Application of Optical Spectroscopy in Biological Systems

Basic optical spectroscopic tools:

Conceptually, a typical spectroscopic experiment is extremely simple. Electromagnetic radiation at a certain nominal wavelength λ is allowed to impinge on the sample. Then some properties of the radiation that emerges from the sample are measured. One of the simplest properties is the fraction of the incident radiation absorbed or dissipated by the sample. (Typical techniques are optical absorption spectroscopy, some modes of NMR spectrometry and various elastic scattering techniques.) Instead, one can examine the radiation emitted by the samples at wavelength other than that used for excitation. (Fluorescence, Phosphorescence, Raman scattering and inelastic light scattering are examples.) Not only the emergent intensity, but also the distribution of emergent frequencies, is sources of information. In more complex techniques, not just intensity is detected, but also the kind and degree of polarization of the radiation emitted by a sample. (ORD, CD and fluorescence polarization fit into this category.)

Qualitative description of spectroscopy:

There is no simple way to explain the interaction of light with matter. Light is rapidly oscillating electromagnetic field. Molecules contain distribution of charges and spins that have electrical and magnetic properties. These distributions are altered when a molecular is exposed to light. In a typical spectroscopic experiment, light is sent through the sample, either continuously or in a pulse. What one must deal with is the rate at which the molecule responds to this perturbation. One must explain why only certain wavelengths cause changes in the state of the molecule. One must calculate how the presence of the molecule alters the radiation that emerges form the sample.

Absorption spectroscopy of electronic states:

The measurement most frequently performed on biopolymers is the absorption of visible or ultra violet light. This technique is used for purposes ranging from simple concentration determinations to resolution of complex structural questions. In this section, we consider first some of the basis features of these measurements, and then particular aspects of absorption relevant to the properties of large molecules.

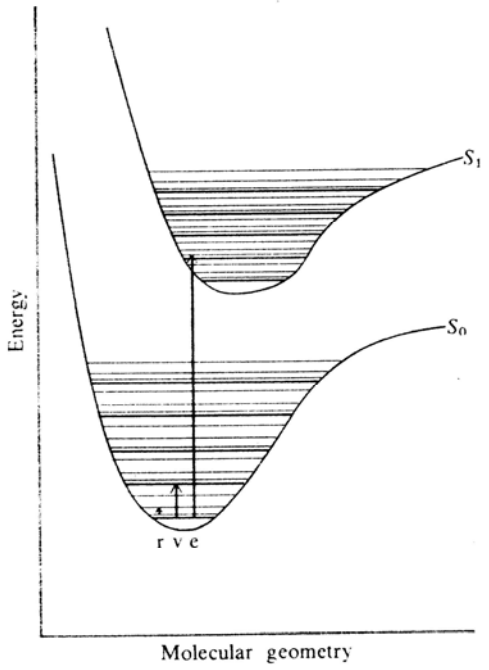


Figure 1 *Energy levels of a small molecule. Selected rotational sublevels of the vibrational levels of each of two electronic states are shown. Transitions corresponding to electronic (e), vibrational (v) and rotational (r) spectra are indicated.*

Energy states of molecules:

Figure 1 shows a section through the potential energy surfaces of the two lowest electronic states of a typical simple molecule. Superimposed on each of these sates is a series of vibrational levels that, in turn, are subdivided into a myriad of rotational levels. The energy spacing between the lowest rotation-vibration states of the two electronic states of the two electronic states S_0 and S_1 typically is 80 kcal mole^{-1}. This energy is much greater than the thermal energies at room temperature. Therefore one knows form the statistical mechanics that, for all practical purposes, in the absence of radiation that can excite a transition, all molecules in a solution are in the lowest electronic state, S_0. The energy spacing between vibrational levels is of the order of 10 kcal mole^{-1}. This energy is also larger than thermal energies so, at least approximately, we can consider only the lowest vibrational level of S_0 to be appreciably populated. However, rotational

energy spacings are only 1 kcal $mole^{-1}$ or less; therefore many rotational levels are populated.

When light of the correct frequency is absorbed, the molecule can be excited to one of many rotation-vibration levels of the electronic state S_1. The energy and corresponding wavelength of the incident light that will be absorbed by a certain transition is determined by the energy difference between the ground and excited states. Yet, when light passes through a sample containing absorbing species it is only a fraction of the light of correct energy that is actually absorbed by the sample molecules. If the given species is uniform, the fraction of light absorbed by the sample is defined by the Beer-Lambart law:

$$A(\lambda) = \log(\frac{I}{I_0}) = \varepsilon(\lambda).C.l$$

Where, $A(\lambda)$ is the absorbance or optical density, I_0 is the initial intensity of light impinging on a sample in cuvette of path length l cm, and I is the final intensity. In practice, the absorption spectrum of macromolecules is the difference in absorbance between the macromolecular solutions against a solvent blank which is employed in a double beam spectrophotometer. For most accurate measurement of absorbance, the value of $A(\lambda)$ usually obtained is in the range 0.1 to 2.

Fluorescence Spectroscopy:

Fluorescence is the radiation in the UV-visible region emitted by a molecule in going from an excited singlet state to the ground state. Compared to the ordinary light absorption, fluorescence process takes place in a slower time scale (~10^{-9} to 10^{-8} sec). During the time a molecule remains in the excited electronic state, it may be subjected to a wide variety of interactions and perturbations, e.g., proton transfer reactions, conformational changes, solvent relaxation etc that might significantly affect the fluorescence spectral characteristics of the molecule and yield useful information. It is this favorable time scale, in conjunction with the intrinsic sensitivity of the technique that makes fluorescence methods generally attractive for investigations of fluorophores and their interaction with other biological macromolecules.

There are several non-radiative processes through which the energy can be lost upon depopulation of the excited state; all the pathways are competing directly with fluorescence. These processes can be schematically described in Jablonski diagram. The deactivation pathways have been schematically shown in figure 2.

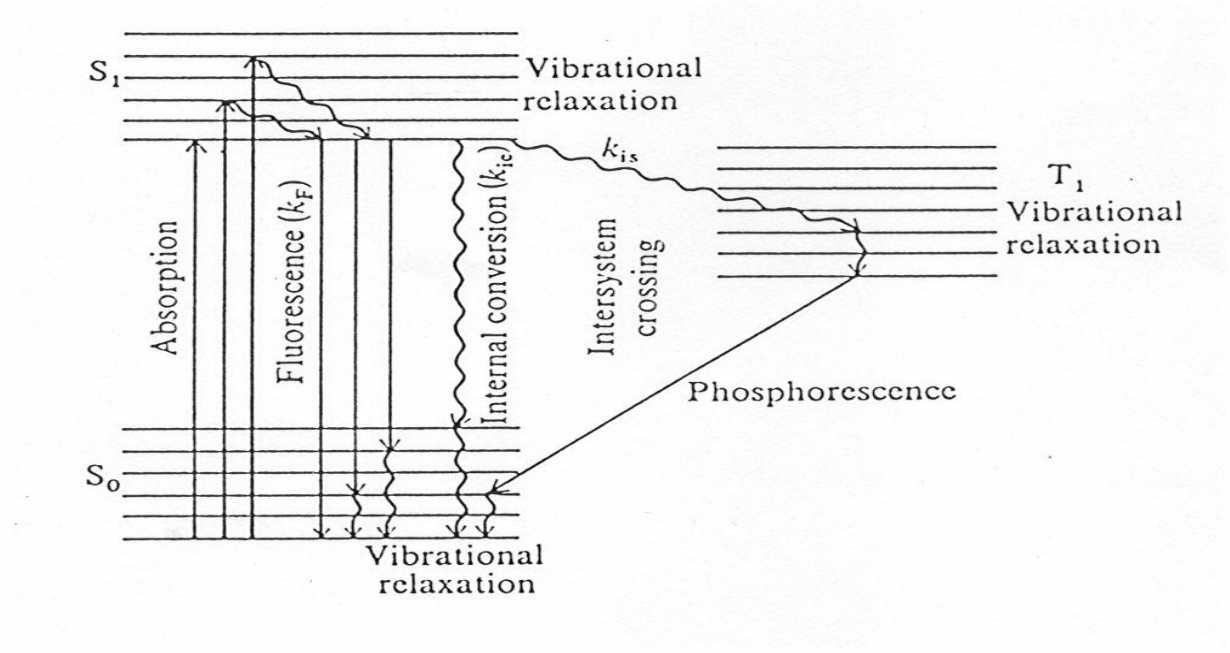


Figure 2 *Pathways for production and deactivation of an excited state*

Here S_0, S_1 are the ground and first excited singlet states respectively. Following light absorption (lifetime $\tau \sim 10^{-15}$ s), a fluorophore is usually excited to some higher vibrational level of S_1. With a few rare exceptions, molecules in condensed phase rapidly relax to the lowest vibrational level of S_1 by a process called internal conversion occurring at a rate k_{ic} . In this process, excitation energy in S_1 is lost by collision with solvent or by dissipation through internal vibrational modes. This is a much faster process ($\tau_{ic} \sim 10^{-12}$ s) than fluorescence hence fluorescence emission generally results from the lowest energy vibrational state of S_1. This is known as "Kasha's Rule". Another process that affects the fluorescence intensity is, intersystem crossing. The process occurs at a rate of k_{is}. In this process, the nominally forbidden spin exchange converts an excited singlet into an excited triplet state (T_1). This state can, in turn convert to the ground singlet state (S_0)

either by phosphorescence (emission of a photon) or by internal conversion. The triplet state generally is lower in energy than the excited singlet and with a different lifetime is of the order of 10^{-3} to 10 sec. Hence, phosphorescence occurs at longer wavelengths and can easily be resolved from fluorescence both by steady state and time resolved measurements.

Fluorescence polarization and anisotropy:

Anisotropy measurements are commonly used in the biochemical application of fluorescence. Anisotropy measurements provide information on the size and shape of a macromolecule and also the rigidity of various molecular environments. Such measurements are based on photoselective excitation of fluorophores by polarized light. Fluorophores preferentially absorb photons whose electric vectors are aligned parallel to the transition moment of the fluorophore. If plane polarized light is used, it will preferentially excite those fluorophores whose molecular axes are oriented in a particular direction with respect to the plane of polarization (photo selection). If the fluorophore remains immobile during its excited state lifetime, then the fluorescent light will be highly polarized. The excitation with polarized light resulting in a population of excited fluorophores that is symmetrically distributed around the z-axis therefore has a very high value of anisotropy (r_0).

If it rotates during its fluorescence lifetime, then the resulting fluorescence will be less polarized or depolarized. The polarization of fluorescence thus acts as a convenient index of the extent of molecular rotation during its excited lifetime. The fluorescence anisotropy (r) and polarization (P) are defined by

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

$$P = \frac{I_{II} - I_{\perp}}{I_{II} + I_{\perp}}$$

Where I_{II} and $I_{\perp}$ are the intensities of the vertically (II) and horizontally ($\perp$) polarized emission, when the sample is illuminated with vertically polarized light.

The anisotropy of the randomly distributed fluorophores, with a collinear absorption and emission dipoles is reduced by a factor of 2/5 due to photo-selection. Generally, the absorption and emission dipoles of a fluorophore are oriented at an angle (α) within the plane of fluorophore. In absence of any depolarizing process, such as rotational diffusion or energy transfer, the observed anisotropy (r_0) of a fluorophore is a product of the factor 2/5 and the loss of anisotropy due to the angular displacement of dipoles. Hence the expression for r_0 is given by

$$r_0 = \frac{2}{5} \times \frac{(3\cos^2 \alpha - 1)}{2}$$

For some molecules, α is close to zero. An anisotropy of 0.39 corresponds to an angle 7.4° between the dipoles, whereas r_0 =0.4 corresponds to an angle of 0°. It is to be noted that the fundamental anisotropy value is zero when α= 54.7°. When α exceeds 54.7°, the anisotropy becomes negative. The maximum negative value (-0.20) is found for α = 90°. Hence, for an isotropic solution with single photon excitation, r_0 lies between $-0.2 \leq A_0 \leq 0.4$. Since the orientation of the absorption dipole differs for each absorption band, the angle α as well as r_0 varies with excitation wavelength. A polarization spectrum is a plot of polarization or anisotropy of the fluorescence versus the excitation wavelength. Generally, the anisotropy is independent of emission wavelength since emission is almost always from the lowest singlet state. However, in case of solvent relaxation during the lifetime of a fluorophore the lowest singlet state relaxes to lower energies resulting in a dependence of anisotropy on emission wavelength. Rotational diffusion of fluorophores is a dominant cause of fluorescence depolarization. The behavior of depolarization is described by the well-known Perrin equation relating the anisotropy of a rotating fluorophore to that of a motionless one.

$$\frac{1}{r} = \frac{1}{r_0}\left(\frac{1+\tau_f}{\tau_c}\right) = \frac{1}{r_0}\left(\frac{1+\tau_f k_B T}{V_h \eta}\right)$$

Where k_B is the Boltzman constant and τ_c is the rotational correlation time of the fluorophore which is governed by the viscosity (η) and temperature (T) of the solution and by the volume of the rotating unit (V_h) as described by the above equation. From this equation it is easily seen that any factor, which affects the size, shape or flexibility of a macromolecule will also affect the anisotropies. These properties of macromolecules can be affected by pH, temperature, viscosity, denaturants and by association reactions. Thus, anisotropy measurement is a useful tool for monitoring molecular motions and microviscosity around a probe.

The most common optical arrangement for the measurement of fluorescence polarization is the 'L' format of the spectrofluorometer. The L-format measurement of fluorescence anisotropy is schematically represented in figure 3.

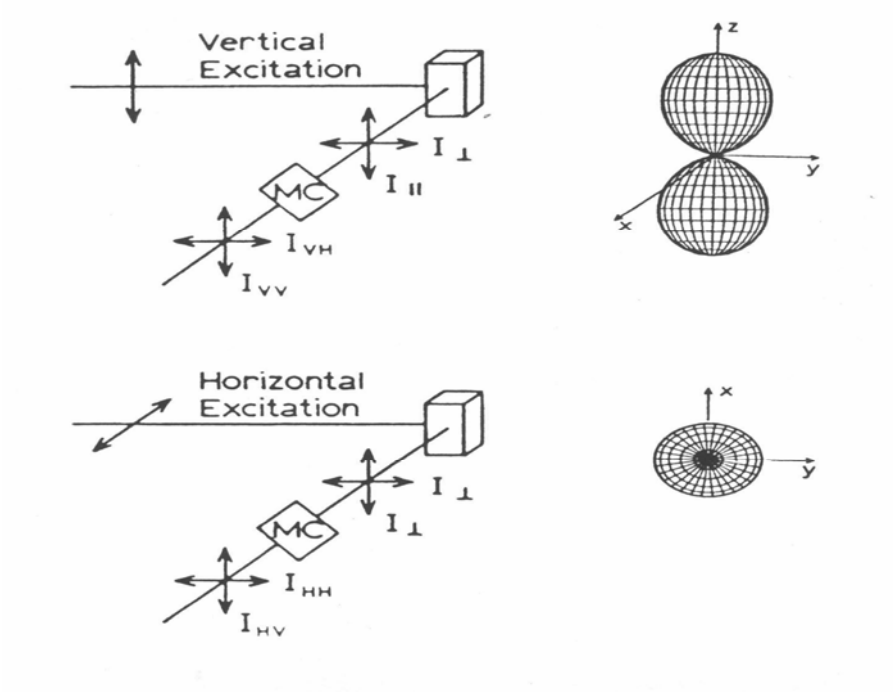


Figure 3 *Schematic diagram for L-format measurements of fluorescence anisotropy. (MC is the monochromator). The shapes at the right are the excited state distributions.*

The sample is excited with the vertically polarized light and the fluorescence intensities are recorded with the analyzing polarizer oriented parallel (I_{VV}) and perpendicular (I_{VH}) to the excitation polarizer and calculated according to equations. One problem

encountered here is that the emission monochromator passes polarized fluorescent light of different orientations with different efficiencies. This means that a correlation factor (G-factor), which is the ratio of the sensitivities of the detection system for vertically and horizontally polarized light, has to be introduced into equations to account for this. It can be easily measured by setting the excitation polarizer to the horizontal orientation and recording the fluorescence in both parallel (I_{HH}) and perpendicular (I_{HV}) orientations. When this is done, both the horizontally and vertically polarized components are proportional to $I_{\perp}$. Therefore,

$$G = \frac{I_{HV}}{I_{HH}}$$

In practice, the correction is incorporated by multiplying $I_{\perp}$ in equations by G, leading to the new formulation

$$r = \frac{I_{VV} - GI_{VH}}{I_{VV} + 2GI_{VH}}$$

$$P = \frac{I_{VV} - GI_{VH}}{I_{VV} + GI_{VH}}$$

The dependence of fluorescence anisotropy upon rotational motion and microviscosity of the environment have resulted in numerous applications of fluorescence spectroscopy in biochemical research like quantification of protein denaturation, protein association with other macromolecules and the internal dynamics of a macromolecule. Moreover, the encapsulation of a fluorophore into a host molecule can also be monitored by using fluorescence polarization and anisotropy.

Fluorescence lifetime measurement

Fluorescence lifetimes were measured using a time-correlated-single-photon counting (TCSPC) spectrophotometer (FLA-900, Edinburgh) with FWHM ~ 1.2 ns, repetition rate 25 kHz, electrode separation 0.25 mm and nitrogen as filler gas.

The single photon counting measurement relies on the concept that the probability distribution for emission of a single photon after an excitation event yields the actual

intensity against time distribution of all the photons emitted as a result of the excitation. The relative timing information of two electronic pulses can be obtained down to a limit of ± 50 ps and it is this advantage of high timing accuracy that makes the TCSPC a unique and widely used technique. Figure 4 shows a block diagram of the TCSPC set-up.

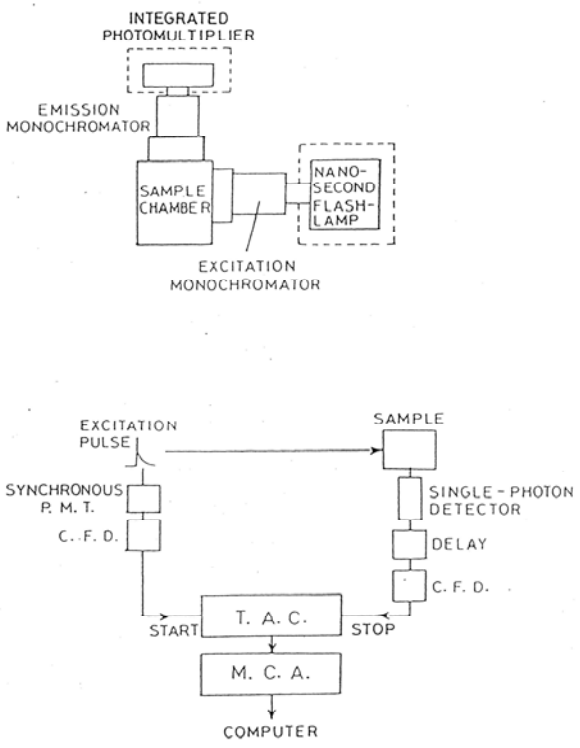


Figure 4 *Block diagram of the TCSPC set-up*

In this technique, a timing pulse (START) is generated synchronous with the excitation pulse. This START pulse switches on the charging of a capacitor in the Time-to-Amplitude-Converter (TAC). The first emitted photon from the sample generates another timing pulse (STOP), which snaps off the switch. The TAC thus maps the time-delay information into a voltage level. The statistical nature of the technique requires this procedure to be repeated till a large number of photons have been detected ~ 10^5 or more. This demands that all instrument parameters should remain constant during the several

hours required in a typical experiment to collect these many counts and one way to ensure such long-term stability, as far as the lamp output is concerned, is to clean the electrodes after every 72 hours of operation. The TAC output pulses are stored and displayed in a histogram in the Multi-Channel analyzer (MCA). For statistical validity of the data and to prevent excess weightage of the early arriving pulses, the START to STOP ratio is limited to 100:1. The decay curve is then obtained over a time zone of five lifetimes. The use of a constant fraction discriminator (CFD) provides improved time resolution by removing the variability due to the amplitude of each pulse. Moreover in a nanosecond flash-lamp system, the radio frequency noise generated in the thyratron and the spark lamp contributes to data distortion. This is normally manifested as a high frequency ripple structure on the decay data in MCA and can be reduced by ensuring a good system ground for the optical bench components and the electronics.

In practice, the decay data on the MCA is a convolution of the excitation pulse, the actual decay and the system response due to temporal effects of the excitation pulse and instrument distortion. So during experiment first a scatterer (Ludox of appropriate concentrations was used in our case) is placed in the sample chamber and the pump function (P(t)) is obtained.

$$P(t) = \int_0^t \left[\int_0^x L(z)H(x-z)dZ \right] E(t-x)dx$$

So, $P(t) = L(t) \otimes H(t) \otimes E(t)$

Where, L(t) is the excitation pulse, H(t) is the photomultiplier response and E(t) is the total electronic processing response. Next, the sample is placed in the sample chamber and the observed decay data I(t) is given by $I(t) = L(t) \otimes G(t) \otimes H(t) \otimes E(t)$

where G(t) is the sample decay. So $I(t) = P(t) \otimes G(t)$ provided that the distortions associated with H(t) and E(t) are linearly propagated and the pump function is identical for both sample decay and lamp profiles. The I(t) and P(t) information stored in the MCA are then transferred to the computer for deconvolution. In our experiments we have used the FL900 analysis program supplied by Edinburgh which uses non-linear least square data fitting based on the Marquardt algorithm. The goodness-of-fit is measured by parameters like reduced chi-square, weighted residuals, autocorrelation function and the Durbin Watson parameter.

Circular Dichroism:

In general, molecules absorb light when they undergo a transition from one state to a higher energy state. Here we are concerned with electronic absorption, that is, the spectra that can be measured for a molecule undergoing a transition form its ground state to some higher energy electronic state. Most biological molecules contain a number of electronic units that absorb light nearly independently called chromophores, which are asymmetrically disposed in space. Such asymmetric molecules will absorb left circularly polarized light differently from right circularly polarized light. Circular Dichroism (CD) is the difference in the absorption between left and right circularly polarized light, so that CD-spectral bands occur wherever there are normal electronic absorption bands in an asymmetric molecules.

Normal absorption spectroscopy obeys Beer's Law:

$A(\lambda) = \varepsilon(\lambda)l.c$

Where A is the measured absorption that is unitless and varies with wavelength λ, l is the path length of the sample cell in centimeters, c is the concentration of the sample in moles/lit, and ε is the characteristic of the molecule in litres/mol/cm called the extinction coefficient. A graph of ε vs.λ gives bands of a finite width and the integrated intensity under these bands is measured by the dipole strength,

$$D = \frac{3\times10^3 hc\ln 10}{8\pi^3 N_0}\int\frac{\varepsilon(\lambda)}{\lambda}d\lambda$$

where h is Planck's constant, c is the speed of light and N_0 is Avogadro's number.

In Cd spectroscopy, both left (L) and right (R) circularly polarized light obey Beer's Law, so that the difference obeys the following equation;

$A_L(\lambda) - A_R(\lambda) = \Delta A(\lambda) = [\varepsilon_L(\lambda) - \varepsilon_R(\lambda)]lc = \Delta\varepsilon(\lambda)lc$

The integrated intensity under CD bands is measured by the rotational strength.

$$R = \frac{3\times10^3 hc\ln 10}{32\pi^3 N_0}\int\frac{\Delta\varepsilon(\lambda)}{\lambda}d\lambda$$

Theoretically the rotational strength can be calculated quantum mechanically from interactions among the transitions in the various chromophores. However, as yet such calculations are not sufficiently accurate for interpretation of CD spectra. Thus CD spectra are usually interpreted empirically.

Membrane

Biological membranes are continuous structures separating two aqueous phases. They are relatively impermeable to water-soluble compounds, show a characteristic trilamelar appearance when fixed sections are examined by electron microscopy and contain significant amount of lipids and proteins. The fact that membranes contain large amounts of lipids, particularly phospholipids, distinguishes them from most other cellular structures and it is not surprising that until recent years membrane research focused primarily on the lipid components. This focus on membrane lipids to the relative neglect of the proteins was encouraged by the fact that the most obvious role of membrane in their barrier function. This requires that unregulated passage of water-soluble materials be prevented, and the lipids, with their markedly hydrophobic hydrocarbon chains, seemed logical candidate for this task.

But now people have realized that the proteins have a major contribution in membrane structure and function. Basically, the proteins give the skeletal structure of the cell membrane. These proteins are either bound to the charged surface of the bilayer or intercalated to varying degrees into the hydrophobic interior of the bilayer.

So the biological membrane is composed of lipids and proteins. The phospholipids, phosphatidylcholine (PC), phosphatidylethanolamine (PE) and phosphatidylserine (PS) are the major lipid component in the mammalian cell membranes. The amphiphilic molecules containing two fatty acid chains of different lengths and degrees of saturation connected to a glycerol backbone to which a headgroup is coupled. Other lipids in the membrane are sphingolipids, where fatty acids are connected to the amino alcohol sphingosin, and cholesterol, which is built up of four fused sterol rings with a hydroxyl group at one end of the molecule and a branched alkyl chain at the other end. Glycoglycerolipids are only available in low concentrations in animal membranes, but are widespread in plant and bacterial membrane. Transmembrane proteins are situated in the

phospholipid bilayer and peripheral membrane proteins are coupled to the bilayer surface. The transmembrane proteins have several functions, e.g., they import and export water, nutrients, ions and xenobiotics; transmit signals and are active in cell-cell recognition. Transmembrane proteins are often built up of membrane –spanning α–helices or β-sheets connected by hydrophilic loops. The three dimensional structures of proteins are unique and crucial for each protein's function.

Fluid mosaic model:

Singer and Nicolson proposed the most accepted model of membrane structure and their function in 1972. The fluid mosaic model is formally analogous to a two-dimensional oriented solution of integral proteins in the viscous phospholipid solvent. The thermodynamic considerations and experimental results fit in with the idea of a mosaic structure of membrane in which globular molecules of integral proteins alternate with the sections of phospholipid bilayer in the cross section of the membrane. The globular proteins are postulated to be amphiphatic as are the phospholipids. So, they are structurally asymmetric, with highly one polar end and one non-polar end. The highly polar region is one in which the ionic amino acid residues and any covalently bound saccharides residues are clustered, and which is in contact with the aqueous phase of in the intact membrane. The nonpolar region is devoid of ionic and saccharide residues, contains many of the nonpolar residues, and is embedded in the hydrophobic interior of the membrane. The amphiphatic nature of the protein is determined by the amino acid sequence. The protein-protein interaction may be important in determining the structure and function of the membrane. The phospholipids of the mosaic structure are predominantly arranged as an interrupted bilayer, with their ionic and polar head groups in contact with the aqueous phase.

But from the two-dimensional model it is not evident which is the major constituent in the membrane, is it the protein on phospholipid? This question must be answered if the third dimension of the mosaic structure is specified. According to Singer and Nicolson fluid mosaic model functional cell membranes have a long-range mosaic structure with the lipid constituting the matrix.

In short the membrane has a flexible structure and can reseal spontaneously if disrupted. The two sides of the membrane are different in both lipid and protein composition. The phospholipids and proteins can diffuse laterally and transversely (flip-flop) in the membrane.

Though the membrane does not have any chromophore but the external chromophores are used to study the structure of the membrane, its interaction with the drug and membrane fusion in details. Sometimes the drug molecules are also used as a chromophore.

Protein

Proteins are macromolecules with specific three-dimensional conformation formed by different amino acids using the genetic information into a polypeptide chain. Proteins play many important roles in the biological system. Some of them are important structurally and some functionally. The polypeptide chains are able to fold into a number of regular structures, which are held together by hydrogen bonds. These secondary structures as α-helix, β-sheet and other minor secondary structures and connecting loops folding in three dimensions gives the tertiary structure of the polypeptide. Most polypeptides fold spontaneously into the correct tertiary structure, as it is generally the lowest energy conformation for that sequence. However, correct folding *in vivo* is often assisted by proteins called chaperones that help to prevent misfolding of new polypeptides before their synthesis on the ribosome is complete.

Folding is such that amino acids with hydrophilic side chains locate mainly on the exterior of the conformation that corresponds to the functionally active state (native state of protein) where they can interact with water or solvent ions, while the hydrophobic amino acids become buried in the interior from which water is excluded. This gives overall stability to the structure. Various types of noncovalent interaction between side chains hold the tertiary structure together as van-der waals forces, hydrogen bonds, electrostatic salt bridges between oppositely charged groups and hydrophobic interactions between nonpolar side chains of the aliphatic and aromatic amino acids. In addition, covalent disulfide bonds can form between two cysteine residues that may be far apart in the primary structure but come close together in the folded tertiary structure.

How an unstructured (random coil) polypeptide can rapidly and efficiently finds one functionally appropriate folded structure from countless alternatives is a crucial problem to the scientific community. The Endoplasmic Reticulum is a specialized folding compartment, which ensures that only correctly folded molecules proceed further into the secretory pathway and that persistently misfolded proteins are directed to degradation. Folding may be either cotranslational or posttranslational depending on the organism and complexity of the protein involved. Unfolded proteins are bound to classical molecular chaperones and specialized lectines. This association favors protein folding by, as yet unknown mechanisms. *In vitro* experiments have shown that interaction with molecular chaperones prevent the aggregation of unfolded polypeptides. It has generally been assumed that the same is true *in vivo*.

Folded proteins are easily unfolded by applying high concentrations of urea or GdmCl, extremes of pH or high and low temperatures. Folding is a fast process, generally of the order of a few milliseconds to minutes. Theoretically, Levinthal has shown that within this short time period polypeptide chain can not acquire all possible conformation states i.e., unfolded polypeptide chain goes to the folded state via a definite path way.

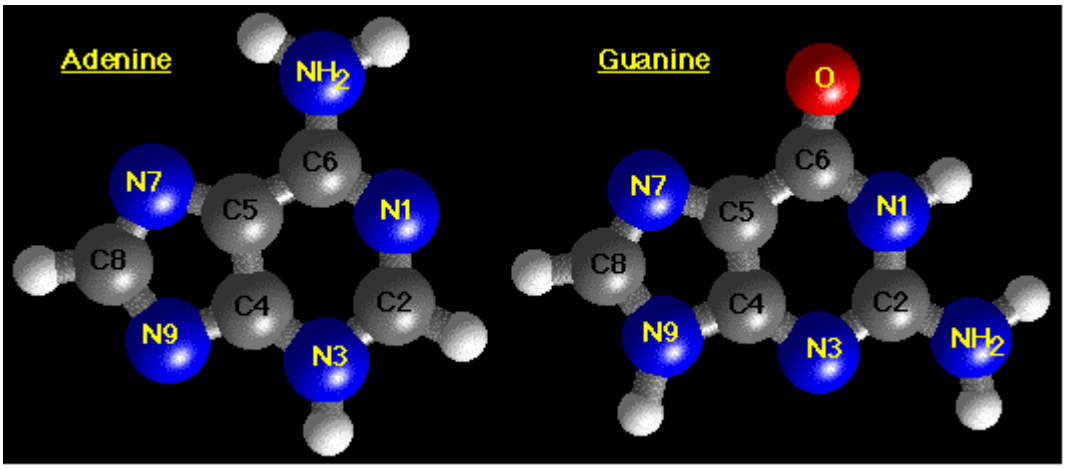
Tryptophan, tyrosine and phenylalanine are spectroscopically active amino acids present in the protein and their spectral properties are used to determine the structure of the protein and their unfolding and folding mechanism.

Nucleic Acid

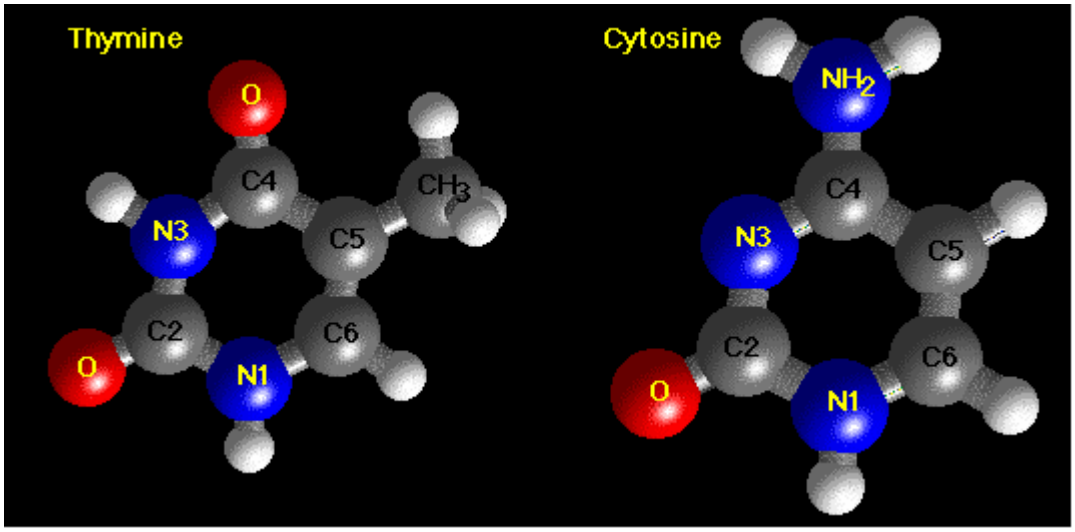
Configuration refers to the covalent bonding in a molecule; conformation refers to the three-dimensional (3D) structure of a molecule. The configuration is constant for a molecule; the conformation depends on the environment (temperature, pH, salt concentration, interactions with other molecules, etc.) the conformation is dynamic; there is always a range of different structures that the molecule samples at equilibrium. The primary structure is the configuration; the secondary structure and tertiary structure are the conformation.

The bases of DNA and RNA that are responsible for coding the genetic information are adenine, cytosine, guanine and thymine for DNA and adenine, guanine, cytosine and uracil for RNA. Each base is essentially planar and its conformations are limited. Of

course, all of the bonds undergo bending and stetching virations, which leads to in-plane and out-of-plane ring breathing modes and amino acid umbrella motions. There is rotation about the bond joining each amino group of a base.

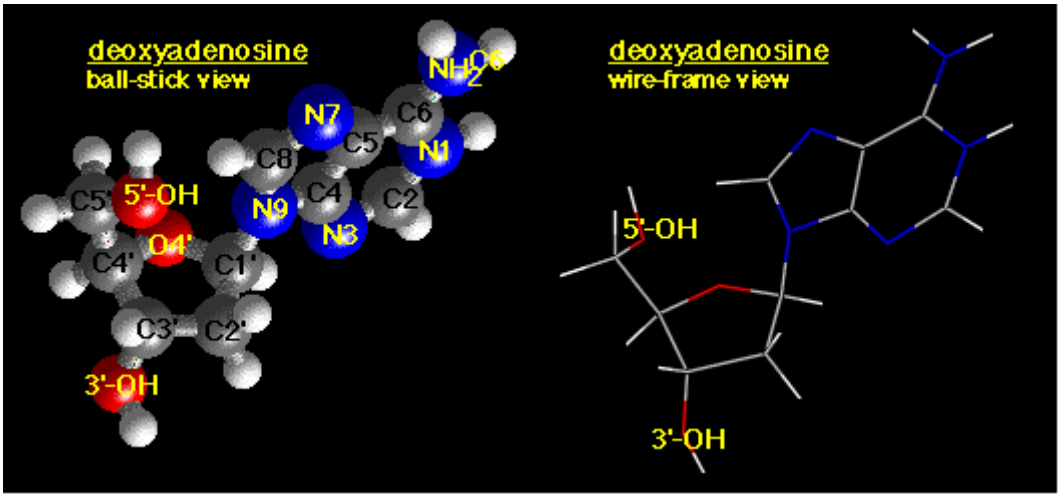


Structure of Adenine and Guanine

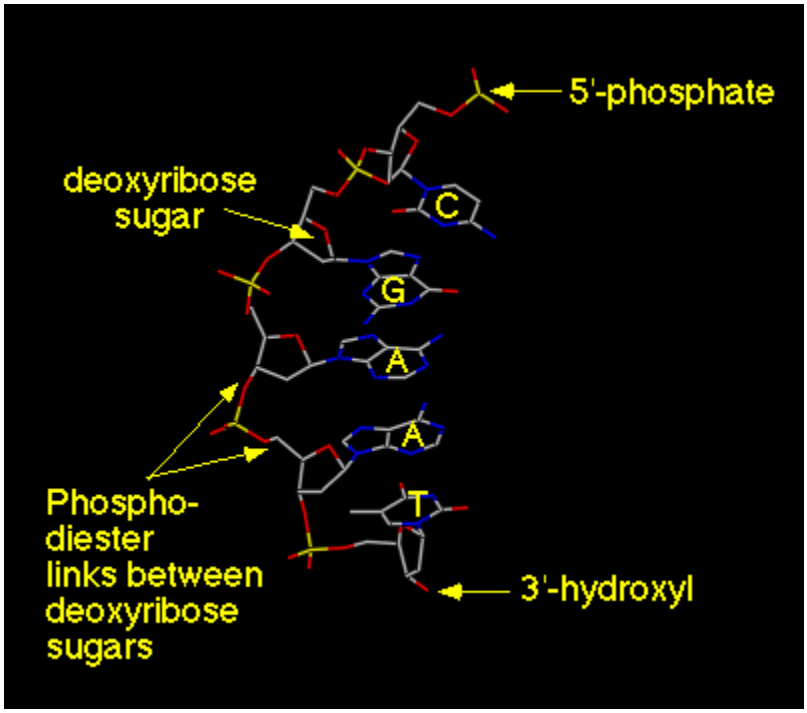


Structure of Thymine and Cytosine

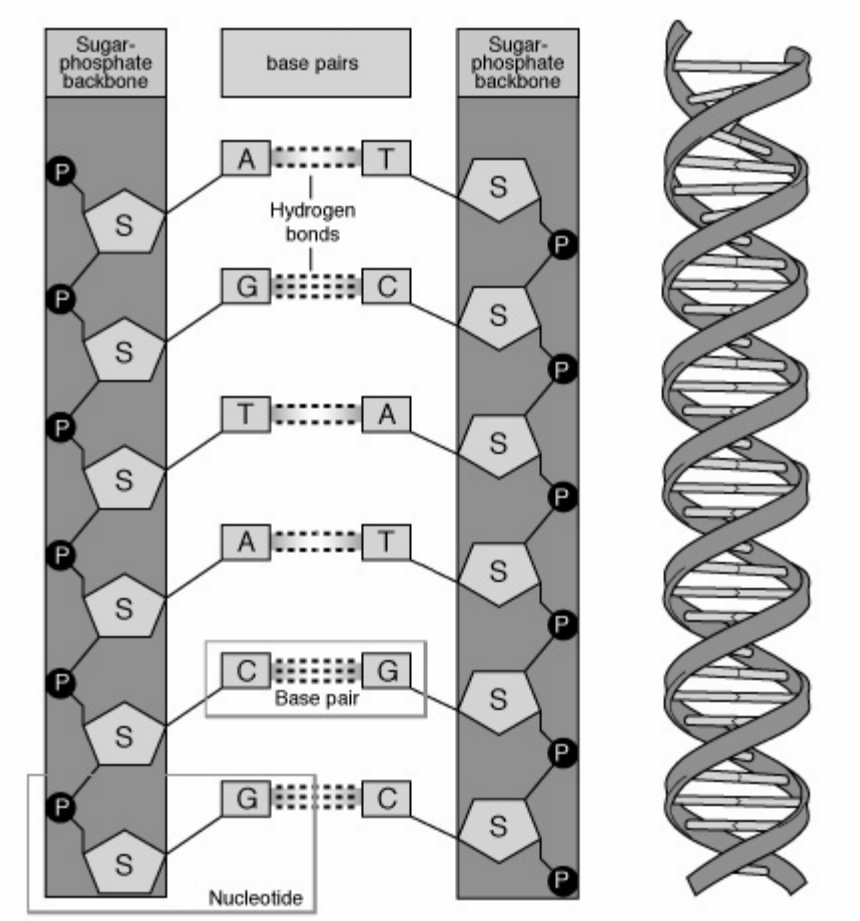
A nucleoside is formed by attaching the base to a sugar and nucleotides are formed by adding phosphates groups to nucleosides.

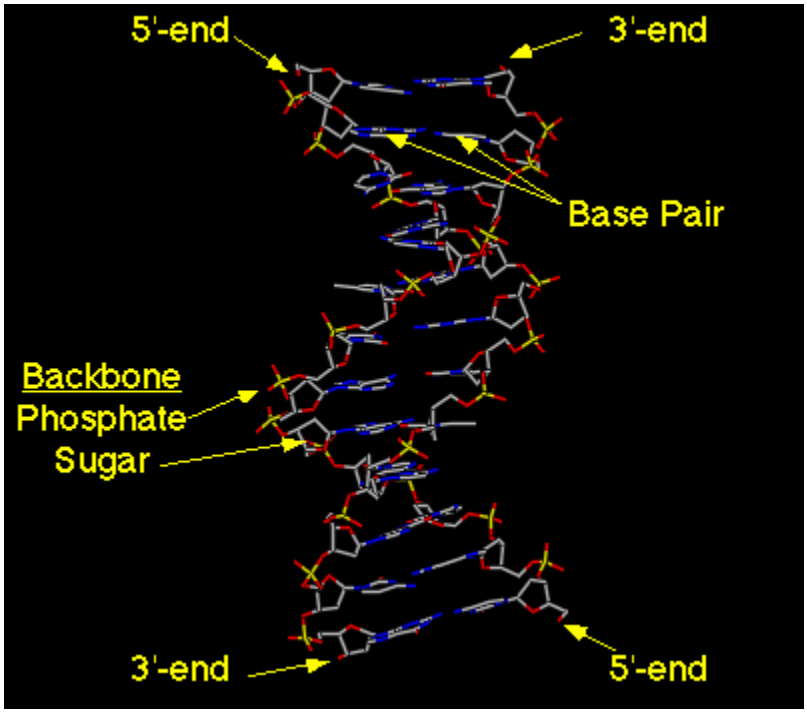


Picture of a nucleoside



DNA Backbone





DNA double helix

The DNA double helix was identified by Watson and Crick in the late 1950's as being the most likely structure for DNA. This structure has been verified in many different ways and its identification is recognized as one of the most significant breakthroughs in modern science.

Method of DNA sequencing:

There are several methods available to determine the actual sequence of nucleotides in a segment of DNA. One procedure uses specially altered nucleotides called dideoxynucloetides, which have been made either radioactive or fluorescent.

When enough of the targeted DNA fragment (marker) has been amplified (multiplied) by PCR (Polymerase Chain Reaction), the mix is put through a series of DNA sequencing reactions that are a variation of PCR. The amplified product from the PCR is added to a reaction tube containing the same Taq DNA polymerase used in the PCR, a primer that can hybridize at the desired location on only one complementary strand of the DNA (as opposed to both strands in PCR), and all four of the nucleotide bases (A, T, C, G.)

In addition, small amounts of fluorescence labeled dideoxynucleotides (A, T, C, G) are added to the mixture. Dideoxynucleotides are human-made nucleotides whose sugar component is slightly different from that of the nucleotides that make up DNA. (There is no OH on the 3' carbon.) Dideoxynucleotides can be picked up and added to a growing DNA chain. However, as a result of this structural difference another nucleotide cannot be added at its 3' end. Consequently, if one of the dideoxynucleotides is added to a growing chain of nucleotides, the strand will be terminated. Each dideoxynucleotide is labeled with a different fluorescent compound so that it will give off an identifying color in a laser beam.

After 20 - 30 cycles of the PCR heating and cooling, the resulting mixture will contain a series of fragments of different lengths depending on how many bases had been added to the chain before one of the dideoxynucleotides sneaked in and blocked further growth.

The mix of billions of short fragments from the sequencing reactions are loaded into glass capillary tubes that contain a gel solution that serves as a sieving matrix. During electrophoresis, a voltage is created across the gel so that one end is made positive and the other negative. Since DNA is slightly negative, its fragments will move to the

positive end of the gel. Not surprisingly, the different length DNA strands migrate at different rates and therefore separate from each other according to size. The smallest strand travels the fastest. As each DNA fragment reaches the bottom of the capillary tube, its fluorescence-labeled end dideoxynucleotide is excited by a laser beam that is directed at the bottom of the tube. Each of the four dideoxynucleotides fluoresces a different color when illuminated by a laser beam. An automatic scanner records each wavelength of light and a computer generates an electropherogram with colored peaks representing each wavelength in the sequence it passes through the beam. The 5' terminal base (the dideoxynucleotide) of the shortest fragment (that moves the fastest) is the first base in the electropherogram. The resolution is so good that a difference of one nucleotide is enough to separate that strand from the next shorter and next longer strand. At the left is a plot of the colors detected in one 'lane' of a gel (one sample), scanned from smallest fragments to largest. The computer even interprets the colors by printing the nucleotide sequence across the top of the plot. Alternatively, the DNA fragments are separated on a flat gel in a tray. The background of the title cell at the top of this page is a photograph of laser scan of a sequencing gel run in a tray.