theta} \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (11.7)$$

where, θ is the angle between the velocity vector, *v* and the magnetic field vector, *B*.

In a mass spectrometer, *v* and *B* are generally orthogonal to each other; in that case:

$$r = \frac{mv}{qB} \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (11.8)$$

Equation 11.8 shows that the deflection caused by a magnetic field in a moving charged particle is proportional to the mass to charge ratio. For the two particles having same charge but different masses, the one with lesser momentum deflects more ($r \propto mv$ and smaller r means larger deflection). In mass spectroscopy, the charge is usually represented as *z* and we shall be sticking to the same convention. A mass spectrum is a two dimensional plot between ion abundance and $\frac{m}{z}$ ratio (Figure 11.3)

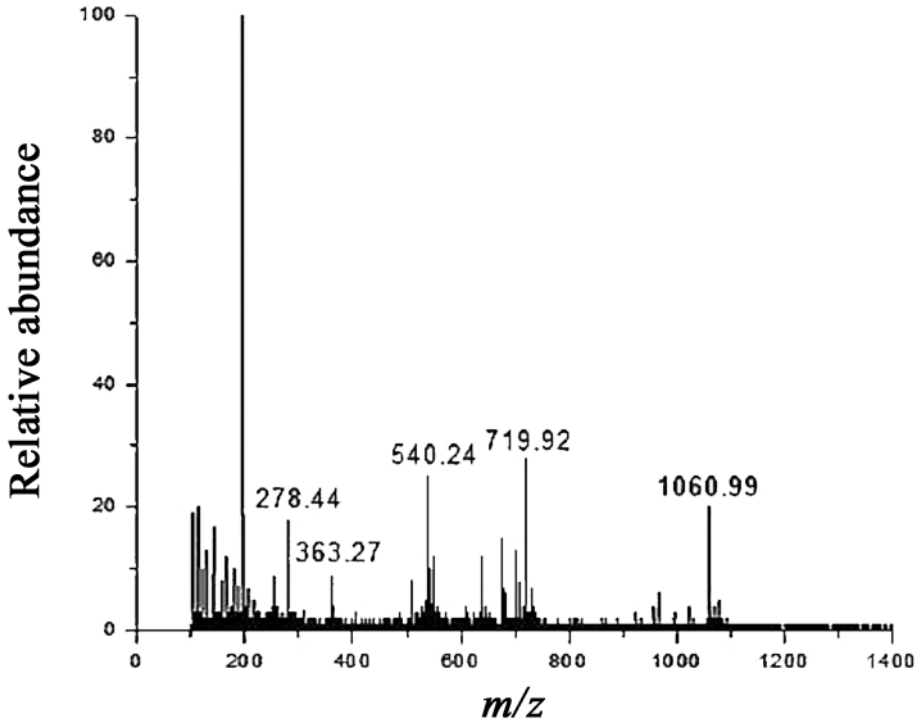


Figure 11.3 A typical mass spectrum

Let us see the design of a typical mass spectrometer (Figure 11.4). The basic requirement for an analyte molecule to be studied using mass spectrometry is that it has to be charged. A large number of molecules, however, may not be charged. The first step in an MS experiment is therefore to ionize the molecules. The spectrometer therefore has an *ionization source*. The ions generated are then separated by one or more *mass analyzers* which are then detected by a *detector*.

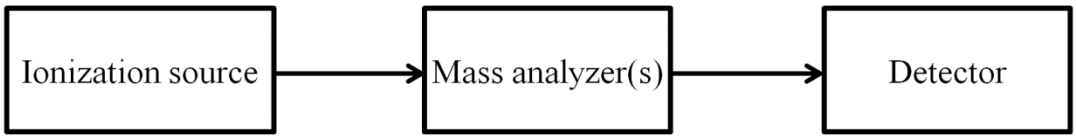


Figure 11.4 The components of a mass spectrometer

Ionization/ionization source

The first step in an MS experiment is to obtain the ions in gas phase. The mass spectrometers, therefore have an ionization chamber (also called ionization source) where the samples are introduced to achieve ionization. Ions are generated through one of the several methods that have their own merits and limitations. Some of the ionization methods are:

Ion mode

Mass spectrometric analyses are usually performed in the positive ion mode *i.e.* only cationic species are detected. It is, however, possible to study the molecules in negative ion mode as well, where anions are detected. Unless mentioned otherwise, it is usually assumed that the analysis is done in positive ion mode.

Electron Ionization (EI)

In electron ionization method (Figure 11.5), a heated filament is used to emit the electrons. The electrons are accelerated through the ionization chamber under the influence of a strong electric field. The sample in gas phase is introduced into the ionization chamber. A high energy electron can knock off an electron from an analyte molecule, M giving a molecular radical cation.

$$M + e^- \rightarrow M^{\bullet +} + 2e^- \qquad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (11.9)$$

$M^{\bullet +}$ is referred to as the molecular ion. Loss of electron is a miniscule loss of mass; therefore mass of $M^{\bullet +}$ equals the mass of the molecule.

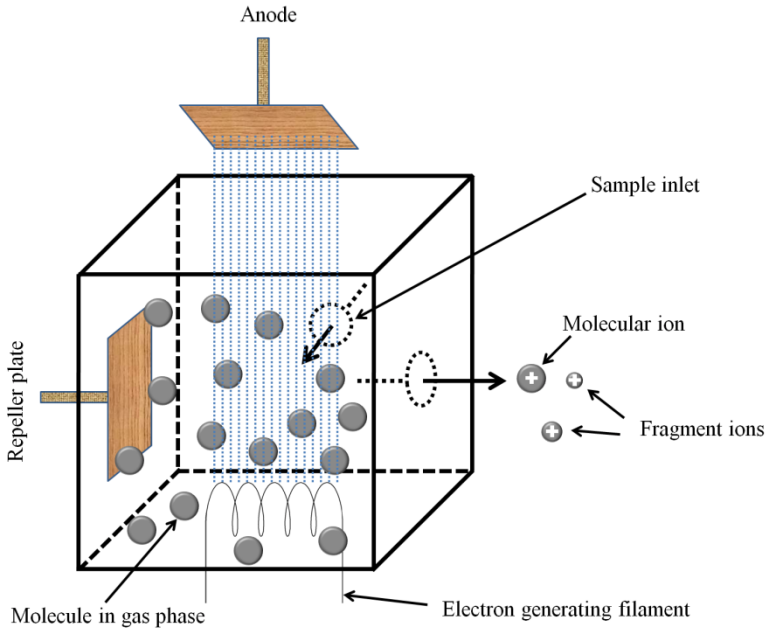


Figure 11.5 Design of an electron ionization source

The kinetic energy of the electrons is usually 70 eV in the electron ionization method. Typically 10-20 eV energy is transferred to the molecules. Around 10 eV energy is sufficient to cause ionization of most organic molecules; the radical cation is therefore left with an excess energy. Electron ionization, therefore often causes extensive fragmentation of the radical cation. Detection of these fragments can provide useful structural information about the molecule but can complicate the data for larger molecules. In some cases, molecular ion may not even be detected at all. The fragmentation is usually hemolytic, resulting in an even-electron cation and a neutral radical (Equation 11.10). Fragmentation into a neutral molecule and a smaller radical cation, however, is not uncommon (Equation 11.11).

$$M^{\bullet+} \rightarrow m_1^{\bullet} + m_2^{+} \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (11.10)$$

$$M^{\bullet+} \rightarrow m_1^{\bullet+} + m_2 \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (11.11)$$

Electron ionization method is limited to samples in the gas phase. Gaseous and highly volatile samples can be directly introduced into the ionization chamber. Liquid and solid samples can be heated to obtain molecules in gaseous phase but it depends on the thermostability of the samples.

Chemical Ionization (CI)

The ions are produced through the collision of the sample molecules with the primary ions produced by a gas (called a reagent gas) in the ionization chamber. The reagent gas is ionized through electron ionization. The radical cations generated will undergo fragmentations and reactions. The most common reaction generating ions is a proton transfer from a gas cation (GH^+) to the molecule. Methane, isobutane, and ammonia are the most common reagent gases. Let us take methane as an example to understand how chemical ionization occurs:

Electron ionization: $CH_4 + e^- \rightarrow CH_4^{\bullet +} + 2e^-$

Fragmentation: $CH_4^{\bullet +} \rightarrow CH_3^+ + H^\bullet$

$CH_4^{\bullet +} \rightarrow CH_2^{\bullet +} + H_2$

The radical cations and the carbocations can react with the reagent gas to give various protonated species: Reactions: $CH_4^{\bullet +} + CH_4 \rightarrow CH_5^+ + CH_3^\bullet$

$CH_3^+ + CH_4 \rightarrow C_2H_5^+ + H_2$

The analyte molecules acquire the protons from any one of these cations:

$$M + CH_5^+ \rightarrow MH^+ + CH_4$$

The ion, MH^+ is called the quasimolecular or pseudomolecular ion as its mass is one unit more than the molecular mass. Saturated hydrocarbons usually ionize through hydride abstraction by the reagent gas cation:

$$RH + CH_5^+ \rightarrow R^+ + CH_4 + H_2$$

The chemical ionization method imparts little excess energy to the molecular ions thereby resulting in lesser fragmentation as compared to the electron ionization method. Furthermore, the degree of fragmentation depends on the reagent gas used for chemical ionization. The fragmentation caused by isobutane and ammonia is considerably less than that caused by methane. Like electron ionization, chemical ionization is also suitable only for gaseous samples limiting it to the gases and volatile liquids.

Fast Atom Bombardment (FAB)

Fast atom bombardment is a soft ionization technique *i.e.* it causes little fragmentation of the molecular ions generated. In fast atom bombardment ionization methods, the sample is dissolved in a non-volatile liquid and the ions are extracted by bombarding the sample with a beam of high energy atoms (~5 keV), usually argon (sometimes xenon). The commonly used liquid matrices include glycerol, thioglycerol, and *m*-nitrobenzyl alcohol. Fast moving Argon atoms are generated as shown in the Figure 11.6. The Argon radical cations ($Ar^{\bullet+}$), generated through electron ionization, are accelerated and focused as a sharp beam. The high energy $Ar^{\bullet+}$ ions are allowed to collide with the Ar atoms resulting in the neutralization of some of the $Ar^{\bullet+}$ ions in the beam. The residual $Ar^{\bullet+}$ in the beam are extracted out by applying an electric field, thereby resulting in a beam of fast moving atoms. The atoms collide with the sample dissolved in the liquid matrix extracting the ions into the gas phase.

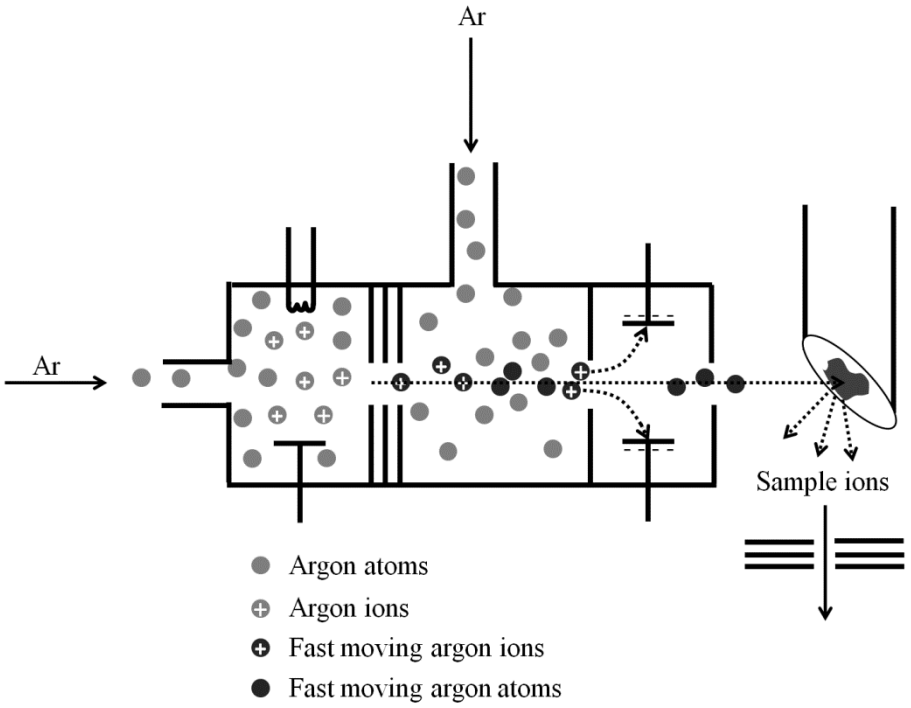


Figure 11.6 Diagram showing the generation of fast moving atoms and sample ionization in FAB ionization source

FAB causes little or no ionization but desorbs the ions already existing in the solution into the gas phase. FAB causes desorption of the ions present on the surface of the matrix; the compounds having higher surface activity are therefore detected better. FAB is particularly good for polar molecules with large molecular weights and molecules up to 10,000 Da can be detected. It is therefore possible to detect biomolecules like peptides, oligonucleotides, and oligosaccharides using FAB ionization.

Liquid secondary ion mass spectrometry (LSIMS)

LSIMS is similar to FAB, with only difference that an ion is used for bombarding the samples. Usually argon, xenon, or caesium ions are used for the LSIMS.

Laser desorption

Laser desorption or laser ablation is the ionization method wherein an intense laser beam is focused on a solid sample resulting in ablation of mass from the surface. The laser pulse causes both desorption and ionization of the molecules. The ions generated are short-lived and therefore detected simultaneously. Laser desorption, however, causes fragmentation for large molecules (Molecular mass >500 Da) therefore restricting its use to small molecules.

Owing to the difficulty in generating gas phase ions of biomolecules like proteins and nucleic acids, the application of mass spectrometry, till late 1980s, was largely restricted to the small organic molecules and biomolecules (amino acids, peptides, oligonucleotides, etc.). Therefore, mass spectrometry was not of much use for biochemists. Advent of two ionization methods around 1987-88 revolutionized the area of biomolecular mass spectrometry. Both these methods are routinely used for identifying and characterizing the biomolecules.

Matrix-assisted laser desorption ionization (MALDI)

MALDI is basically a laser desorption ionization method wherein ionization is assisted by small organic molecules, called matrix. The matrices used in positive ion mode MALDI mass spectrometry are organic acids that have strong absorption for the wavelength of the laser used. Some of the commonly used MALDI matrices are listed in Table 11.1.

Table 11.1 Some of the MALDI matrices commonly used for biomolecules	
Matrix	**Analyte**
α-Cyano-4-hydroxycinnamic acid (CHCA or HCCA)	Proteins, peptides, lipids, oligonucleotides
2,5-Dihydroxybenzoic acid (DHB)	Proteins, peptides, oligonucleotides, oligosaccharides
3,5-Dimethoxy-4-hydroxycinnamic acid (Sinapinic acid)	Proteins, peptides, lipids
3-Hydroxypicolinic acid (HPA)	Oligonucleotides
Trihydroxyacetophenone (THAP)	Oligonucleotides, oligosaccharides

The sample for MALDI is usually prepared in one of the following ways:

i. mixing the analyte solution with the matrix solution → deposition of the mixture on a metallic plate → complete drying of the sample
ii. Deposition of the matrix on the metallic plate → drying of matrix → addition of analyte solution → drying of analyte solution

Desorption and ionization is achieved by applying the laser pulse on the dried sample. Although the exact mechanism behind MALDI is not completely understood, it is believed that the absorption of light by the matrix molecules causes sublimation of matrix crystals carrying along with them the analyte molecules into the gas phase (Figure 11.7).

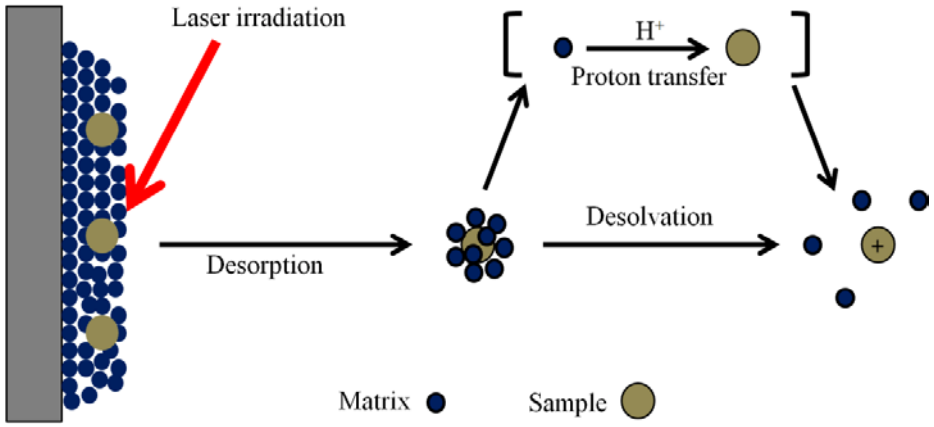


Figure 11.7 Diagram showing ionization in a MALDI source

MALDI can very efficiently generate the gas phase ions from a variety of non-volatile and thermolabile molecules such as proteins, carbohydrates, nucleic acids, and synthetic polymers. MALDI can desorb and ionize the molecules as large as 300 kDa. MALDI usually results in the molecular species having only one charge. In positive ion mode, a quasimolecular ion is formed by protonation of the molecule (MH^+). The samples or the matrices can have trace amounts of alkali metal ions, often resulting in the quasimolecular species, MNa^+ or sometimes MK^+ (Figure 11.8). Multiply charged species are also observed sometimes.

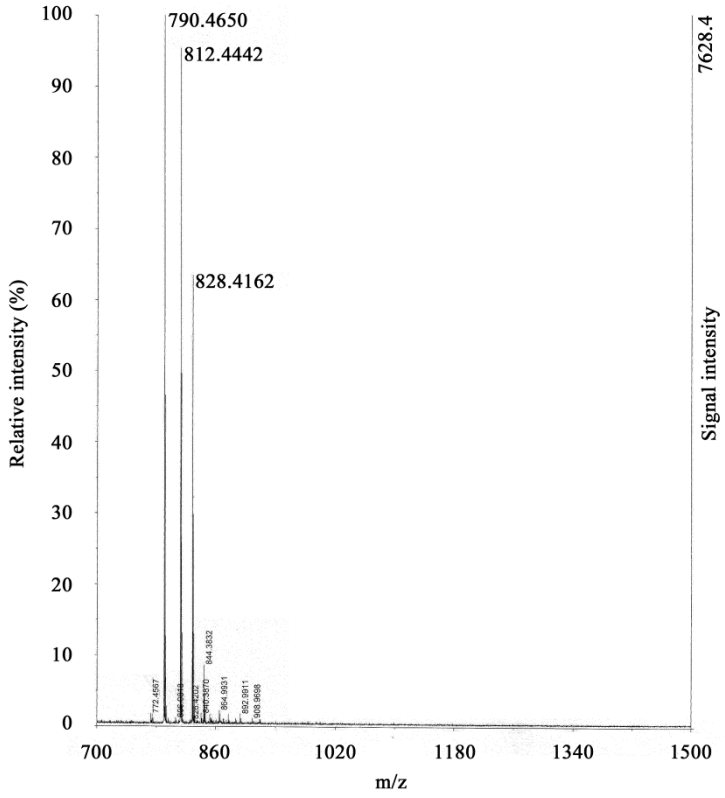


Figure 11.8 MALDI-TOF mass spectrum of a peptide with theoretical mass of 790.448 Da. The peaks at 812.4442 and 828.4162 represent [peptide–Na^+] and [peptide–K^+], respectively.

Time-of-flight (TOF) mass analyzers are well suited for the pulsed ionization techniques such as MALDI (In MALDI, laser pulses produce the ions in pulses). We shall be discussing the different mass analyzers in the next lecture.

Electrospray ionization (ESI)

In electrospray ionization, the analyte solution enters the ionization chamber, maintained at atmospheric pressure, through a fine capillary. The typical flow rates are ~1-20 µl/min. A potential difference of ~3-6 kV is applied between the capillary and the counter-electrode that is ~0.3 – 2 cm away. Under this electric field, the sample droplets appearing at the capillary end accumulate large amount of charge. If the potential of the capillary is above a threshold voltage, the drop will be dispersed into a very fine spray. A coaxial sheath is present around the capillary through which dry nitrogen is supplied for better nebulization and restricting the dispersion of the spray in space. Evaporation of the solvent causes the droplets to diminish in size and their charge density to increase (Figure 11.9). The high charge density on these droplets can further result in the production of smaller droplets. The droplets keep losing the solvent ultimately resulting in the desorption of molecular ions from the surface. As the ions are generated from the surface, a surface active molecule will be detected better. For large molecules such as proteins, the molecules do not desorb from the droplets but become ionized through complete evaporation of the solvent.

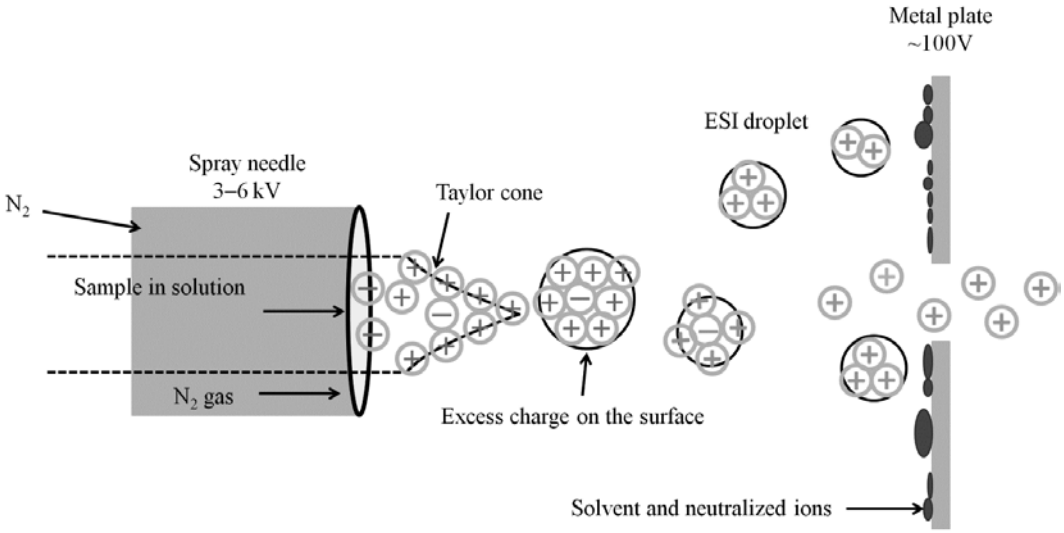


Figure 11.9 Diagrammatic representation of ionization in an ESI ion source

ESI typically results in multiply-charged molecular species. Biological macromolecules usually result in the mass spectra wherein the consecutive peaks differ by one charge unit. This information allows one to calculate the mass of the biomolecule.

Lecture 12 Mass Spectrometry-II

Following ionization, the gas phase ions are accelerated towards the mass analyzers. A great variety of mass analyzers are utilized in the mass spectrometers. All of these analyzers separate the molecules using static or dynamic electric fields and magnetic fields, alone or in combination. After mass analysis, the ions are detected and the mass spectra are generated.

Imagine what would happen if the ions collide with the air molecules in the MS tube. This can lead to the loss of charge to the molecules in the air, deviation in the ion trajectory which might lead to the collision with the MS tube, reactions with the air molecules, etc. A mass spectrometer, therefore operates under very high vacuum to ensure that the ions reach the detector without colliding with the air molecules. The mean free path of a particle is the average distance a particle travels before colliding with other particles and is inversely proportional to the pressure of the gas and the size of the colliding molecules. The mean free path, according to the kinetic theory of gases is given by:

$$\lambda = \frac{kT}{\sqrt{2}\,\pi d^2 p} \qquad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (12.1)$$

where, λ is the mean free path, k is the Boltzmann constant (1.38×10^{-23} JK^{-1}), T is the temperature, d is the sum of the radii of the ion and the colliding molecule, and p is the pressure.

Let us try calculating the mean free path for a small ion in the air. We need to make a few assumptions: Air is largely nitrogen (~78%) and oxygen (~21%). The Van der Walls radii for nitrogen and oxygen are 155 pm and 152 pm, respectively. As the Van der Walls radii of oxygen and nitrogen are very close, we can assume air to be composed of a particle with ~150 pm. Let us calculate the mean free path for a small molecular ion, CH_4^+ (molecular radius ~380 pm). The mean free path at room temperature ($T \approx 298$ K) and atmospheric pressure (1.01×10^5 Pa) can be given by:

$$\lambda = \frac{(1.38 \times 10^{-23}) \times 298}{\sqrt{2}\,\pi \times (530 \times 10^{-12})^2 \times (1.01 \times 10^5)} \qquad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (12.2)$$

$$\lambda = \frac{4.11 \times 10^{-21}}{1.26 \times 10^{-13}} \qquad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (12.3)$$

$$\lambda = 3.26 \times 10^{-8} \text{ m} = 32.6 \text{ nm} \quad \cdots\cdots(12.4)$$

In a mass spectrometer, the ions have to travel large distances (usually >1 m). It is therefore absolutely essential to apply large vacuum for increasing the mean free path by several orders of magnitude. Let us now have look at some of the important mass analyzers:

Magnetic sector

Figure 12.1 shows a diagram of the magnetic sector analyzer mass spectrometer. The ions (say, cations) generated in the ionization chamber are accelerated under a strong electric field. The accelerated ions are allowed to pass through a narrow slit resulting in a sharply focused ion beam. The ions in the beam can be deflected by applying a magnetic field perpendicular to the velocity of the ions.

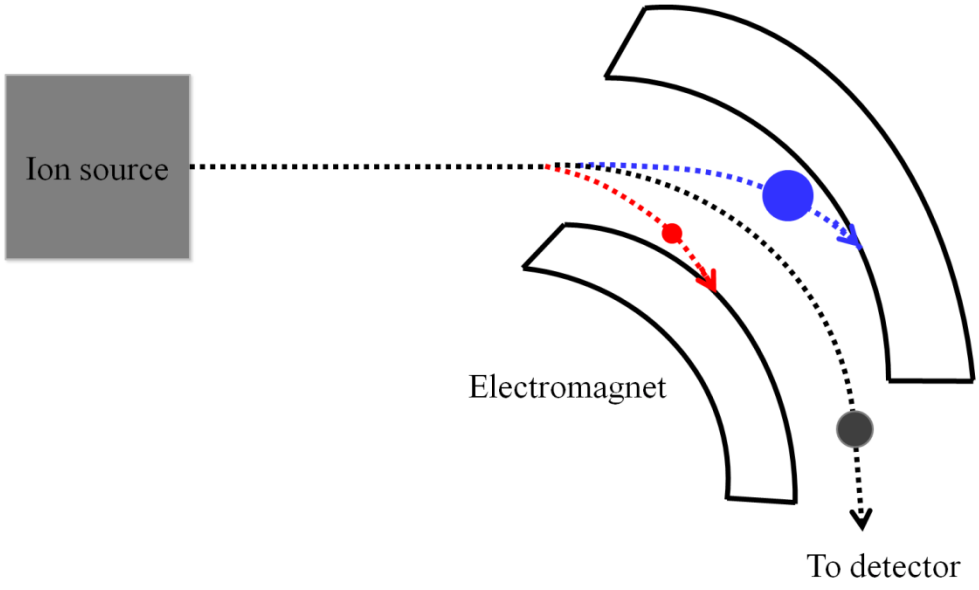


Figure 12.1 Diagram of a magnetic sector mass analyzer.

Too small (blue dotted line, Figure 12.1) or too much deflection (red dotted line, Figure 12.1), which is determined by the charge and the momentum of the ions (Equation 11.8), causes the ions to collide with the MS tube. Therefore, the ions within a small $\frac{m}{z}$ range will be allowed to go to the detector. It is possible to sequentially allow all the ionic species to reach the detector by gradually varying the magnetic field. Assume that the magnetic field strength is zero initially. All the ions will move straight (Equation 11.8) and collide with the curved MS tube losing their charge. If the magnetic field is slowly increased from zero to the maximum value, the

ions with lowest momentum and highest charge will appear first while the ions with highest momentum and lowest charge will appear last.

Time of flight (TOF)

In a time of flight mass analyzer, the time taken by the ions to reach the detector is measured. The ions are generated in bundles *e.g.* by MALDI. The ions are then accelerated towards the flight tube. The flight tube does not have any electric field and the ions drift in the flight tube according to their velocities (Figure 12.2).

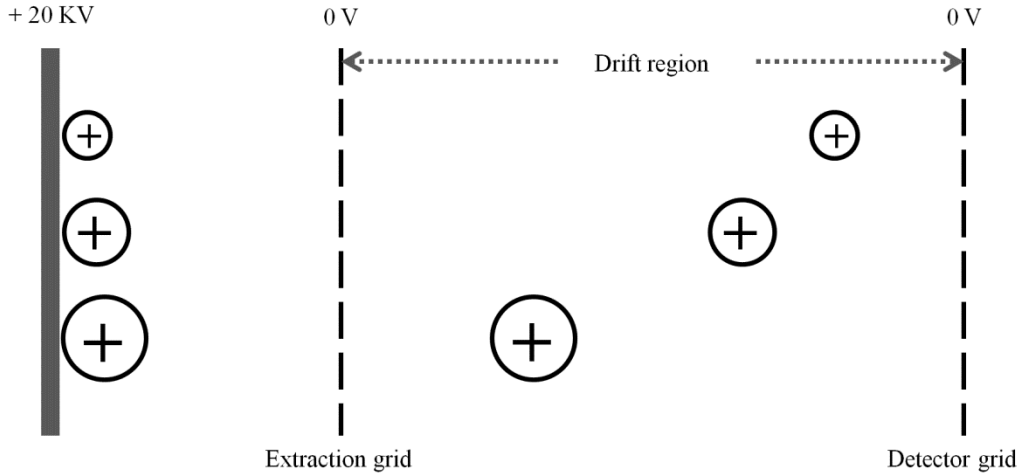


Figure 12.2 Separation of the ions in a TOF tube

If a particle with charge, q (= ez) and mass, m is accelerated towards the flight tube by electrostatic potential, V, the kinetic energy (KE) of the particle can be given by:

$$KE = qV \quad \cdots\cdots (12.5)$$

$$\frac{1}{2}mv^2 = zeV \quad \cdots\cdots (12.6)$$

where, v is the velocity of the particle.

If, L is the length of the tube, the time taken by an ion with velocity v is given by:

$$t = \frac{L}{v} \quad \cdots\cdots (12.7)$$

Substituting the value of v from equation 12.6 into equation 12.7 gives:

$$t = \sqrt{\frac{mL^2}{2zev}} \quad \cdots\cdots (12.8)$$

Equation 12.8 shows that the time taken by the particle to reach the detector is directly proportional to $\frac{m}{z}$ ratio. The $\frac{m}{z}$ ratio of the particle can therefore be calculated from the time of flight of the particle:

Rearranging equation 12.8:

$$\frac{m}{z} = \frac{2evt^2}{L^2} \quad \text{.....} \quad (12.9)$$

A serious problem with the linear TOF analyzers is their poor resolution. This happens due to difference in the flight times among the ions having same $\frac{m}{z}$ ratio. The factors responsible for the poor resolution include length of the ionization laser pulse, space distribution of the ions formed, and spread in the initial kinetic energies of the ions. In MALDI-TOF, for example, these factors severely affect the resolution. Two techniques have considerably improved the resolution in MALDI-TOF:

i. Delayed extraction: In continuous extraction, the ions generated are continuously extracted towards the TOF tube. An ion with higher initial kinetic energy reaches the detector earlier than the ion with smaller initial kinetic energy. Delayed extraction improves this situation substantially. Following ionization, the ions are allowed to move in the field free region according to their kinetic energies during a short delay. An extraction pulse is then applied; the pulse gives more energy to the ions that are nearer to the ionization source as compared to those that have moved away. The ions with small initial kinetic energies, therefore gain more energy and catch up the ions moving ahead (Figure 12.3).

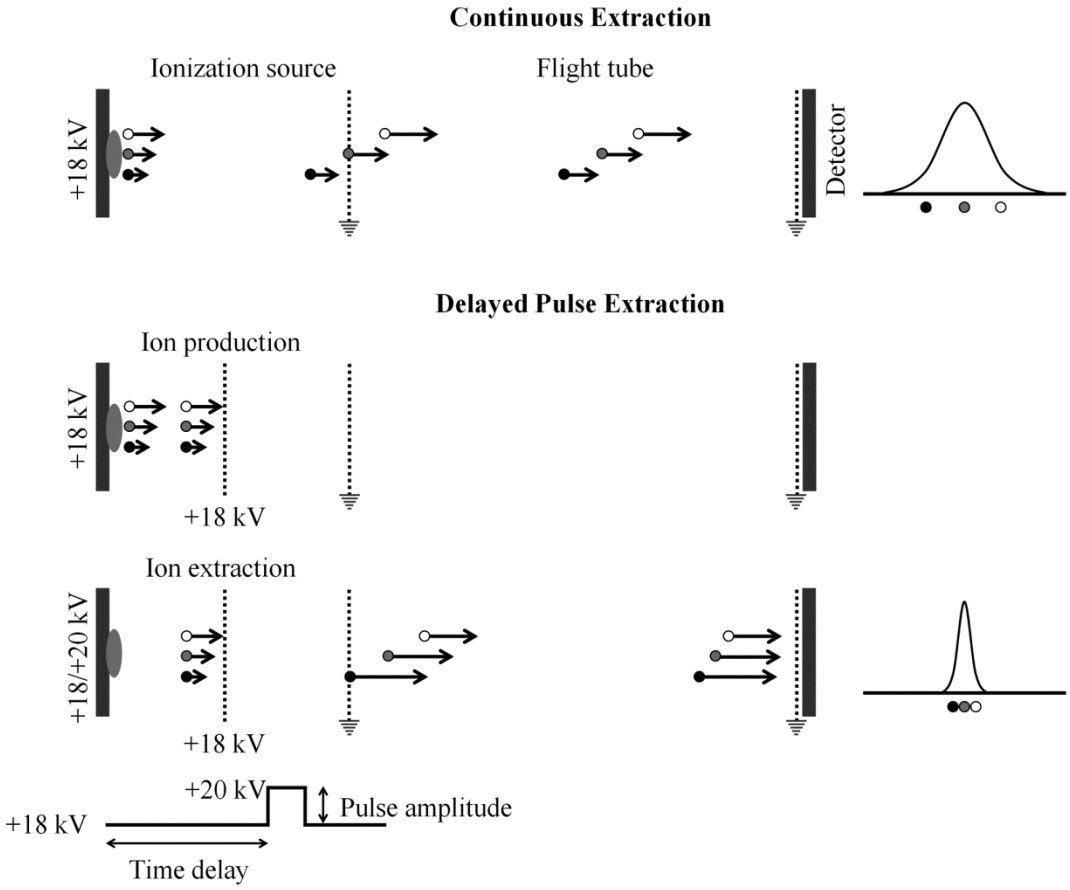


Figure 12.3 A diagrammatic representation of delayed extraction TOF

ii. Reflectrons: A simple reflectron is composed of a series of equally spaced grids. The reflectron is placed at the tube end opposite to the ion source. A potential is applied to the reflectron so as to reflect the incoming ions. A reflectron therefore acts as an ion mirror. The ions with higher kinetic energy will travel longer distance before reflecting back than those have smaller kinetic energy. The ions with higher kinetic energy are therefore made to travel longer distances thereby correcting for the spread in the peaks (Figure 12.4).

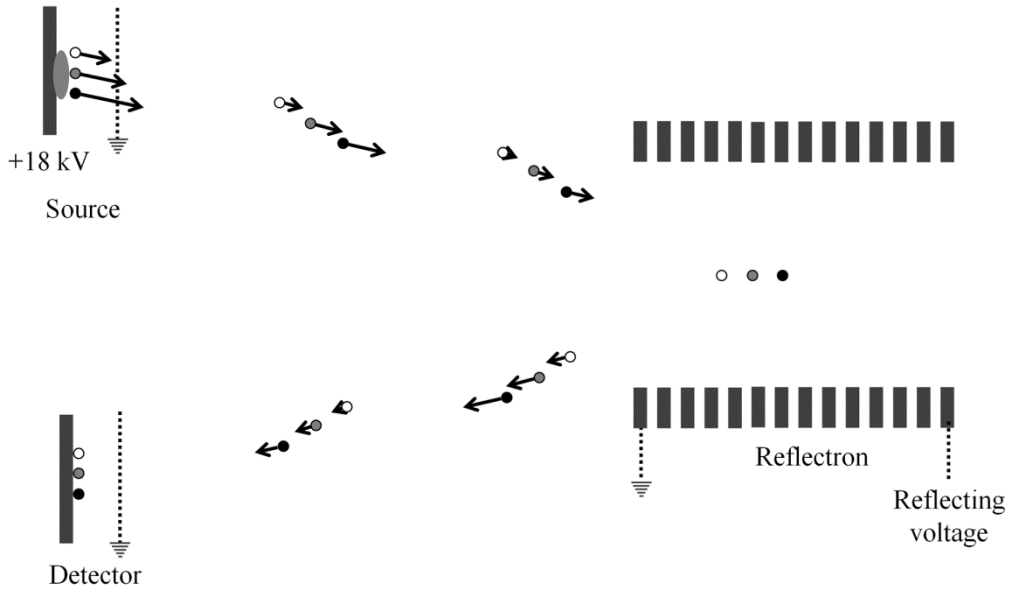


Figure 12.4 A diagrammatic representation of reflectron-mode TOF

Quadrupole analyzers

A quadrupole mass analyzer is made up of four rods arranged parallel to each other as shown in Figure 12.5. The principle of a quadrupole was proposed by Paul and Steinwegen in 1953 wherein hyperbolic cross-section of the rods was described as necessary. In practice, however, rods with circular cross section have also proved effective and have replaced the rods with hyperbolic cross-section in modern quadrupole detectors.

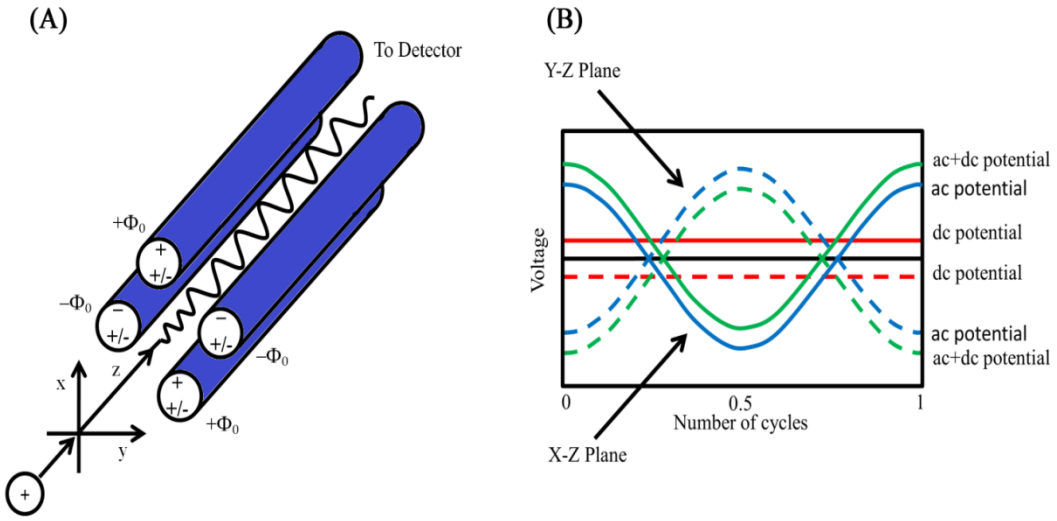


Figure 12.5 A quadrupole mass analyzer (A) and the potential on the rods as a function of time (B).

Consider the rods in x-z plane to be at a positive potential, U and the rods in y-z plane to be at a negative potential –U. An *a.c.* potential is subsequently applied to the rods such that the *a.c.* potential on the rods in x-z plane is 180° out of phase than that applied on the rods in the y-z plane. The potential on the rods in x-z and y-z plane can therefore be represented as

$$\Phi_{x-z} = U + V cos\omega t$$

$$\Phi_{y-z} = -U - V cos\omega t$$

The values of U range from 500 – 2000 volts while V ranges from 0 – 3000 V. The gas phase ions are accelerated and introduced into the quadrupole as a focused beam. To understand how the ions are separated inside a quadrupole, let us consider the rods in the x-z plane and those in the y-z planes separately. The rods in x-z plane have a positive *d.c.* potential and an *a.c.* potential that will periodically make the overall potential negative (Figure 12.5B). The ions will respond to the changes in the potential. If a cation is very heavy or the frequency of the *a.c.* potential is very high,

the cation remains largely unaffected and experiences only the average potential. This causes the ion to get focused towards the centre. If a cation is very light, it will readily respond to the changes in the potential and can accelerate towards the rods during negative potential and collide with them. Collision of the cation with the quadrupole rods during negative potential depends on the magnitude of the potential on the rods, frequency of the *a.c.* potential, mass of the cation, charge on the cation, and the position of the cation in the quadrupole. Let us now turn our attention towards the rods in the y-z plane. These rods have a negative average potential; the heavier cations will therefore get accelerated towards the rods and collide with them. Lighter cations will respond to the *a.c.* potential and get focused towards the centre. We can say that the x-z rods filter out the lower masses while y-z rods filter out the higher masses. The four rods together can therefore be used to allow the passage of a very small range of masses. A quadrupole therefore acts as a mass filter.

Ion trap

As the name suggests, ion trap mass analyzers trap the ions inside them. An ion trap can either be a 2D or a 3D ion trap. A 3D ion trap is basically a quadrupole in 3 dimensions. It has a circular electrode (also called a ring electrode) with two ellipsoid electrodes as its caps (Figure 12.6).

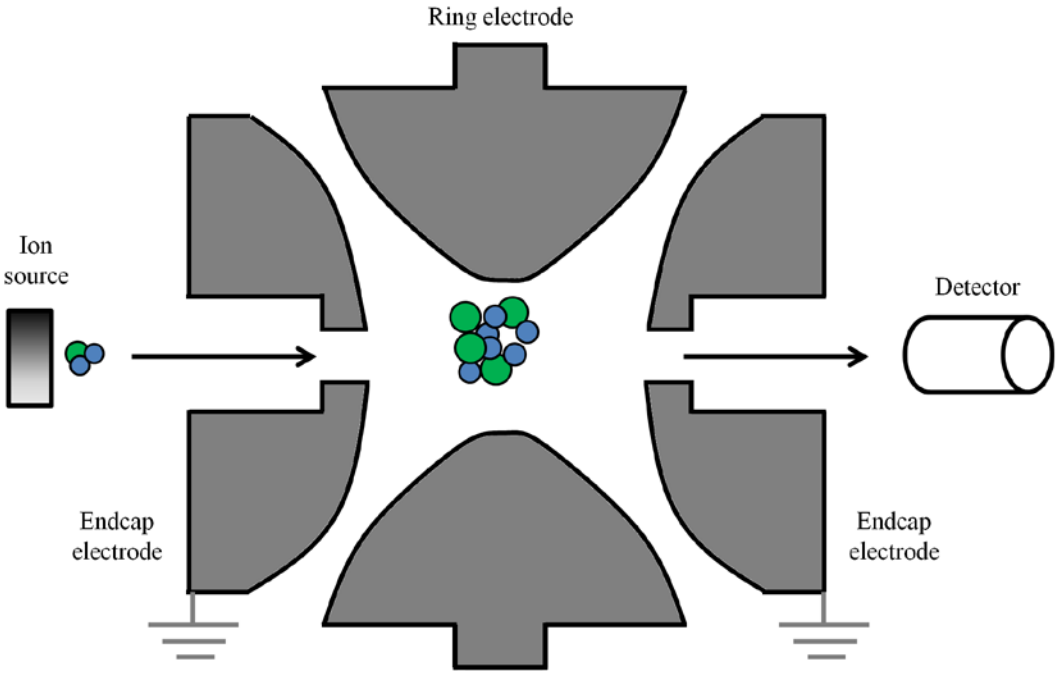


Figure 12.6 A diagrammatic representation of an ion trap

All the three electrodes have hyperbolic surfaces. Unlike linear quadrupole (quadrupole with four parallel rods) wherein electrostatic forces act in two axes, forces act in all the three axes in an ion trap. This implies that the stable trajectories of the ions cause them to be trapped. Ion traps provide higher sensitivity and are useful in tandem mass spectrometry; ions of desired mass can selectively be allowed to escape the trap by varying the ac potential; the escaped ions can then be analyzed by another mass analyzer attached in tandem.

Orbitrap

Orbitrap is an electrostatic ion trap. It has a barrel containing a spindle-shaped electrode at the centre. The spindle electrode is held at a constant negative voltage (-3200 V) for positive ion mode MS. Ions enter the orbitrap tangentially and get trapped by revolving around the spindle shaped electrode (Figure 12.7). The ions can be ejected out by applying the radiofrequencies of suitable frequency to the central electrode.

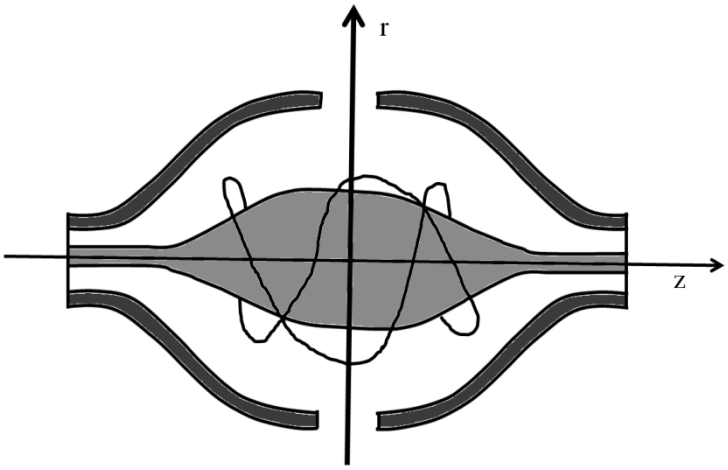


Figure 12.7 A schematic diagram of an orbitrap

Mass spectrometry coupled with chromatography

Mass spectrometers have been successfully coupled with the liquid and gas chromatographic methods. Chromatographic methods separate the compounds based on the differences in their physico-chemical properties. A complex mixture of compounds can be resolved into pure components using one or more chromatographic methods. Though excellent in separating the molecules, chromatographic methods do not allow identification of the unknown compounds in the mixture. The separated

components can be collected and identified using MS. It is also possible to couple the mass spectrometers with the chromatographic methods. Gas chromatography (GC) and liquid chromatography (LC) coupled with mass spectrometers have emerged as very powerful analytical tools. GC allows easy interfacing with the mass spectrometers; a gas chromatographic column can directly be coupled to the ionization source of the MS. Interfacing the LC with MS, however, is not as straightforward. We shall not be discussing the different types of interfaces of LC and MS. As it allows studying the large, polar, non-volatile, and thermolabile compounds, LC-MS is more widely used compared with GC-MS. Electrospray ionization is one of the softest ionization methods and probably the most widely used ionization method for LC-MS. LC-MS and GC-MS therefore allow determining the molecular masses of the eluants in real-time. The ions separated by MS can be fragmented in a collision cell; identification of these fragments by another mass analyzer allows identification of the compounds (tandem mass spectrometry). We shall see in the next lecture how fragmented ions help in the identification of the compounds.

Tandem mass spectrometry

Tandem mass spectrometry (also known as MS/MS) involves more than one mass analyzer. As we have just seen, incorporating more than one mass analyzer greatly enhances the capabilities of a mass spectrometer. Furthermore, it improves the sensitivity, mass resolution, and the mass accuracy of the spectrometer. The most common MS/MS experiment involves selecting an ion using first mass analyzer, which is then fragmented into daughter ions in a collision cell. The daughter ions are then detected by a second analyzer. It is possible to further fragment the daughter ions into granddaughter ions that are then analyzed by a third mass analyzer. It is in principle possible to do an MS^n experiment. We shall see in the next lecture how powerful tandem mass spectrometry is in identifying and characterizing the biomolecules.

Detectors

A variety of ion detectors presently exist, some of which are:

Electron multiplier: An electron multiplier is perhaps the most commonly used ion detector in mass spectrometers. It consists of a series of electrodes (dynodes). When an ion strikes the first dynode, it causes release of electrons from the dynode (the first dynode, therefore is a conversion dynode that converts the ion signal into electrons) that strike the second dynode releasing more electrons and so on. This cascading effect causes a large amplification in the electrical current (Figure 12.8A). Another design of the electron multiplier uses a continuous dynode (12.8B).

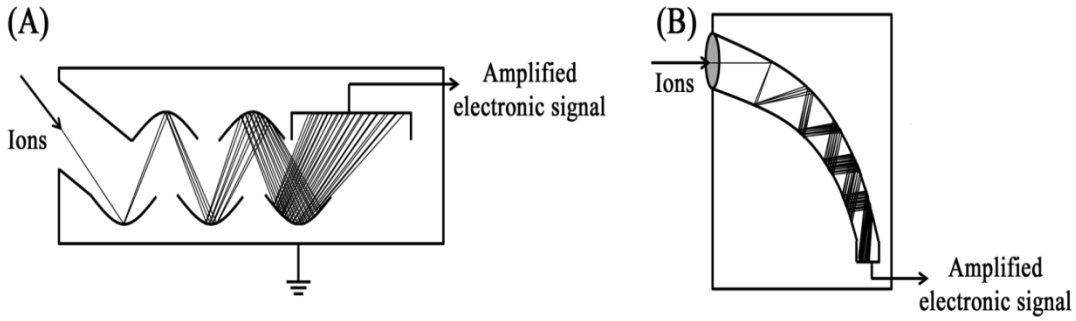


Figure 12.8 A diagrammatic representation of current amplification in an electron multiplier

Faraday cup: Faraday cup consists of a metallic cup that is connected to the earth through a resistor. An incoming ion strikes the cup and gets neutralized. This results in an electric current through the resistor that is proportional to the ion abundance.

Daly detector: A Daly detector is a type of electro-optical ion detector. The detector has one or more conversion dynodes that generate electrons in response to the ion strike. These electrons are accelerated towards a phosphor screen (scintillator) that generates photons in response to the electron strike. The photons, thus generated are detected by a photomultiplier tube (Figure 12.9).

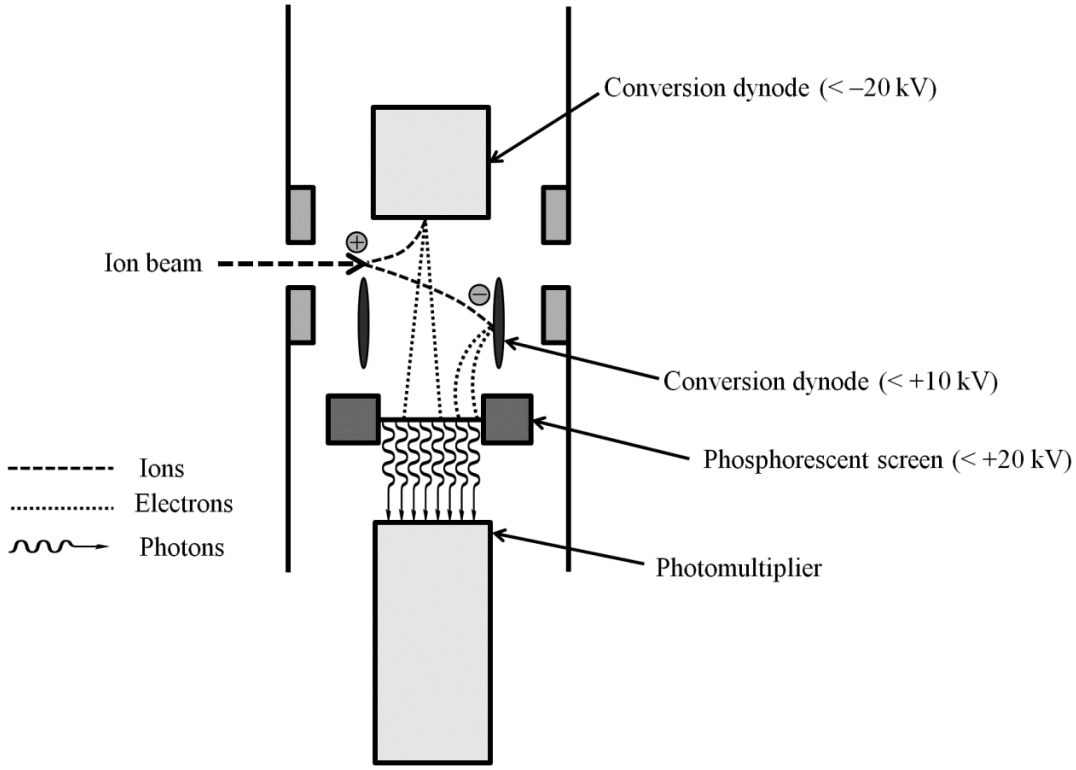


Figure 12.9 Schematic diagram showing working principle of a Daly detector

Focal-plane detectors: All the above-mentioned detectors fall in the category of point ion detectors *i.e.* the ions with different $\frac{m}{z}$ are resolved in time and detected at a single point. Focal-plane detectors detect the ions simultaneously; the ions of different $\frac{m}{z}$ are resolved in space, therefore strike at different points in array detectors (Figure 12.10).

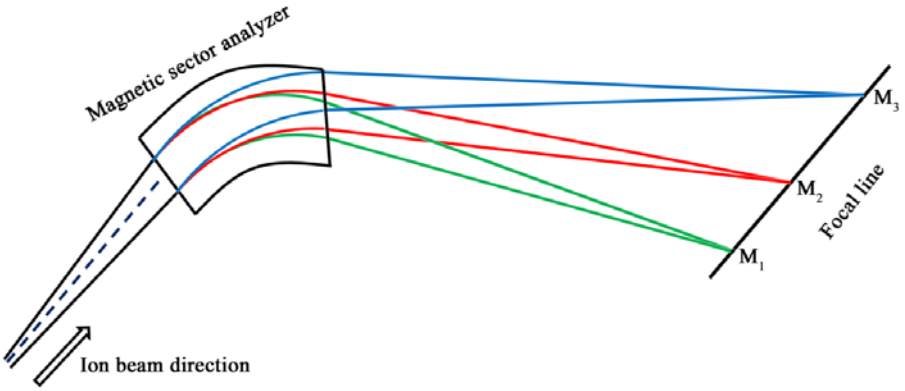


Figure 12.10 Schematic diagram showing a focal-plane detector for a magnetic sector mass analyzer

Lecture 13 Mass Spectrometry-III

We have already studied the various methods of ionization and mass analysis in lectures 10 and 11. This lecture discusses the properties of the mass spectra, their interpretation, and their applications, particularly in biomolecular analysis.

Characteristics of mass spectra

There are three basic characteristics of mass spectrometry: *exact mass*, *isotopic abundances*, and *fragmentation*.

Exact mass: When we talk of molecular weights in general chemistry, we typically refer either to the nominal mass or the average molecular mass. Nominal mass is calculated by adding the atomic masses of the predominant isotopes of all the elements rounded off to the nearest integer. Average molecular mass is the weighted average of the masses of all the isotopes (without rounding off). Stable isotopes of hydrogen, carbon, nitrogen, and oxygen as well as their relative abundances are listed in table 13.1.

Table 13.1 Natural abundance of the stable isotopes of selected elements

Isotope	Relative abundance	Exact mass	Nominal mass
^{1}H	99.985	1.0079	1
^{2}H (D)	0.015	2.0140	2
^{12}C	98.90	12.000	12
^{13}C	1.10	13.003	13
^{14}N	99.63	14.003	14
^{15}N	0.37	15.000	15
^{16}O	99.76	15.995	16
^{17}O	0.04	16.999	17
^{18}O	0.20	17.999	18

The nominal and average molecular masses of, say methane are 16 Da and 16.0428 Da, respectively. A mass spectrometer, however, detects the exact masses of the ions.

Isotopic abundances: Isotopic abundances are reflected in the high resolution mass spectra of the compounds and allow easy identification of small organic compounds. Let us take an example of methane. The different possible isotopologues of methane are listed in table 13.2 along with their natural abundances.

Table 13.2 Isotopologues of methane

Isotopologue	Relative abundances
$^{12}CH_4$	= 0.989 × 0.99985 × 0.99985 × 0.99985 × 0.99985 = 0.9884 = **98.84 %**
$^{12}CH_3D$	= 0.989 × 0.99985 × 0.99985 × 0.99985 × 0.00015 = 1.483×10^{-4} = **1.483×10^{-2} %**
$^{12}CH_2D_2$	= 0.989 × 0.99985 × 0.99985 × 0.00015 × 0.00015 = 2.247×10^{-8} = **2.247×10^{-6} %**
$^{12}CH_1D_3$	= 0.989 × 0.99985 × 0.00015 × 0.00015 × 0.00015 = 3.337×10^{-12} = **3.337×10^{-10} %**
$^{12}CD_4$	= 0.989 × 0.00015 × 0.00015 × 0.00015 × 0.00015 = 5.007×10^{-16} = **5.007×10^{-14} %**
$^{13}CH_4$	= 0.0110 × 0.99985 × 0.99985 × 0.99985 × 0.99985 = 0.011 = **1.1 %**
$^{13}CH_3D$	= 0.0110 × 0.99985 × 0.99985 × 0.99985 × 0.00015 = 1.649×10^{-6} = **1.649×10^{-4} %**
$^{13}CH_2D_2$	= 0.0110 × 0.99985 × 0.99985 × 0.00015 × 0.00015 = 2.474×10^{-10} = **2.474×10^{-8} %**
$^{13}CH_1D_3$	= 0.0110 × 0.99985 × 0.00015 × 0.00015 × 0.00015 = 3.712×10^{-14} = **3.712×10^{-12} %**
$^{13}CD_4$	= 0.0110 × 0.00015 × 0.00015 × 0.00015 × 0.00015 = 5.569×10^{-18} = **5.569×10^{-16} %**

$^{12}CH_4$ and $^{13}CH_4$ are the two predominant isotopologues of methane. Other isotopologues are too small in quantities to detect. A methane mass spectrum will therefore look like as shown in Figure 13.1

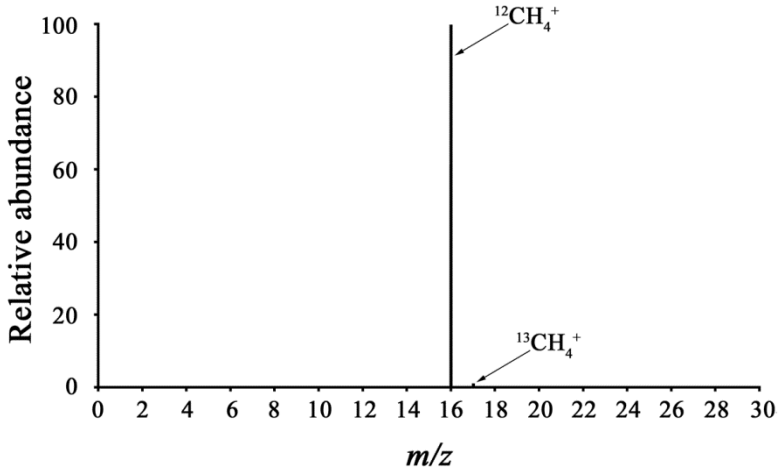
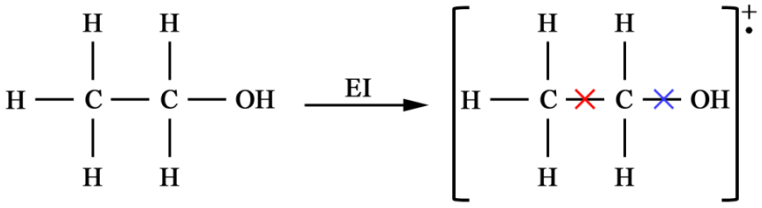


Figure 13.1 Electron ionization mass spectrum of methane showing the two predominant isotopologues

Fragmentation: We have already studied that electron ionization imparts large amount of energy to the cations that it generates. The radical cations thus generated undergo extensive fragmentation giving smaller cations. The fragments generated from molecular ions can provide important structural information about the molecules. Let us take ethanol as an example to see how this works:



The cross signs in the molecular radical cation represent the cleavage sites of the molecular ion. Cleavage between methyl and methylene carbons (red cross) can result in $[CH_3]^+$ or $[CH_2OH]^+$ ions while cleavage between methylene carbon and oxygen can result in $[C_2H_5]^+$ or $[OH]^+$ ions. The electron ionization mass spectrum for ethanol will therefore look like as shown in Figure 13.2. The idea behind identifying the structure of the molecules is to look at the differences between the peaks. A difference of 15 Da will be due to methyl loss, a difference of 17 is suggestive of the hydroxyl group, a difference of 29 can be due to the loss of an ethyl or aldehyde group. This information can be used to identify the molecules. In fact, there are softwares

available which can provide the possible molecular formulae of the compound when fed with the MS peaks.

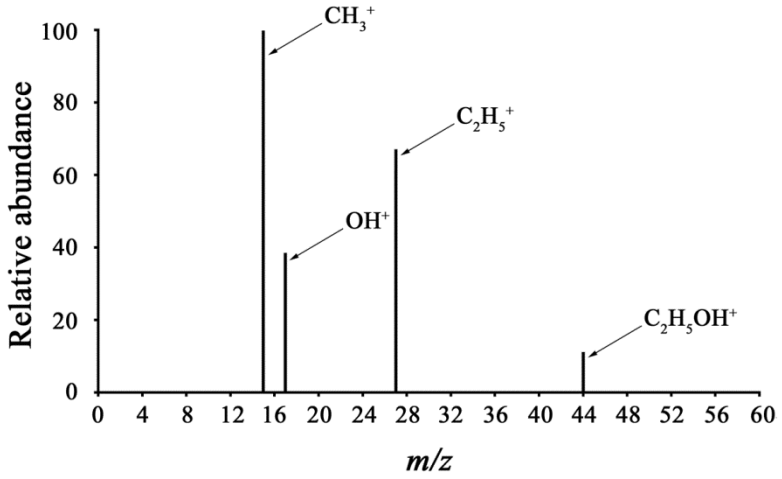


Figure 13.2 Electron ionization mass spectrum of ethanol

Analysis of biomolecules

Since the advent of MALDI and ESI ionization method, mass spectrometry has become a routine method for analyzing biomolecules. MS has been successfully utilized for obtaining a large amount of information about biomolecules including information that is difficult to obtain using other tools. Let us go through the various applications of MS in biomolecular analysis:

Molecular weight determination: Determination of molecular weight of a biomolecule is the most straightforward application of MS.

Structure verification and purity: Chemically synthesized molecules, such as peptides and oligonucleotides are often characterized by liquid chromatography and mass spectrometry. Suppose a chemically synthesized peptide, *DAKLRYFNQP* gives a MALDI mass spectrum as shown in Figure 13.3; the monoisotopic mass of the peptide is 1250.63 Da. The peak at 1180.59 is due to deletion of alanine.

THINK TANK??

In figure 13.3, where do you think the peaks at 1202.58 and 1273.62 m/z values arise from?

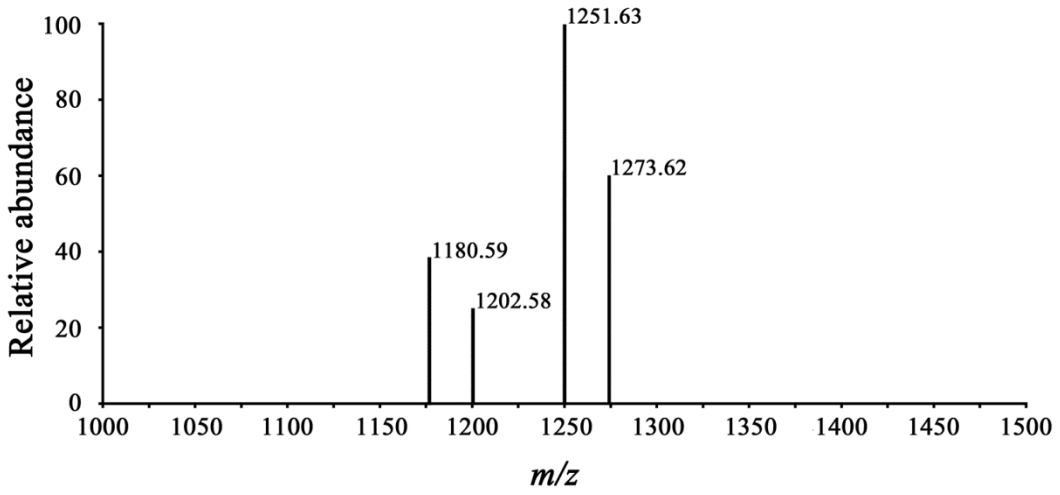


Figure 13.3 A hypothesized MALDI mass spectrum of the synthesized peptide, DAKLRYFNQP (theoretically calculated mass: 1250.63 Da)

Identification of chemical modifications: Biomolecules, especially proteins, can undergo a variety of chemical modifications such as phosphorylation, acetylation, methylation, fatty acylation, glycosylation, etc. These modifications are involved in biological processes like regulation of enzyme activity, signal transduction, gene expression, etc. It is therefore important to identify these species for understanding their function. Owing to its sensitivity and resolution, mass spectrometry has emerged as the method of choice for identification of small molecule modifications in biomolecules.

Protein sequencing:

Proteins are usually ionized using soft ionization techniques, MALDI and ESI. These methods yield quasimolecular ions that allow identification of proteins in complex mixtures. For determining their sequences, however, proteins need to be fragmented. The idea behind protein sequencing using MS is very straight forward and is summarized in Figure 13.4. Briefly, a protein quasimolecular ion is selected using a mass analyzer. The selected ion is then fragmented, typically in a collision cell (collision induced dissociation). Collision induced dissociation results in a large number of fragments that also have overlapping amino acid sequences. These daughter ions are then detected by a second mass analyzer. Mass of a fragment comprises the information for its amino acid composition. Masses of the overlapping fragments allow sequencing of the complete molecule.

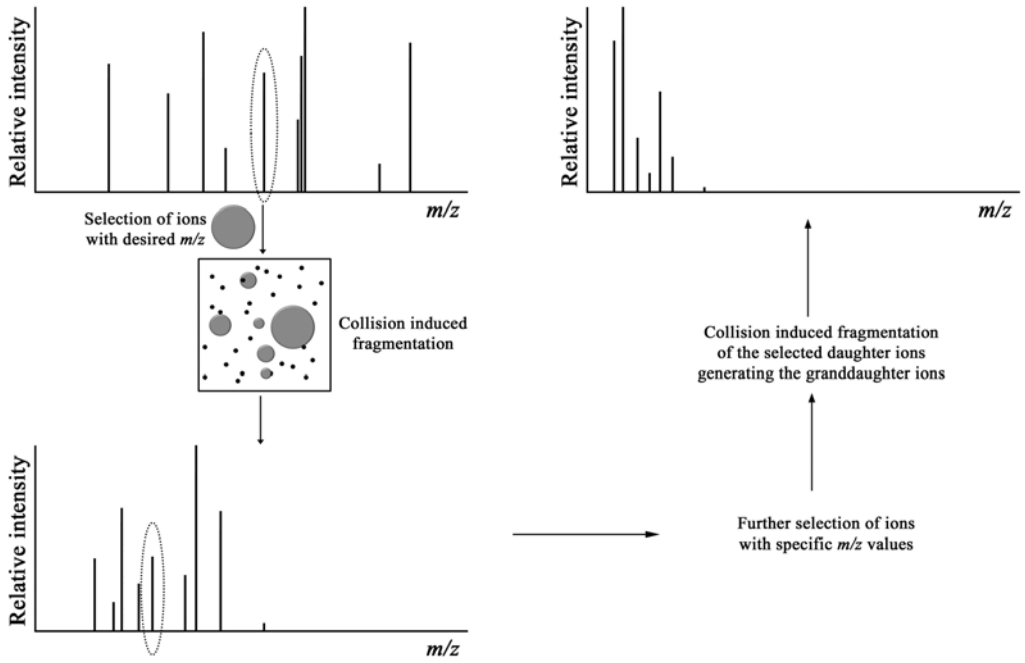
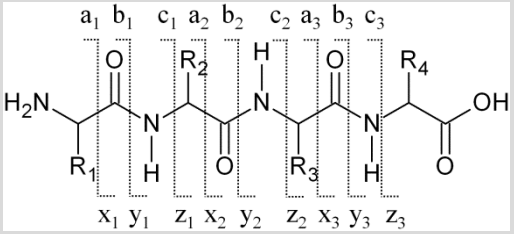


Figure 13.4 Principle of protein sequencing using tandem MS

Box 13.1: Peptide fragmentation

Fragmentation of peptides in a collision cell occurs in a well-defined manner; the fragmentation largely occurs along the peptide backbone with some side-chain fragmentation. The cleavage along the peptide backbone can occur at NH–C_α, C_α-CO, and CO–NH bonds. Each cleavage produces two species, one of which is charged and the other one neutral. As the charge can lie on any of the species produced, six different types of ions are generated. The nomenclature of these ions is shown in the figure below:



CO-NH is the most common cleave site. The difference in the masses of the adjacent b ions or y ions enables identification of the terminal amino acids thereby providing sequence information. Side-chain fragmentation is also useful as it provides information about side-chain modifications, if any.

Protein identification: Identification of a protein classically requires complete or partial protein sequence. A partial sequence can allow protein identification by comparing it with the sequences of the proteins available in protein sequence databases. A protein can therefore be identified by doing sequencing using MS. However, it may not be required to determine the sequence of a protein for its identification. A typical scheme for protein identification is shown in Figure 13.5.

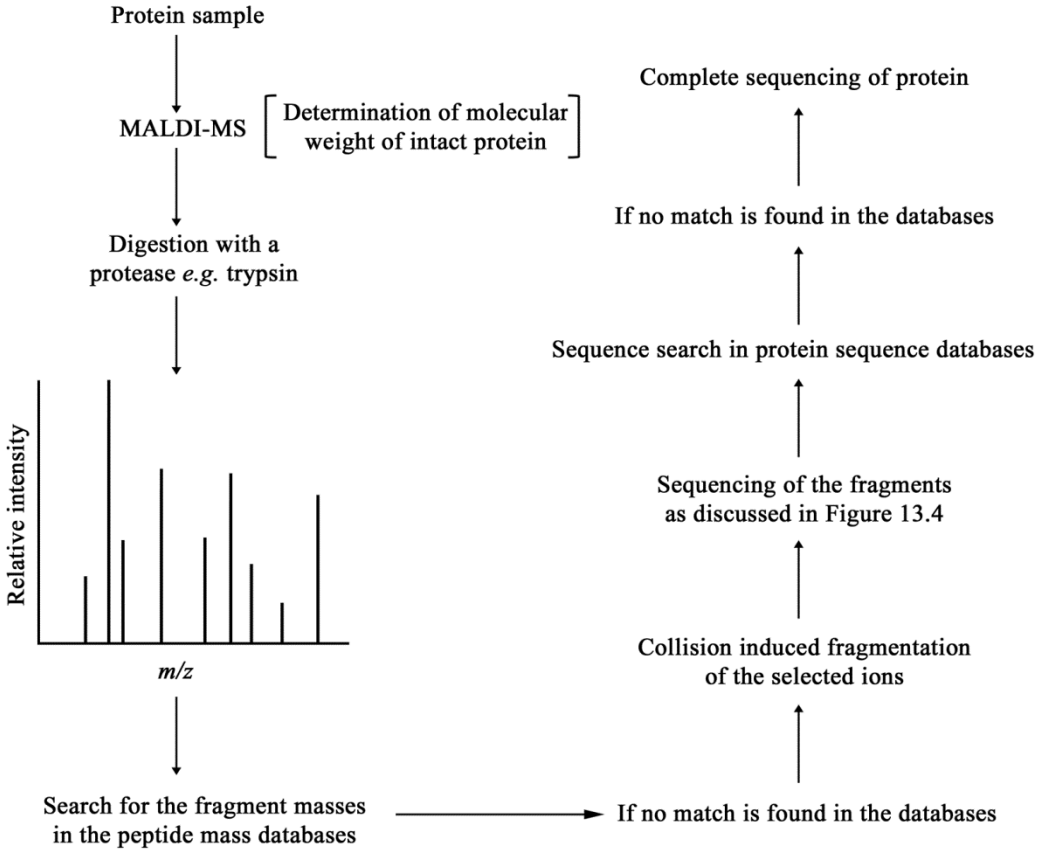


Figure 13.5 Protein identification using mass spectrometry

A protein is cleaved into smaller peptide fragments using a sequence specific enzyme, usually trypsin or chymotrypsin. Trypsin cleaves at the C-terminal side of lysine and arginine residues while chymotrypsin cleaves at the C-terminal side of aromatic amino acids, phenylalanine, tyrosine, and tryptophan. The masses of these fragments are then searched in peptide mass databases. This method is termed as the *peptide mass fingerprinting*. If no match is found in the databases, the peptides are sequenced using another MS as discussed earlier; the protein is identified by searching the

sequences of the fragments in the protein sequences databases. Protein identification using MS is central to the proteomic studies. A typical proteomic analysis is briefly summarized in Figure 13.6

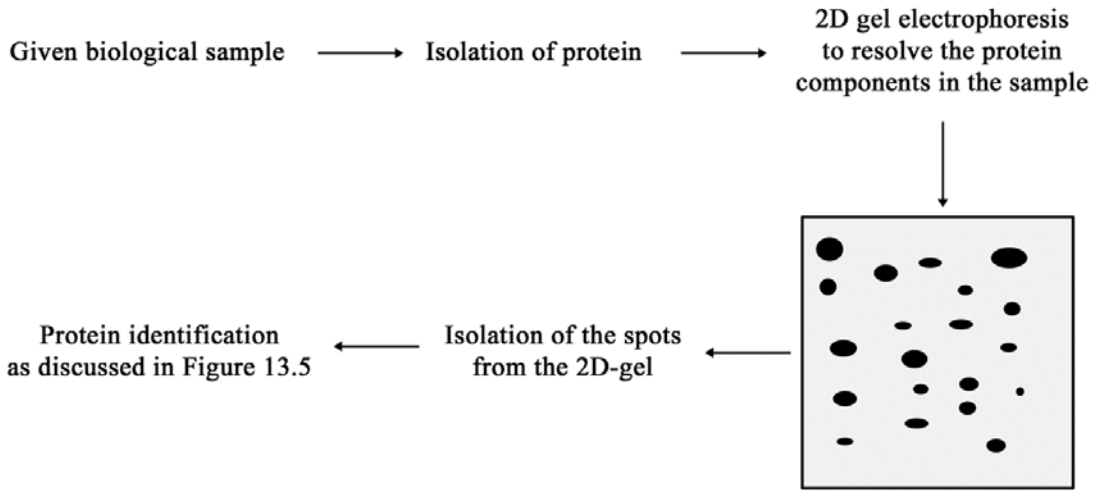


Figure 13.6 Outline of a proteomics experiment

Non-covalent protein complexes: Non-covalent interactions between the molecules are usually disrupted during ionization. ESI, however, has proved to be sufficiently soft for allowing detection of the protein complexes. It is, however, important to note that the complexes of biomolecules detected in gas phase may not represent the true solution complexes. The structures of the molecules in solution and the interactions they are involved in are determined by the complex interplay of hydrogen bonding, hydrophobic interactions, van der Waals forces, and electrostatic interactions. In gas phase under high vacuum, the electrostatic interactions are usually the most predominant ones. One therefore has to be careful while analyzing the data for the non-covalent complexes.

Three dimensional structures of proteins and peptides: Mass spectrometry can provide information about the three dimensional structure of a protein. Consider a protein that is being analyzed using ESI-MS. An unfolded protein usually shows a broader charge distribution with higher amount of charge due to larger solvent-accessible area and the exposed ionization sites. Another approach includes structural analysis of the proteins after subjecting them to the conformation sensitive reactions. Hydrogen-deuterium exchange (H/D exchange) is the most widely used such reaction.

The exposed regions on a protein will readily exchange their amide hydrogens with deuterium. The amides that are involved in backbone H-bonding (*i.e.* those involved in secondary structures) exchange their hydrogens very slowly. The folded regions in the protein can then be identified by analyzing the peptide fragments using tandem mass spectrometry. These approaches also allow study of the folding processes of proteins.