Circular Dichroism

Norma J. Greenfield
9/20/04

I. Background and Definitions of Terms.

Optical rotatory dispersion (ORD) and circular dichroism (CD) are two phenomena that result when asymmetrical molecules interact with plane polarized light.

A. Definitions of asymmetry:

1. Molecules which lack a center plane or n fold axis of symmetry, e.g. carbon atom with 4 different substituents, disulfides and allenes (Figure 1).

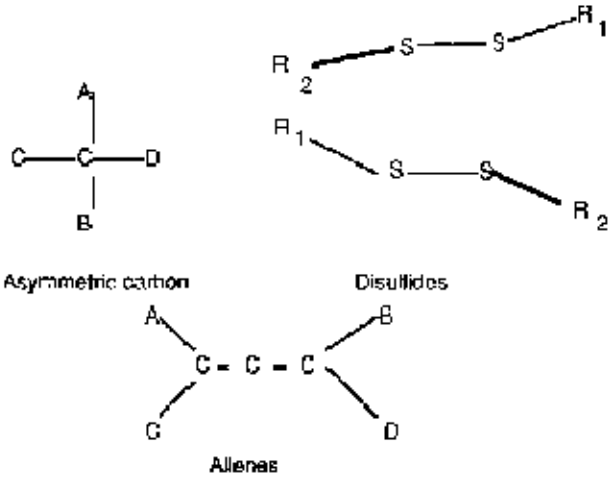


Fig. 1. Examples of asymmetric molecules

2. Arrangement of molecule in space is asymmetric, e.g. restricted rotation or spiral staircase arrangement such as a helical coiled-coil peptide (Figure 2).

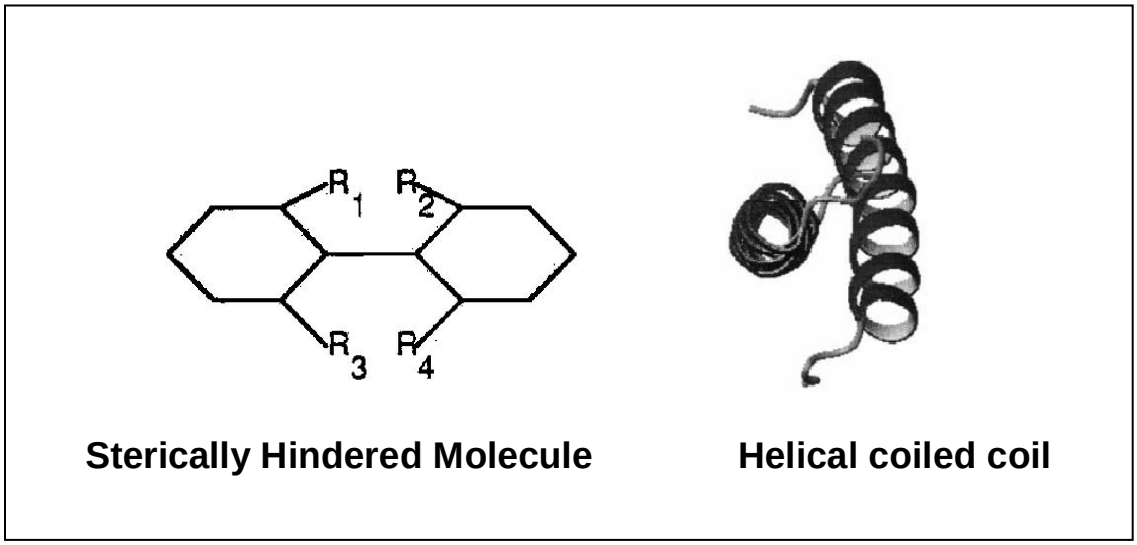


Fig. 2. Examples of asymmetric oligomers.

B. *What is* ***plane polarized light.***

Light can be separated into two forms of plane polarized light, parallel and Perpendicular polarized light, by several methods.

Classical methods:
1. Reflect light off a transparent media at an angle of 52^o 45' (Figure 3). The reflected ray is polarized in the plane of incidence and the refracted rat is polarized perpendicular to the plane of incidence.

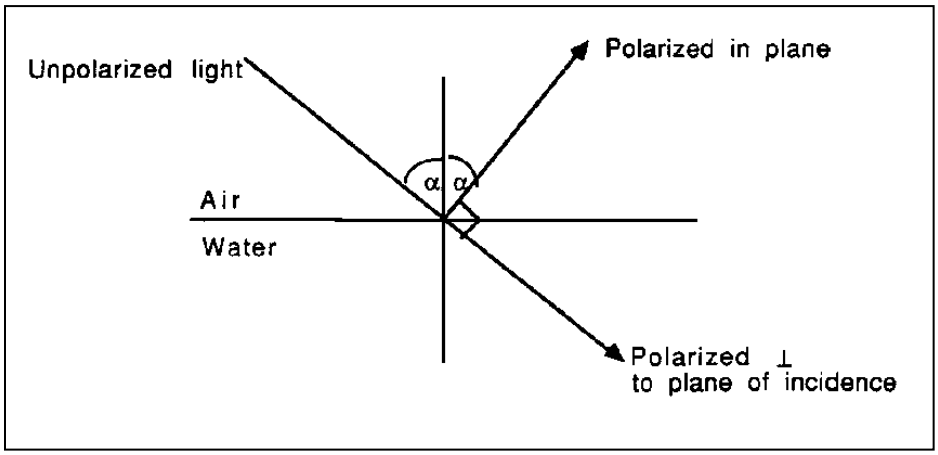


Fig. 3. Creating plane polarized light by bouncing light off of water.

2. Reflect light off of two glass mirrors (Figure 4).

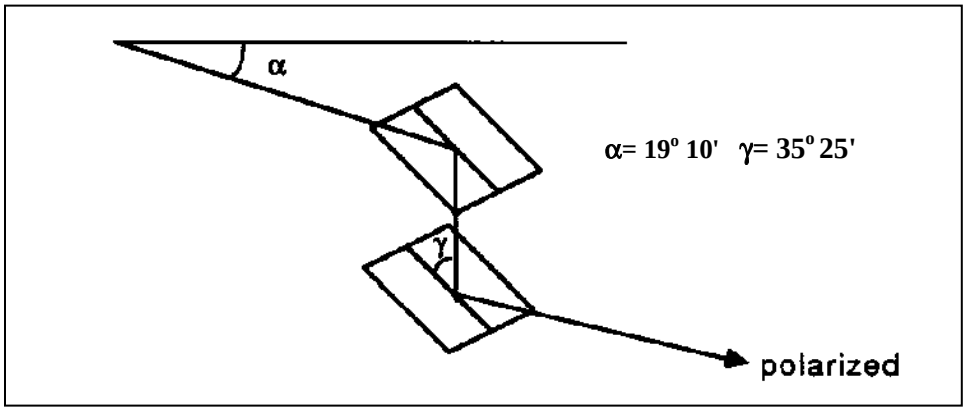


Fig. 4. Create polarized light by reflecting light off of a mirror.

3. Pass light though a Nicol Prism (Figure 5).

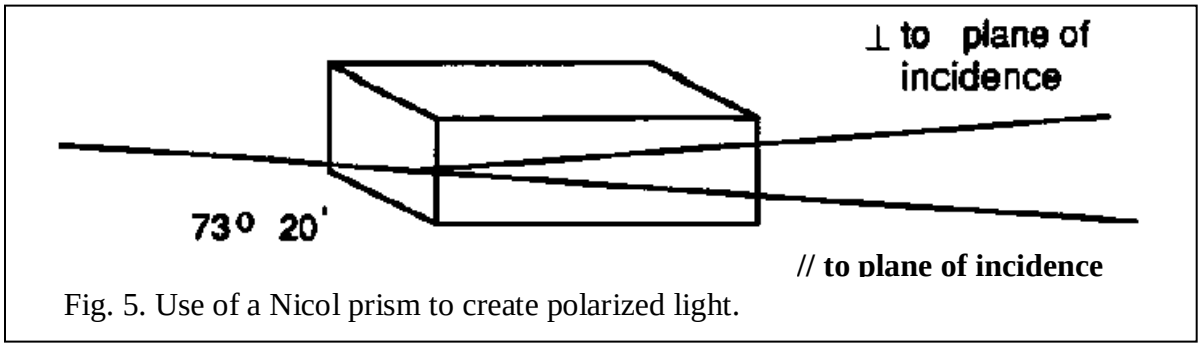


Fig. 5. Use of a Nicol prism to create polarized light.

Plane polarized light can be visualized as a sign wave oscillating in a plane (Figure 6).

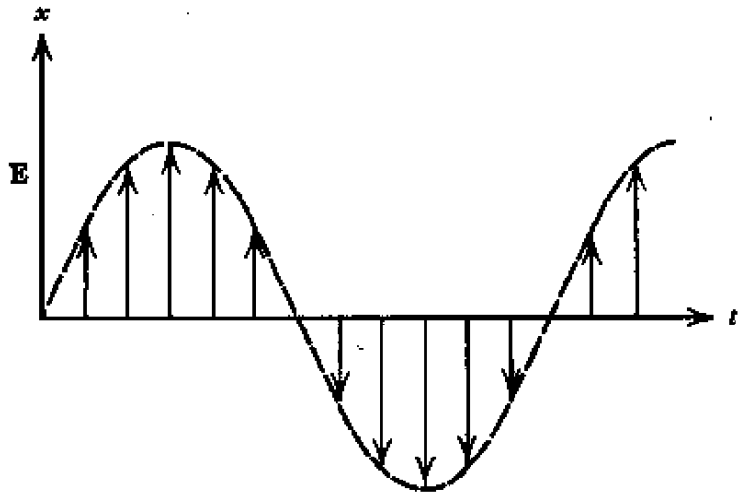


Fig. 6. Electric (E) as a function of time in a monochromatic linearly polarized wave. After Moscowitz, A C. Djerrasi, ed. Optical Rotatory Dispersion, McGraw-Hill Book Co, 1960

The sine wave can be visualized as the sum of two equal vectors of circularly polarized light, one going in a right handed direction and one in a left handed direction. It can also be thought of as the sum of two perpendicular planes of polarize light which are 90 degrees out of phase with one another (Figure 7).

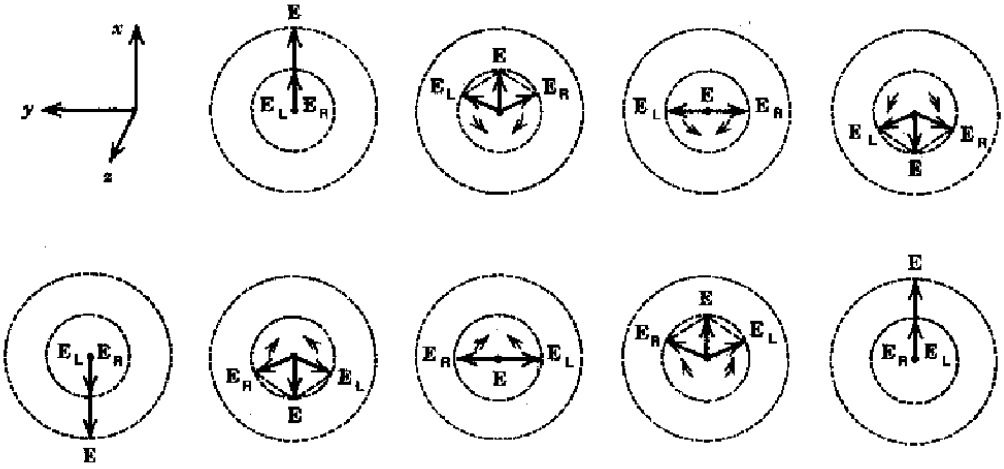


Fig. 7. A schematic representation of the separation of light into right and left circularly polarized light. The sum of the vectors traces the sine wave shown in Fig 6. After Moscowitz, A in C. Djerrasi, ed. Optical Rotatory Dispersion, McGraw-Hill Book Co, 1960

There is a very neat website which illustrates CD waves as a function of time in three dimensions. http://www.enzim.hu/~szia/cddemo/edemo0.htm

The two beams of right handed and left handed polarized light can be separated from one another by several methods:

1. A Fresnel rhomb is a quartz crystal which utilizes two internal reflections to retard the two components of plane polarize light by 90°.

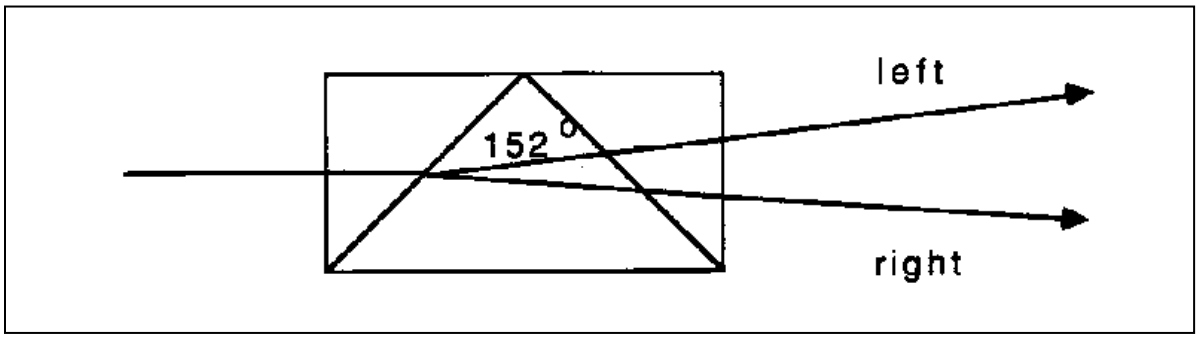


Fig. 8. Separation of plane polarized light into left- and right-handed circularly polarized light.

2. A quarter wave plate consists of a thin lamina of a uniaxial or biaxial crystal cut parallel to the optic axis. It is of such a thickness that the ordinary and extraordinary rays suffer a retardive phase difference of 1/4 wavelength when passing though it. It only works for light of a fixed wavelength.

3. A Pockles cell is a crystal of ammonium dihydrogen phosphate. Application of an electric field can modify its optical properties. The change in properties depends on the voltage applied and will give the crystals the properties of a quarter wave plate. Where V=1350 + 1.06 (A-2200).

C. What is optical Activity?

When a molecule is optically active, it absorbs right-and left-handed circularly polarize light to different extents. This is called circular dichroism (CD). It also has a different refractive index for the two forms of light. This results in the rotation of the plane of polarization of the light. The rotation is dependent on the wavelength, so the effect is called optical rotatory dispersion (ORD). When a molecule exhibits a combination of ORD and CD in the region of absorbance, then the transmitted light is said to be elliptically polarized.

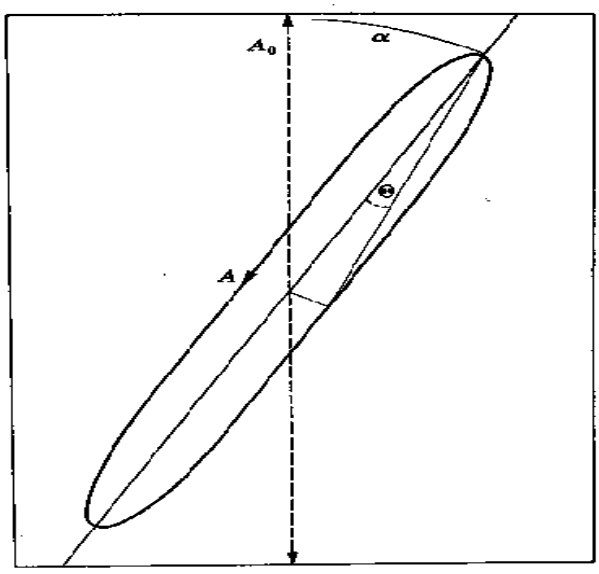


Fig. 9. A schematic representation of circular dichroism, θ, and optical rotation, α, of polarized light. The dashed line Ao represents the plane-polarized monochromatic ray striking the sample. The transmitted ray *A* is elliptically polarized. After Tinoco, I. Jr .and Cantor, C.R. in Glick, D., ed. Methods of Biochemical Analysis, Vol 18, John Wiley and Sons, New York, 1970

CD and ORD are phenomena result from the same electronic transitions in a molecule and are related to one another by the Kronig-Kramer transform.

$$\phi(\lambda)=2/\pi \int_o^\infty [\theta(\lambda')][\lambda'^2/ (\lambda^2 -\lambda'^2)]\delta\lambda' \qquad (1)$$

$$\theta(\lambda)=-2/(\pi\lambda) \int_o^\infty [\phi(\lambda)][\lambda'^2/ (\lambda^2 -\lambda'^2)]\delta\lambda' \qquad (2)$$

where $\phi(\lambda)$ is the rotation at wavelength λ and $\theta(\lambda)$ is the ellipticity at wavelength λ' in equation 1, and $\theta(\lambda)$ is the ellipticity at wavelength λ and $\phi(\lambda')$ is the rotation at wavelength λ' in equation 2. Thus if you have the CD of a molecule you can calculate its ORD and visa versa.

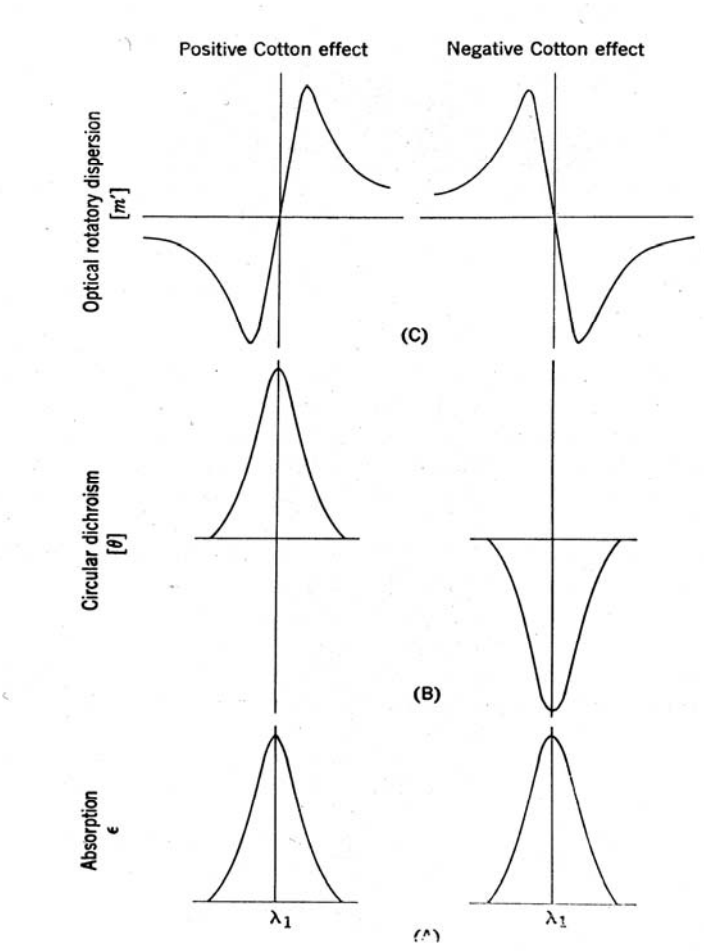


Fig. 10. The absorption band of a typical UV transition of an optically active molecule is illustrated on the left. The Fig. 10 also illustrates the associated CD band (B), and the associated ORD of the band (C). CD spectra are coincident with the absorption spectra of molecules. However, ORD spectra are dispersive and have long "tails" of rotation to longer and shorted wavelengths than the absorption spectra. Thus ORD measurements can be made of the rotation of light by molecules in regions where they do not absorb light. From Greenfield, N.J. and Fasman, G.D. Polypeptides, characterization. Encyclopedia of polymer science and technology. J. Wiley and Sons, N.Y. 1971; 15:410-429.

D. Origins of optical activity.

An electronic transition is the result of the movement of charges when a molecule is exposed to light. The electronic transition has an associated magnetic transition that is perpendicular to it. The energy of a transition depends on the electric dipole moment and the magnetic dipole moment induced by the action of light on the electrons in the molecule. The rotational strength, $R_{K,}$ of a transition is the imaginary part of the dot product of the electric dipole induced by the light and the magnetic dipole induced by the light. (The dot product is the product of two vectors times the cosine of the angle between the vectors.)

$$R_K = \mathrm{Im}\ u_{ek} \bullet u_{mk}$$

If a molecule has a plane or center of symmetry either the sum of all the induced electric and magnetic dipoles is zero, or the vectors representing the magnetic and electric dipoles are perpendicular (orthogonal) to one another. The result is that there is no optical activity since the cosine of 90^o equals 0. There are several cases when the magnetic and electric fields will not be orthogonal and in these cases the molecule is optically active.

Case 1. There are two transitions in a single chromophore, one is magnetic and one is electric. The asymmetry in the molecule perturbs the field so the symmetry of the chromophore breaks down and the two transitions are no longer perpendicular to one another. This is the "one-electron theory" of Condon, Alter and Eyring. The n-π^* transition is electrically forbidden but is magnetically permitted. In an asymmetric molecule such a transition can have very low total absorption of light but high optical activity. This is found in the case of helical polypeptides and proteins.

Case 2. Two groups have simple electric transitions. These are coupled by their proximity to produce a magnetic moment. (This type of phenomena is an example of an exciton interaction and it is responsible for most of the optical activity of helices.)

Case 3. There are two chromophores within a molecule. One has a magnetic transition and one has an electric transition. These are coupled in the molecule to give strength to both transitions.

Quantum mechanical treatments of optical activity are very difficult. It is relatively easy to calculate the electric dipoles of a molecule in the ground state but these are perturbed in the excited state. [For an introduction to the quantum mechanics of circular dichroism see Cantor, C.R. and Shimmel, P.R. (1980) Biophysical Chemistry Part II, W.H. Freeman and Co., New York, 1980.]

The rotational strength of an optical transition is R_k

$$R_k = (3\ h\ c / 8\ \pi^3 N) \int_o^\infty (\pi\phi_k(\lambda)\delta\lambda)/\lambda$$

where ϕ_k is the optical rotation at wavelength λ, h is Planck's constant and c is the speed of light.

In terms of circular dichroism, this is approximately equal to:

$$R_k \equiv 1.23 * 10^{-42} [\theta_K] / \lambda_{max}$$

where θ_K is the ellipticity of the K^{th} band at wavelength λ and Δ is the half band width of the transition at θ_k/e.

E. Unit of ORD and CD.

Specific rotation is classically defined as:

$$\alpha_\lambda = 100\alpha_{obs} / l\,[c]$$

where [c]=grams/100 ml and l=decimeters (10 centimeters).

Molar rotation is defined as:

$$[M] = MW /100[\alpha_\lambda] = 10\lambda_{Obs}[C]_{(Molar)}$$

where MW = the molecular weight of the compound. Since the original measurement were made on 100 ml of solution and the concentration was in moles/liter the units are expressed as per decimole, i.e. deg·cm^2/dmole.

Residue rotation is defined as:

$$[M] = MRW\ \alpha_\lambda/100$$

where MRW is the mean residue weight of the protein or polymer.

Reduced residue rotation is defined as:

$$[M'] = (3/ (\eta^2 + 2)) [M]$$

where η is the refractive index of the medium.

The same units are used in circular dichroism. No correction for refractive index is made. The mean residue ellipticity of a polypeptide or protein is defined as:

$$[\ \theta_\lambda] = 100\ \theta_{obs} / (cl).$$

Here the concentration is in moles/liter of residues and the pathlength is in centimeters. The units are in deg.·cm^2 /dmole as in rotation. When θ_{obs} is in millidegrees, and the pathlength, l, of the optical CD cell is one millimeter, and the concentration of residues, c, is in moles per liter, $[\theta]$**,**the mean residue ellipticity in degrees · cm^2/dmol is easy to calculate:

$$[\theta] = \theta_{obs} / (cl)$$

CD and ORD are products of the same phenomena. The degree of rotation is closely related to the degree of ellipticity.

$$[M_{peak} - M_{trough}] \cong 1.2\ [\theta]_{max}$$

V. Applications of CD spectroscopy

CD spectroscopy has many applications detailed below:

A. Estimation of protein and nucleic acid conformation

B. Determination of conformational changes due to the interactions of asymmetric molecules e.g.

1. Protein-protein interactions
2. Protein-DNA interactions
3. Protein-Ligand interactions
4. DNA-Ligand interactions

C. Determination of the thermodynamics of folding and unfolding of proteins and nucleic acids

D. Determination of binding constants by:

1. direct titrations
2. serial dilutions of complexes
3. changes in stability to thermal or chemical denaturation

E. Kinetics of folding and unfolding of macromolecules

A. Conformational dependence of the ORD and CD of polypeptides and polynucleotides

1. CD and ORD of Polypeptides and Proteins

a. Transitions of the Amide Bond

In peptides, the chromophore which interacts with light is the amide bond. In simplified molecular orbital theory the 2Px orbitals of the nitrogen, carbon and oxygen atoms of the amide linkage are combined to form three orthogonal linear combinations, the π^+, π^o, and π^- bonds. These three orbitals have the plane of the peptide bond as there nodal plane. The four π electrons fill the first two molecular orbitals, the π^+ which is strongly bonding and π^o which is almost non-bonding. The remaining two electrons are non-bonding electrons on the oxygen atoms of the amide bond which are conventionally labeled the 2Py electrons. These electrons have energy which is close to that of the atomic orbitals of hydrogen. The top lobes of the amide bond are illustrated below in Fig. 11. These lobes are above the plane of the amide bond. The bottom lobes are mirror images of the upper lobes but are opposite in sign. The n-π* transition involves the promotion of an electron from the n to the π^- orbital. The transition is very weak because the ground and excited states have nodal planes which are perpendicular to one another. Thus the n-π* transition is said to be forbidden. The transition is forbidden electronically but permitted magnetically. Thus the absorption band of the n-π* transition is very weak, but the CD bands can be very large. The π -π* transition involves the promotion of an electron from the π^o to the π- orbital and is strong with an extinction coefficient equal to approximately 7100. A transition to an σ orbital is called a "Rydberg" transition. In this case an electron from the n orbital is excited essentially to atomic orbitals of higher energy than the ground state on the carbon, nitrogen and oxygen atoms of the amide. These are also called transitions to "big orbit" orbitals.

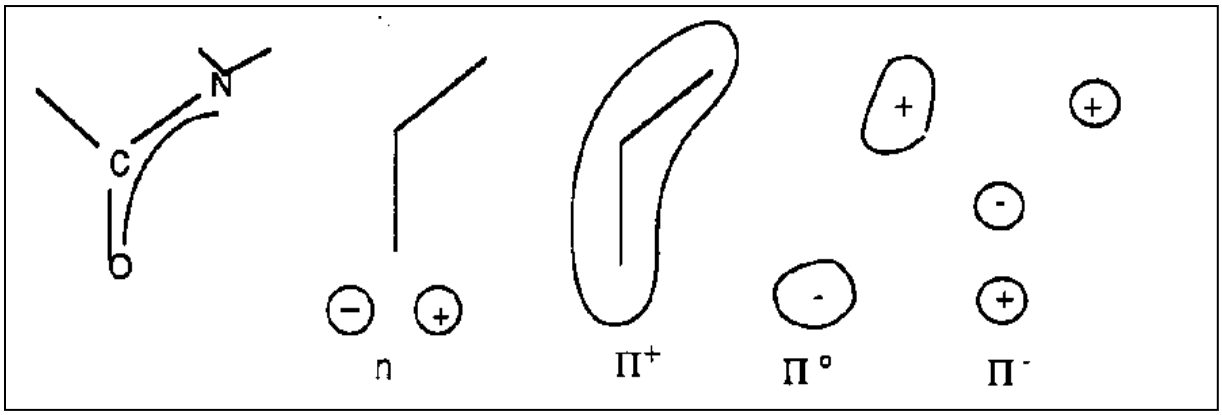


Fig. 11. The upper plane of the orbitals of the amide bond. The lower plane has the same lobes which are opposite in sign. Redrawn from Schellman, J.A. and Schellman, C. (1964) Conformation of polypeptide chains" in the Proteins, 2nd ed. Vol 2. Neurath H. ed. Academic Press, 1-137.

The first study of the UV spectra of amides were performed by Ham and Platt in 1953 on the spectra of glycyl glycine. They observed a peak at 185 nm and postulated that it was due to a π -π* transition. However they felt than an n-π* transition was also present but obscured by the larger π-π* band. Simpson and coworkers studied the spectra of formamide and myristamide and found a weak band at 320 nm and a strong band at 170-190 nm depending on the substituents which they ascribed to the n-π* and π -π*transition of the amide bond respectively References: Ham, J.S. and Platt J.R. (1953) J. Chem. Phys. 20:355; D.L. Peterson and W.T. Simpson (1957) Polarized Electronic Absorption Spectrum of Amides with Assignments of

Transitions, J. Am. Chem. Soc.; 1957; 79:2375-2382; H. D. Hunt, W. T. Simpson (1953) Spectra of Simple Amides in the Vacuum Ultraviolet; J. Am. Chem. Soc. 75: 4540-4543; D. L. Peterson, W. T. Simpson (1955) Polarization of the 1850 Å. Band of amides, J. Am. Chem. Soc.,77:3929-3930.

The transitions of the amide bond are affected by solvent and environment. The wavelength of the n-π* transition blue shifts in water compared the wavelength in organic solvents. This happens because the lone pair electrons in the ground n state are hydrogen bonded by water. Higher energy is therefore needed to promote electrons from the ground state to the excited state. The π-π* transition, on the other hand, red shifts in water because the excited state of the π-π* transition is more polar than the ground state. The π-π* transition is a charge transfer from the nitrogen to the oxygen atoms. The charged nitrogen atom forms a hydrogen bond with the aqueous solvent. The stabilization of the excited state leads to the effective lowering of the ionization potential of the amide. [Nielsen, E.G. and Schellman, J.A. (1967) The absorption spectra of simple amides and peptides, J. Phys. Chem. 71;2297-2304]

The n-π* and π -π* transition are seen in the circular dichroism and optical rotatory dispersion of small cyclic amides such as 3-amino-pyrrolidone [B. J. Litman and J. A. Schellman, J . Phys. Chem., 69,978 (1965); Greenfield, N., Fasman, G.D. Optical activity of simple cyclic amides in solution Biopolymers 7,595-610, 1969] and 3-methyl pyrrolidone [Greenfield, N. and Fasman, G.D. The Circular Dichroism of 3-Methylpyrrolidin-2-one, J. Am. Chem.Soc. J. Am. Chem. Soc.; 1970; 92(1); 177-181]

b. The conformational dependence of the spectra of polypeptides.

Polypeptides often exist in highly ordered structures. Therefore, their chromophores are in ordered arrays. In a system of N non interacting chromophores if one chromophore is excited by light, the energies of the other N non interacting molecules are unchanged. Thus an array of N chromophores would have N-fold degeneracy, as any chromophores could be equally excited by the light. In biopolymers, however, the chromophores are in ordered arrays, and when electrons are excited they do interact. For example, if one chromophore is excited, its transition dipole may interact with the transition dipole of a neighboring unexcited molecule. The net result is an exchange of excitation. The energy of the transition will be perturbed by the potential which allows excitation energy of one molecule to migrate to another.

If one had only two identical chromophores in a crystal, and one was excited and the other was not, there could be resonance of the excitation between the two chromophores. It would be unclear, at any given time, which of the two chromophores was in the excited state. The exchange of excitation energy between the chromophores leads to two states, one of which has higher energy, and one of which has lower energy with respect to two isolated non interacting chromophores. In the case of an ordered array with two chromophores, the transition dipole of one of the isolated chromophores will be split into two transitions, one of which having lower energy and one of which having higher energy than the original transition. This splitting of one transition into two is called "exciton" splitting.

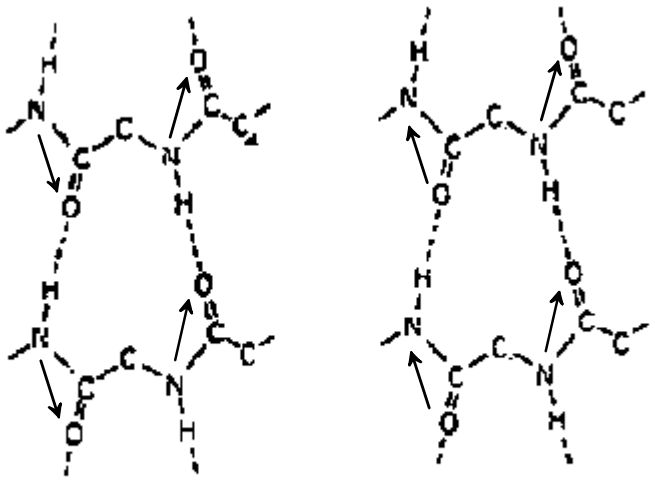
The monomer units of the two chromophores in an ordered array are described by their wave functions Ψ_1 and Ψ_2. The oscillation of excitation energy will be between $\Psi^*\Psi_2$ and $\Psi_1\Psi^*$. The all over excited state of the dimer will be given by:

$$\Psi = (1/\sqrt{2})\,(\Psi_1\Psi^*_2 +/- \Psi^*_1\,\Psi_2)$$

Exciton splitting can be seen in the ORD and CD spectra of amides in dipeptides, e.g the cyclic peptide alanyl-glycyl [Greenfield, N., Fasman, G.D. Optical activity of simple cyclic amides in solution Biopolymers 7,595-610, 1969].

If one had N such chromophores in a crystal it would look as if one could then get splittings into N such transitions. Since a typical polypeptide has hundreds of amide groups the situation could get out of hand and one could expect optical transitions to have very broad absorption bands. Fortunately in an ordered array, such as in a crystal or a polypeptide one only has to look at the smallest number of molecules than will generate the whole array by simple translation. The smallest number of chromophores in a crystal or polymer of repeating units is called the unit cell. If an isolated polypeptide chain is in the α-helical conformation, two amino acid residues can generate the whole structure by simple translation. Thus an optical transition of an amide bond in an isolated α-helix can be split into a maximum of two transitions. Usually polypeptides in the β-strand conformation are arranged in rows which are either parallel or antiparallel (see below). For the parallel and antiparallel β–pleated sheets one needs a minimum of 4 amino acid residues to generate the whole structure. One therefore might expect the π-π* transition to be split into 4 transitions. However, while the maximum number of theoretical transitions might be four, the only transitions that can be observed are those that are in phase with each other. In order for an electronic transition to occur it is necessary that a charge moves, or that a dipole be created. If a dipole cannot be created than the transition is forbidden.

By looking at the symmetry of a crystal (or polymeric molecule) one can make some predictions about whether a given transition will be allowed, and how it will be polarized. The predictions are made on the basis of the symmetry of the unit cells. All crystals have only certain types of symmetry. These symmetry types are called space groups. By determining the symmetry elements of allowed protein conformations, and determining in which space groups they belong, one can predict whether a transition will be allowed and what the polarization of a transition will be. Figure 12 illustrates how transitions are allowed or forbidden in parallel and antiparallel beta sheets.



Parallel Perpendicular

Fig. 12A. A diagram illustrating the allowed transitions of polypeptides in the parallel β-pleated sheet conformation.

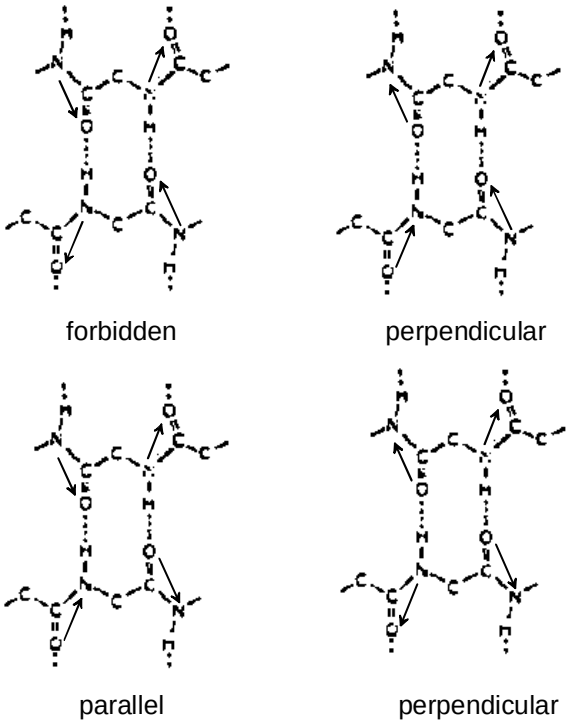


Fig. 12B. A diagram illustrating the allowed transitions of polypeptides in the antiparallel β-pleated sheet conformation.

A polypeptide in a parallel pleated sheet conformation has only one element of symmetry, a screw axis of symmetry. This puts it into space group C_2. This symmetry will permit an optical transition of the parallel pleated sheet to be split into only two allowed transitions, one polarized perpendicular to the axis of symmetry and one polarized parallel to the axis of symmetry (Figure 12A, above)

The antiparallel β-pleated sheet has three elements of symmetry, a screw axis of symmetry in the x direction, a screw axis of symmetry in the y direction and a simple axis of symmetry in the z direction. These symmetry properties place the antiparallel pleated sheet into space group D_2. In this space group structure will have 3 allowed transitions per monomer, one parallel to the fiber axis and two perpendicular to the fiber axis (Figure 12B, above).

The geometry of the α-helix is such that viewed end on, the transition moment vectors lie approximately along the sides of a square. Taking all additive and subtractive combinations of these vectors will lead to an absorption band parallel to the helix and two a degenerate pair of absorption bands perpendicular to the helix. The splitting of the π-π* transitions in the α-helix and β-pleated sheet are quite obvious in absorption spectra and the circular dichroism spectra of polypeptides.

In addition to splitting of electronic transitions, the placement of chromophores in ordered arrays will lead to hypo or hyperchromism of their optical transitions. What this means is that the absorption of a dimer will not be simple but will be equal to twice that of a monomer, but will be

higher or lower. Several generalizations about hypo and hyperchromism have been made by several groups. These apply to both polypeptides and polynucleotides.

1. When chromophores are aligned side by side (card stack) such in a helix the lowest energy transition will be hypochromic. The hypochromism results from intensity interchange with transitions of higher energy. (The total oscillator strength summed over all transitions of a molecule is constant).

2. When chromophores are aligned tail to tail, the lowest energy transitions will be hyperchromic, intensity being gained at the expense of the transitions of higher energy.

The UV spectrum of poly-L-Iysine in the α-helical and antiparallel β-pleated sheet is illustrated in Fig. 13, below

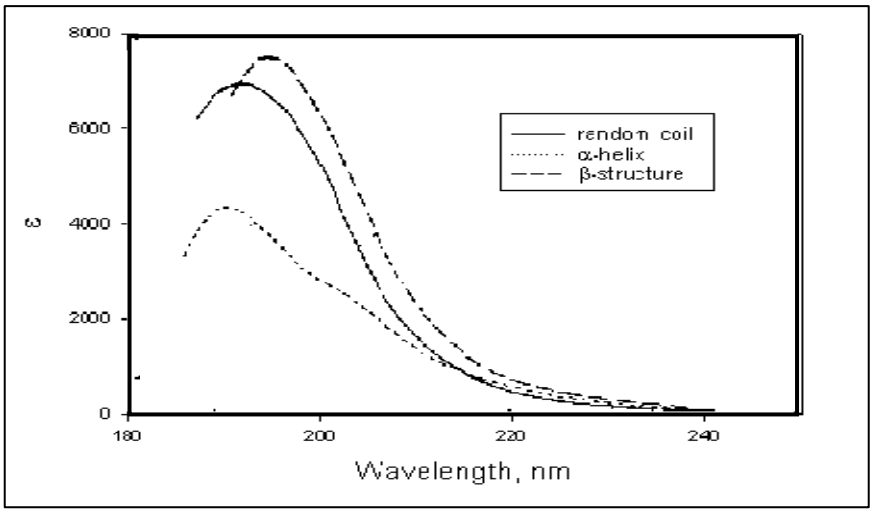


Fig. 13 Ultraviolet absorption spectra of poly-L-Iysine in aqueous solution. The α-helical peptide is hypochromic while the peptide in the antiparallel β conformation is hyperchromic.

The π-π* transition of the a helix displays hypochromism. Moreover it displays the splitting into transitions which are polarized. When the polymer is oriented in a film and exposed to plane polarized light it absorbs light maximally when the peptide is aligned parallel or perpendicular to the light source. In addition to the strong π-π*. transition, the a helix displays a large n-π* transition. The n-π* transition is enhanced in the α- helix because while the n and π orbitals of an isolated amide bond are orthogonal (i.e. perpendicular to each other), and the transition is forbidden, the orbitals of neighboring chromophores are not necessarily orthagonal and thus the transition is more likely to occur in a polymer. In addition, the helix environment shields the transition from the solvent, causing a red shift of the transition. In simple amides, the small n-π* transition often overlaps the larger π-π* transition. The antiparalleI β-pleated sheet form of poly-L-Iysine shows the predicted hyperchromism of the π-π* transition. Moreover, the splitting of the transition into two transitions is also evident when oriented films are irradiated with polarized light.

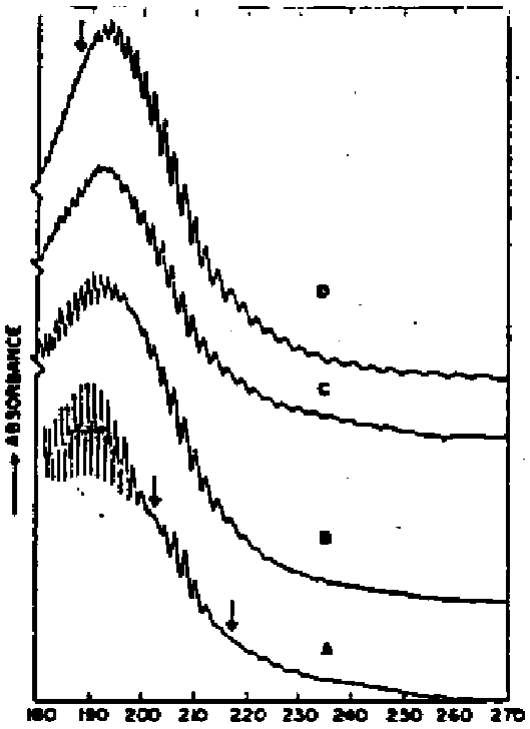


Fig. 14. Polarization spectra *of* oriented films or unionized poly-L-lysine.These are linear dichroism spectra in which the osciliations arise from the introduction into the sample channel of a quarter-wave retardation plate. The extrema of the oscillations correspond to parallel and perpendicular absorption at these wavelengths; the polarizaion spectra are then defined by the curves joining the extrema for either phase. Changes of polarization occur at the indicated points. A: *a*-helix showing the perpendicularly polarized *n-π** transition at about 215 to 240 nm and the split *π-π** transition. B. C. and D: samples containing increasing proportions or *β*-structure. From K. Rosenheck and B. Sommer (1967) Theory of the Far-Ultraviolet Spectrum of Polypeptides in the β-Conformation. The Journal of Chemical Physics 46: 532-536.

The same phenomena which affect the ultraviolet spectra of a polypeptide affect its circular dichroism and optical rotatory dispersion. The ORD and CD spectra of poly-L-Iysine in the α-helical, antiparallel β-pleated sheet and random coil conformations are illustrated in figure 15

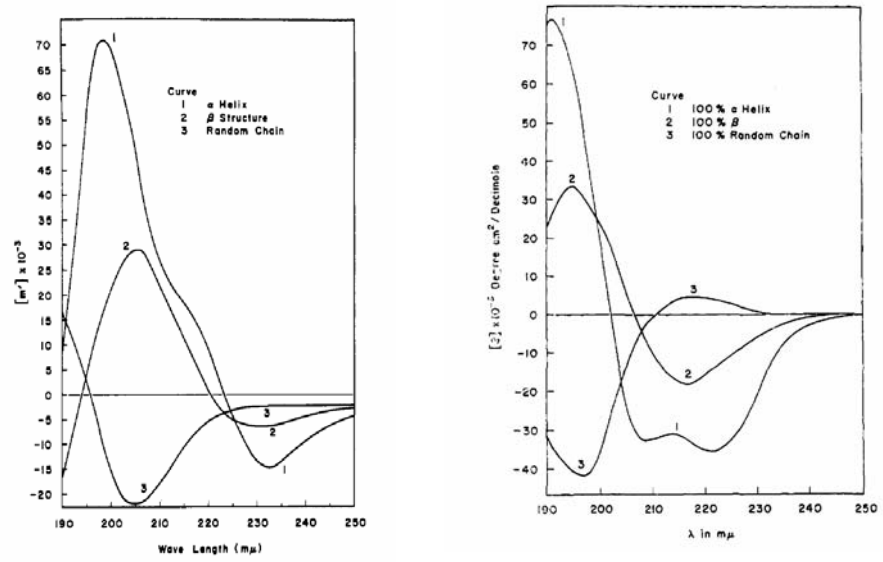


Fig 15. The optical rotatory dispersion (ORD left) and Circular Dichroism (CD right) of Poly-L-lysine in the α-helical, β- and random coil forms.

A polypeptide which is random has a CD spectrum which is similar to that of a simple amide, but is bigger in magnitude. The spectrum displays a small positive n- π∗. transition at approximately 230 nm and a large single π–π∗ transition at approximately 195 nm. The spectrum of the α-helix shows a large negative n-π∗ transition at 222 nm and a π–π∗ transition which is

split into two transitions because of exciton coupling. This transition has a negative band at approximately 208 nm and a positive band at approximately 192 nm.

The CD spectrum of the antiparallel pleated sheet also shows evidence of exciton coupling in the $\pi-\pi^*$. transition. However the splitting of the transitions are different than in the case of the α-helix. The spectrum of the antiparallel β-structure shows a negative band at 218 nm and a positive band at 195 nm.

Collagen is an unusual protein which contains a great many proline, hydroxyproline and glycine groups. This allows it to form two very special conformations which are not found in typical proteins. Unfolded collagen has a structure which is very similar to that of a homopolymer of poly-L-proline. This polymer can exist in two single stranded forms. The PI and PII. in the PI conformation the prolines are in the cis conformation and in the PII conformation they are trans. Native collagen has three chains, where the prolines are in the trans conformation that are supercoiled.

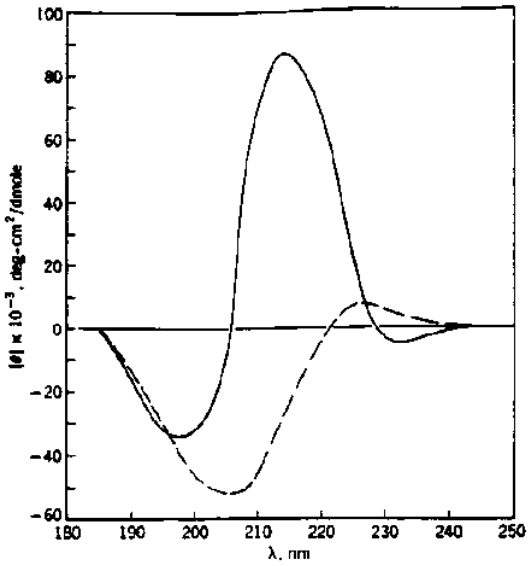


Fig. 16 The circular dichroism of poly-L-proline I and II in triflnoroethanol: poIy-L-proline I (solid line), poly-L-proline II (dashed line).

c. Historical methods of estimating the secondary structure of proteins using ORD and CD.

The first measurement of the optical activity of polypeptides and proteins were ORD measurements that were made in the visible region of the spectrum. It was only in the 1960's when instrumentation improved sufficiently to accurately measure the small differences in absorbance between right-handed and left-handed polarized light that CD measurement of amides could be performed.

The Drude equation describes the rotation of light by a chromophore which absorbs light at wavelength λ

$$[\alpha_\lambda] = K/ (\lambda^2 - \lambda c^2)$$

In the 1950's it was found by Yang and Doty W. [J. T. Yang and P. Doty (1957), J.A.C.S. 79:761] and by Moffitt [Moffit and J. T. Yang (1956), P.N.A.S. 42: 596] that while the Drude equation was adequate to describe the rotation of simple amides and random polypeptides it could not describe the optical activity of fibrous proteins or polypeptides known to be helical.

Moffitt proposed that in a alpha helix the π–π*. transition of the amide bond was split into two transitions because of exciton coupling. He added up the terms of the two Drude equations due to the split transition of the helix and the term due to a random polypeptide to come up with a general equation for describing the optical rotation of a protein.

$$[M'] = a_o \lambda_o^2 / (\lambda^2 - \lambda_o^2) + b_o \lambda_o^4 / (\lambda^2 - \lambda_o^2)^2$$

Plotting $M'(\lambda^2 - \lambda o^2)$ versus $1/(\lambda^2 - \lambda_o^2)$ gave a straight line when λ_o was chosen correctly. λ_o =212 nm usually worked. Fully helical polypeptides gave values of bo = -630 while random polypeptides gave a value of 0. It was attempted to estimate the fraction of alpha helix in a protein using these values.

In the 1960 Simmons et al. [N. S. Simmons, C. Cohen, A. G. Szent-Gyorgyi, D. B. Wetlauter and E. R. Blout, (1961)J. Am. Chem. Soc., 83: 4766] first detected the trough of a Cotton effect at 233 nm in the ORD of tobacco mosaic virus protein. They attributed this band to an optical transition of a helical polypeptide and for a while it was attempted to estimate the helicity of a protein from its optical rotation at 233 nm.

In the 1960's the first good ORD and CD measurements of the total CD spectra of the α-helix, β-form and random coil were obtained for a model polypeptide, poly-L-Iysine. It was thus possible to try to determine the secondary structure of proteins based on their ORD and CD spectra using poly-L-Iysine as a reference. It was attempted to fit the ORD and CD spectra of proteins by a linear combination of the α-helical, β and random forms. The method of least squares was used to determine the best fit. At the time the ORD measurements were performed the X-ray crystallography of only 2 proteins had been performed, myoglobin and lysozyme. Attempts to fit the ORD data of these two proteins to the spectra of poly-L-Iysine gave only poor fits. There was too much overlap of the ORD spectra of the reference conformations.
When the early CD measurements were obtained there was good data on only a few proteins, myoglobin, lysozyme and RNAase. In addition, preliminary data had been obtained on carboxypeptidase A., chymotrypsin and chymotrypsinogen. Fitting the CD spectra of these proteins with the poly-L-Iysine reference spectra gave remarkably good fits considering the limitations of the method. See table 1 and Figs 17 and 18.
[References :Greenfield, N., B. Davidson, G. D. Fasman. 1967. The use of computed optical rotatory dispersion curves for the evaluation of protein conformation. Biochemistry 6:1630-1637; Greenfield, N., G. D. Fasman. 1969. Computed circular dichroism spectra for the evaluation of protein conformation. Biochemistry 8:4108-4116]

β

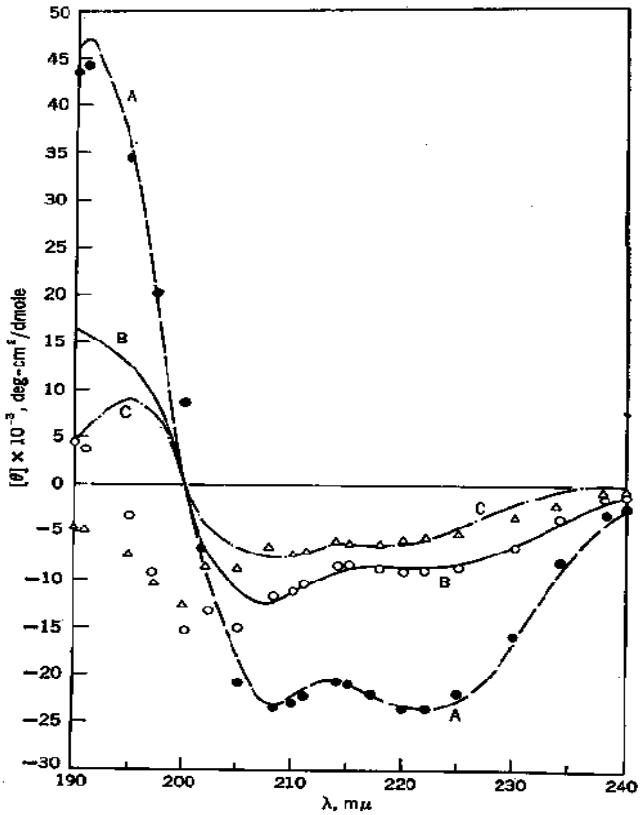


Fig. 17. The circular dichroism of (A) myoglobin (curve) and 68.3% α-helix, 4.7% structure and 27.0% random coil, calculated from poly-L-lysine reference spectra in water (●●●). (B) Lysozyme(curve) and 28.5% α-helix, 11.1% β-structure, and 60.4% random coil (OOO), calculated as in part A.(C) RNase (curve) and 9.3% α-helix, 32.6% structure, and 58.1% random coil (ΔΔΔ), calculated as in part A

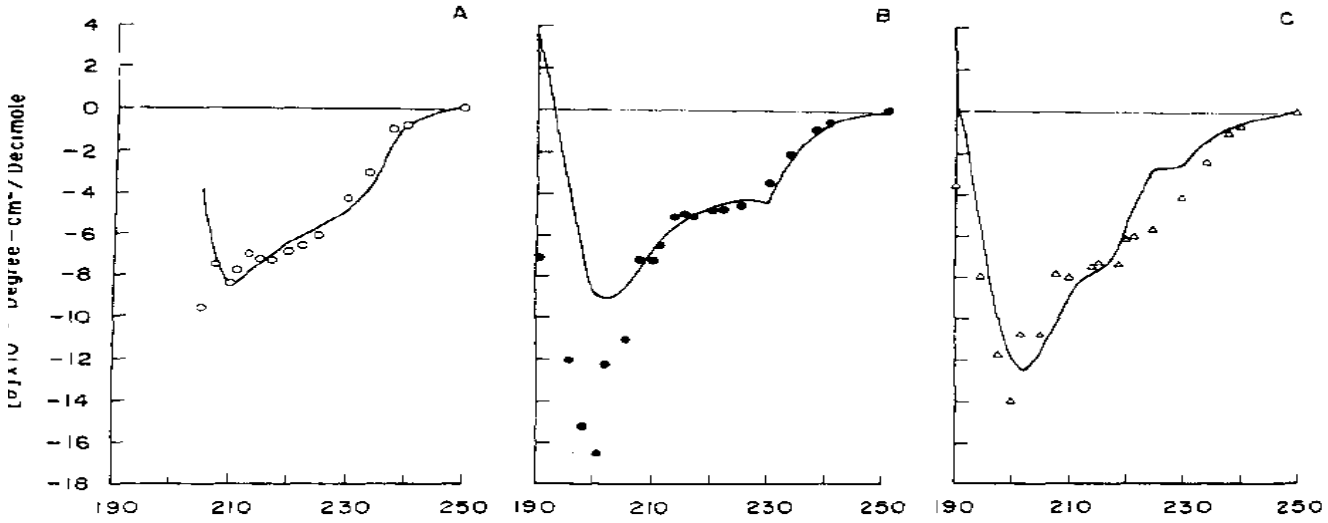


Fig. 18 The circular dichroism of: **(A)** carboxypeptidase **A** (solidcurve) and 13% α– helix, 30.6% β–structure, and 56.4% random coil, calculated from poly-L-lysine spectra in water (OOO). (B) α–Chymotrypsin (solid curve), and 11.8 % α–helix, 22.8 % β– structure and 65.5 % random coil, calculated as in part **A** (• • • **).** (C) α–chymotrypsinogen and 13.8% α-helix, 25.2 β-structure, and 60.9 random coil. (□□□) calculated as in part **A**

Table 1

Agreement of the Secondary Structure of Proteins Determined by CD Using Poly-L-Lysine in the α-helical, β-structure and "random coil" Conformations with their X-ray Structures

	X-ray Structure			Calculated Structure		
Protein	α	β	Other	α	β	Random
myoglobin	65-78	0	22-32	68.3-68.2	4.7-7.9	27.0-23.9
lysozyme	28-42	11	48-62	28.5-29.9	11.1-9.3	60.4-61.0
ribonuclease	6-24	31-36	40-63	9.3-12.0	32.6-43.4	58.1-44.5
carboxypeptidase A	17	25	58	13.0-15.9	30.6-39.9	56.4-44.2
chymotrypsin	14	33	53	11.8-13.4	22.8-31.9	65.5-54.8
chymotrypsinogen	16	32	52	13.8-17.3	25.2-28.9	60.9-53.8

Once the crystallography of may proteins had been obtained several groups (Saxena and Wetlaufer (Saxena, V. P., D. B. Wetlaufer. 1971. A new basis for interpreting the circular dichroic spectra of proteins. Proc Natl Acad Sci U S A 68:969-972.)and Yang and coworkers (Yang, J. T., C. S. Wu, H. M. Martinez. 1986. Calculation of protein conformation from circular dichroism. Methods Enzymol 130:208-269) used the CD spectra of proteins to try to determine the contributions due to the α–helix, random coil, β-pleated sheet and β-turns. Again they used the method of least squares to determine the best estimates of the spectra of the component conformations. They then used these values to calculate the structure of unknown proteins. The CD spectra of the a-helix, α–helix, random coil, β-pleated sheet and β-turns and random coil, derived from the CD spectra of 17 proteins whose structure was derived from X-ray crystallography is shown in Fig. 19.

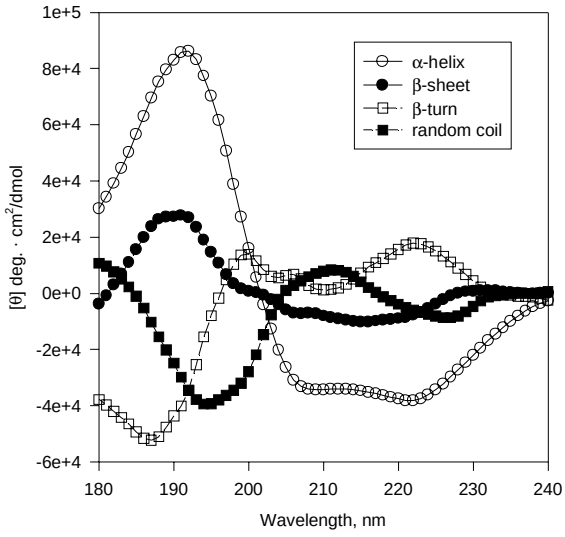


Fig. 19. CD basic spectra extracted from the spectra of 16 protein plus poly-L-glutamate by the method of least squares.

d. Current methods for estimating protein secondary structure of proteins in solution.

In addition to linear regression, more recently singular value decomposition (SVD) and convex constraint analysis (CCA) have been used to extract the basis curves corresponding to helix and β-sheets and turns. The basis curves obtained from deconvoluting a set of protein CD spectra may vary greatly depending on the choice of reference proteins. This occurs because some proteins have unusual CD spectra in the far UV, due to aromatic amino acids, disulfide bridges, or rare conformations. To overcome these difficulties, some methods use selection procedures so that only proteins with spectral characteristics that are similar to those of the protein to be evaluated are used as standards. Methods that use selected reference data include ridge regression, variable selection and neural network procedures.

All methods of analyzing CD spectra assume that the spectrum of a protein can represented by a linear combination of the spectra of its secondary structural elements, plus a noise term which includes the contribution of aromatic chromophores.

$$\theta_\lambda = \Sigma \varepsilon_1 S_l + \text{noise} \qquad (3)$$

where θ_λ is the CD of the protein as a function of wavelength, ε_i is the fraction of each secondary structure, i, and S_i is the ellipticity at each wavelength of each i_{th} secondary structural element. In constrained fits the sum of all the fractional weights, εi, must be equal to 1.

1. Linear Regression.

Linear regression is useful for evaluating the effects of mutations, ligands and solvents on protein conformation.

Standards for Linear Regression

1. Polypeptides with known conformations (Greenfield and Fasman, Brahms and Brahms).
2. Standard spectra extracted from a data base of known proteins using the method of least squares (Saxena and Wetlaufer, Yang and coworkers).
3. Standard spectra extracted from a data base of protein with known conformations using the convex constrain algorithm (Fasman and coworkers).

Non-Constrained Multilinear Regression (MLR)

In non-constrained least squares fit the sum of the conformations is not constrained to =1

$$\theta_\lambda = \Sigma \varepsilon_1 S_l + C$$

Advantages:
1. Don't need to know protein concentration.

2. Invariant standards.

Drawbacks:
1. Poor estimates of β-sheet and turns.
2. Overestimates β-sheets if data below 200 nm is used
3. Useful for analyzing difference spectra

Reference: Brahms, S., and Brahms, J. (1980) Determination of protein secondary structure in solution by vacuum ultraviolet circular dichroism J Mol Biol 138, 149-78.

Constrained Least Squares Analysis

In constrained least squares fits the sum of the contribution of each spectra are set to equal 1.

$$\theta_\lambda = \Sigma \varepsilon_1 S_l \qquad \Sigma \varepsilon = 1$$

Advantages:
1. Invariant data base (useful for direct comparisons).
2. Better fits than non-constrained least squares method.

Disadvantages:
1. No good single standard for β-turns
2. Not as good as more modern methods.

References:
Greenfield, N., and Fasman, G. D. (1969) Computed circular dichroism spectra for the evaluation of protein conformation. Biochemistry 8, 4108-16.
Chen, Y.H., Yang, J.T and Martinez, H.M. (1972) Determination of the secondary structures of proteins by circular dichroism and optical rotatory dispersion. Biochemistry.11:4120-31.
Perczel, A., Park, K., and Fasman, G. D. (1992) Analysis of the circular dichroism spectrum of proteins using the convex constraint algorithm: a practical guide. Anal Biochem 203, 83-93..15

2. Ridge Regression *(Contin)*

Contin, developed by Provencher and Glöckner, fits the CD of unknown proteins by a linear combination of the spectra of a large data base of proteins with known conformations. In this method the contribution of each reference spectrum is kept small unless it contributes to a good agreement between the theoretical best fit curve and the raw data.

Advantage: Relatively good estimate of β-turns.
Drawback: References are different for every fit.

Reference: Provencher, S. W., and Glockner, J. (1981) Estimation of globular protein secondary structure from circular dichroism Biochemistry 20, 33-7.

3. Singular Value Decomposition *(SVD)*

SVD, developed by Hennessey and Johnson, extracts basis curves with unique shapes from a set of spectra of proteins with known structures. The basis curves are each characterized by a mixture of secondary structures, and are then used to analyze the conformation of unknown proteins. The sum of weights is not constrained to equal 1.

Advantage: Best estimate of α-helical content of proteins.
Problem: Poor fit of β-sheet and turns if data are not collected to at least 184 nm.

Reference :Hennessey, J. P., Jr., and Johnson, W. C., Jr. (1981) Information content in the circular dichroism of proteins Biochemistry 20, 1085-94..20

4. Variable Selection *(VARSLC).*

In variable selection an initial large data base of proteins with known spectra and secondary structures is selected. Some of the protein spectra are then eliminated systematically to create new data bases with a smaller number of standards. SVD is performed using all of the reduced data sets and the ones fulfilling selection criteria for a good fit are averaged.

Advantage: Superior Fits
Disadvantage: Very Slow

Reference : Manavalan, P., and Johnson, W. C., Jr. (1987) Variable selection method improves the prediction of protein secondary structure from circular dichroism spectra Anal Biochem 167, 76-85

5. Self Consistent Method *(Selcon)*

The SELCON program of Sreerama and Woody is a modification of VARSLC that is faster to use. With suitable references, the SELCON program can be used to estimate the contribution of the P2 conformation in proteins and the length of helical and beta segments.

Advantage: Good estimates of β-sheet and turns in proteins.
Drawback: Poor fits of spectra of polypeptides with high β-sheet content.

References: Sreerama, N., and Woody, R. W. (1993) A self-consistent method for the analysis of protein secondary structure from circular dichroism. Anal Biochem 209, 32-44.
Sreerama, N., and Woody, R. W. (1994) Protein secondary structure from circular dichroism spectroscopy. Combining variable selection principle and cluster analysis with neural network, ridge regression and self-consistent methods J Mol Biol 242, 497-507.
Sreerama, N., and Woody, R. W. (1994) Poly(pro)II helices in globular proteins: identification and circular dichroic analysis Biochemistry 33, 10022-5.
Sreerama, N., Manning, M. C., Powers, M. E., Zhang, J. X., Goldenberg, D. P. and Woody, R. W. (1999) Tyrosine, phenylalanine, and disulfide contributions to the circular dichroism of

proteins: circular dichroism spectra of wild-type and mutant bovine pancreatic trypsin inhibitor Biochemistry 38, 10814-22.
Sreerama, N., Venyaminov, S. Y., and Woody, R. W. (1999) Estimation of the number of alpha-helical and beta-strand segments in proteins using circular dichroism spectroscopy Protein Sci 8, 370-80.
Sreerama, N., and Woody, R. W. (2000) Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set Anal Biochem 287, 252-60..22

6. Neural Networks *(CDNN and K2D)*

A neural network is an artificial intelligence program which can detect patterns and correlations in data. Two widely used programs are the CDNN program of Böhm et al. and the K2D program of Andrade et al. A neural network is first trained using a set of known proteins so that the input of the CD at each wavelength results in the output of the correct secondary structure. The trained network is then used to analyze unknown proteins. The method works very well and the fits seem to be relatively independent of the wavelength range.

References:
CDNN: Bohm, G., Muhr, R., and Jaenicke, R. (1992) Quantitative analysis of protein far UV circular dichroism spectra by neural networks Protein Eng 5, 191-5.
K2D: Andrade, M. A., Chacon, P., Merelo, J. J., and Moran, F. (1993) Evaluation of secondary structure of proteins from UV circular dichroism spectra using an unsupervised learning neural network Protein Eng 6, 383-90.

7. Convex Constraint Algorithm *(CCA)*

The CCA algorithm deconvolutes a set of spectra into a desired number of basis spectra, which when recombined generate the entire data set with a minimum deviation between the original data set and the reconstructed curves. It is very useful for determining whether there are intermediate states in thermal and perturbant denaturations. The method has also be used to estimate protein conformation, but is poorer than least squares, SVD or neural net analysis. It is very useful, however to determine how many different states contribute to the changes in CD as a function of ligand or temperature.

References:
Perczel, A., Hollosi, M., Tusnady, G., and Fasman, G. D. (1991) Convex constraint analysis: a natural deconvolution of circular dichroism curves of proteins Protein Eng 4, 669-79.
Perczel, A., Park, K., and Fasman, G. D. (1992) Deconvolution of the circular dichroism spectra of proteins: the circulardichroism spectra of the antiparallel beta-sheet in proteins Proteins 13, 57-69.
3. Perczel, A., Park, K., and Fasman, G. D. (1992) Analysis of the circular dichroism spectrum of proteins using the convex constraint algorithm: a practical guide Anal Biochem 203, 83-93.

Comparison of Various Methods of Analyzing CD spectra .

The various methods of estimating secondary structure are compared in table 2 below.

Table 2

Comparisons of Methods of Analyzing Protein Conformation from Circular Dichroism Data*

Computer Program	Standards	Wavelength Range, nm	α-Helix P	α-Helix σ	β-Sheet P	β-Sheet σ	β-Turn P	β-Turn σ
Linear Regression-Unconstrained Fit								
MLR	4 Peptides	240-178	0.91	0.13	0.43	0.21	0.07	0.16
MLR	4 Peptides	240-200	0.92	0.14	0.74	0.16	0.23	0.16
Linear Regression-Constrained Fit								
G&F	Poly-L-lysine	240-208	0.92	0.13	0.61	0.18	ND	ND
LINCOMB	4 Peptides	240-178	0.93	0.11	0.58	0.15	0.61	0.11
LINCOMB	4 Peptides	240-200	0.94	0.11	0.71	0.13	0.53	0.14
LINCOMB	17 Proteins	240-178	0.94	0.09	0.62	0.14	0.21	0.13
LINCOMB	17 Proteins	240-200	0.92	0.10	0.09	0.28	0.52	0.12
Convex Constraint Algorithm								
CCA	17 proteins	260-178	0.96	0.10	0.62	0.18	0.39	0.18
CCA	17 proteins	240-200	0.97	0.10	0.42	0.20	0.52	0.22
Ridge Regression								
CONTIN	17 Proteins	260-178	0.93	0.11	0.56	0.15	0.58	0.08
CONTIN	17 Proteins	240-200	0.95	0.13	0.60	0.15	0.74	0.07
Variable Selection								
VARSLC	17 Proteins	260-178	0.97	0.07	0.81	0.10	0.60	0.07
Variable Selection-Self Consistent Method								
SELCON	17 Proteins	260-178	0.95	0.09	0.84	0.08	0.77	0.05
SELCON	17 Proteins	260-190	0.94	0.09	0.73	0.09	0.84	0.05
SELCON	17 Proteins	240-200	0.93	0.10	0.73	0.11	0.71	0.06
SELCON	33 Proteins	260-178	0.93	0.09	0.91	0.07	0.53	0.09
SELCON	33 Proteins	260-200	0.88	0.12	0.86	0.09	0.46	0.09
Neural Network Analysis								
K2D	19 Proteins	240-200	0.95	0.09	0.77	0.10	ND	ND
CDNN	17 Proteins	260-178	0.95	0.09	0.73	0.11	0.82	0.05
			mean	std	mean	std	mean	std
			0.36	0.27	0.20	0.16	0.22	0.08

*P is the correlation coefficient between the CD estimated and X-ray conformations and σ is the mean square error between the estimated and X-ray conformations. When protein data bases

were used as standards each protein analyzed was excluded from the data set used as the references. When the σ value is higher than the standard deviation (std) of the mean value of each conformation found in the 17 samples which were analyzed, the program does a relatively poor job of analyzing that conformation.

9. Recommendations:

1. For determination of globular protein conformation in solution: SELCON, CDNN and K2D.

2. For determination of polypeptide conformation: CONTIN or LINCOMB with a suitable set of references.

3. For determining the effects of mutations, ligands and perturbants on protein structure: MLR or LINCOMB.

4. For evaluating the number of folding states giving rise to a set of spectra: The CCA algorithm and SVD.

2. The Use of CD for analyzing the conformation of nucleic acids.

The same sort of transitions that affect the transitions of the amide bond, split the aromatic transitions of nucleic acids. Therefore their CD spectra also shows conformational dependence. The CD of polynucleotides in various conformations is shown below:

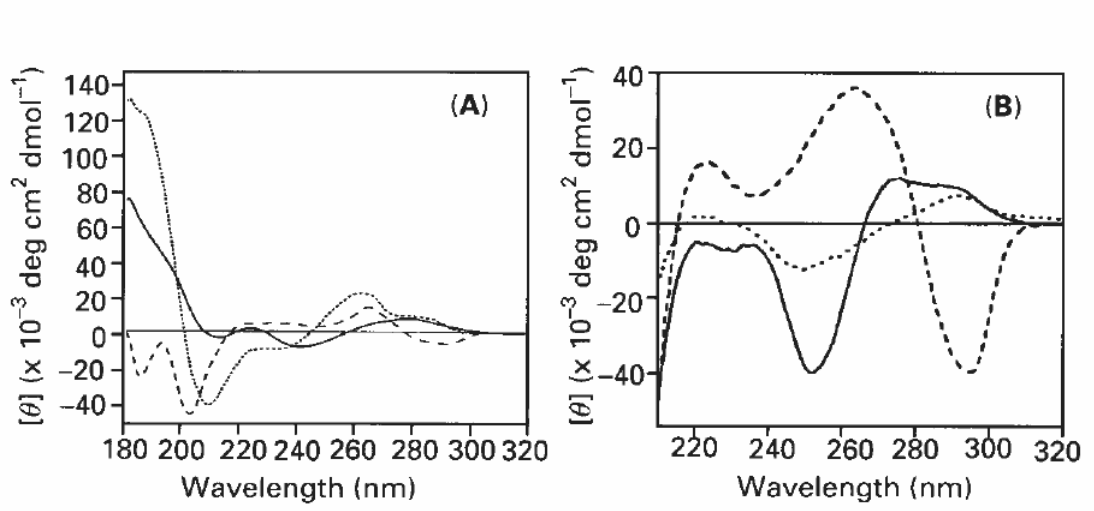


Fig 20. The CD spectra of poly (dAdC)_ poly(dGdT) in the (....) A, (____) B and (---) Z conformations. Redrawn from data in Riazance-Lawrence JH and Johnson WCJr (1992) *Biopolymers* **3**2, 271-276, 1992) B. CD spectra of poly(dGdC)-poly(dGdC) in the (____) B-form, (- - -) Z-form and (....) in a form in which the bases assume the Hoogsteen base pairing conformation. Redrawn from data of Seger-Nolten GMJ, Sijtsema NM and Otto C (1997) *Biochemistry* 36: 13241-13247

B. Determination of conformational changes due to the interactions of asymmetric molecules

When a protein or nucleic acid interacts with another chromophore there are often changes in either the *intrinsic* CD due to changes in secondary structure, the environment of aromatic or prosthetic groups, or the *extrinsic* CD due to changes in the environment of the bound ligand. CD is a quantitative technique and the change in CD is directly proportional to the amount of complex formed. The change can then be used to determine the association or dissociation constant of the complex.

1. Protein Interactions

When a protein interacts with another protein or a ligand there may be a change in the intrinsic CD due to a conformational change, or the ligand might become optically active. These changes can be used to follow binding and estimate binding constants. Figures 21-24 illustrate changes in intrinsic CD spectra due to protein-protein interaction. Figure 21 shows how binding may induce proteins to become more helical, Figure 22 illustrates how an interaction may increase beta structure and Figures 23-24 show how interactions may change the near UV spectrum of the aromatic or heme region of a protein spectrum.

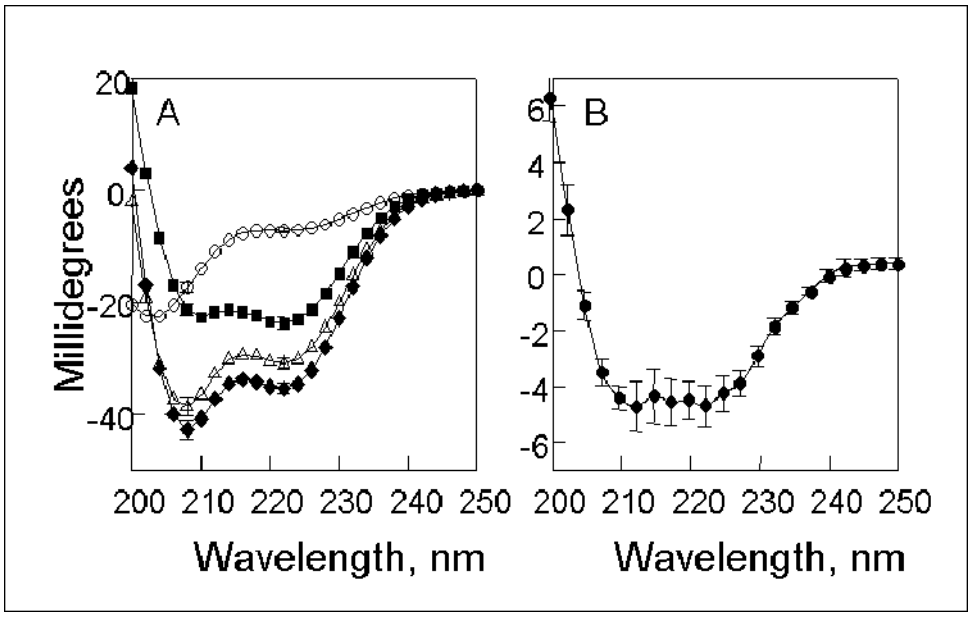


Figure 21. A) Circular dichroism spectra of (o) E-Tmod$_{1\text{-}130}$, (■) AcTM1bZip, (Δ) the sum of the spectra of E-Tmod1-130 and AcTM1bZip and (♦) the spectrum of the mixture E-$_{Tmod1\text{-}130}$ and AcTM1bZip.. (B) Difference between the spectrum of the mixture of the components and the sum of the unmixed components in A. Data are the average and standard error of four measurements.. B Analysis of the difference spectrum using the MLR program. Analyzing the data using the MLR program, with peptides with known conformations as references, suggests that the CD change is 58% due to an increase in α-helix formation, 41% due to an increase in β-sheet and 1% due to an increase in turns. Greenfield, N. J., and Fowler, V. M. (2002) Biophys J 82, 2580-91.

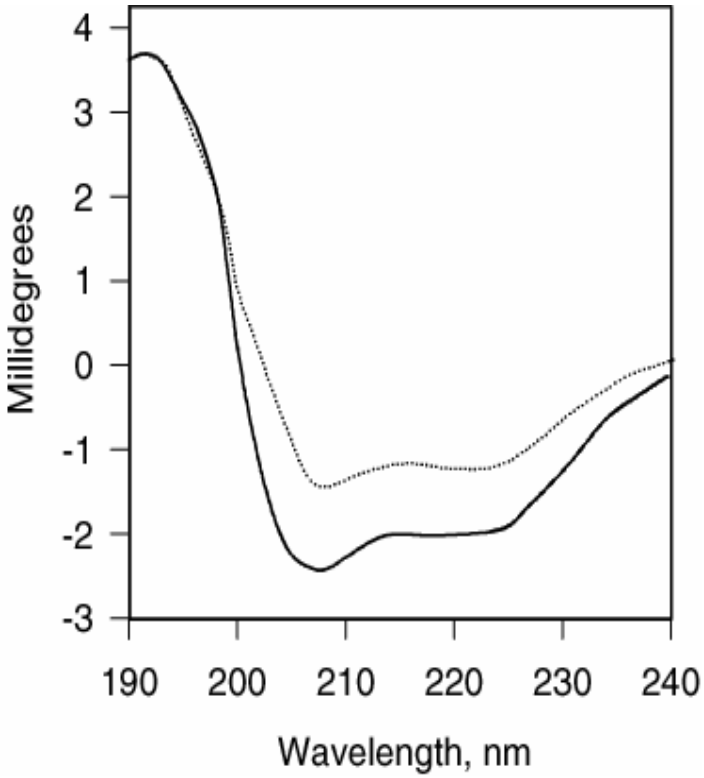


Figure 22. Change in the far UV spectrum of the catalytic Unit of a cAMP- dependent protein kinase upon binding a peptide substrate The conformations of the bound and free protein were analyzed using Contin.

Free Protein:
49 ± 2 % α-helix
20 ± 6% β-sheet
31 ± 7% remainder

Protein + Peptide:
31 ± 2 % α-helix
55 ± 6% β-sheet
15 ± 7% remainder

Redrawn from data in Reed, J., and Kinzel, V. (1984) Biochemistry 23, 1357-62

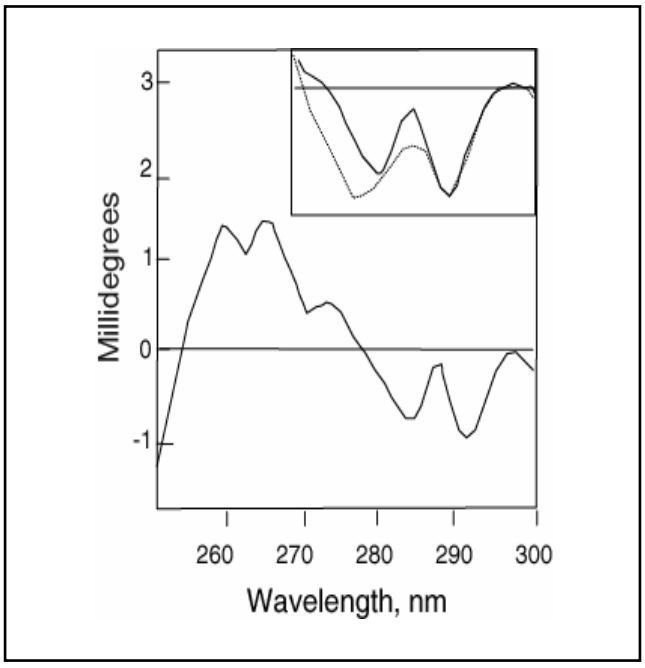


Fig. 23. Far UV CD of the catalytic subunit of adenosine cytclic 5'monophosphate dependent protein kinase. (Inset) CD spectrrum of the catalytic subunit from 275-300 nm (____) enzyme alone; (- - -) enzyme plus 250 mM Kemptide, a synthetic peptide substrate (buffer plus Kemptide base line subtracted) Redrawn from Reed, J., and Kinzel, V. (1984) Biochemistry 23, 1357-62.

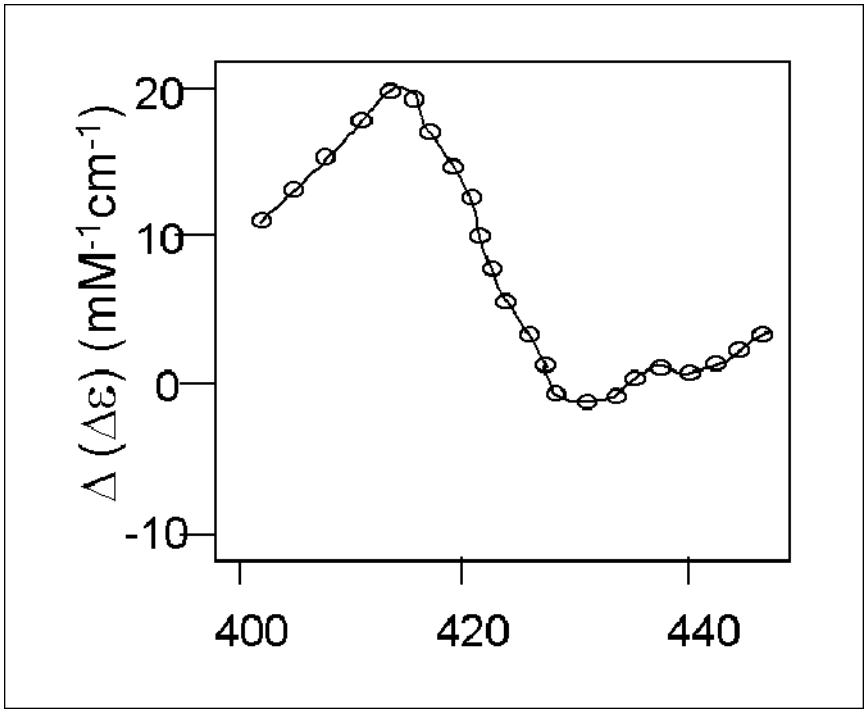


Figure 24. CD difference spectra of native cytochrome C upon binding to cytochrome c oxidase
Michel, B., Proudfoot, A. E., Wallace,C. J., and Bosshard, H. R. (1989) Biochemistry 28, 456-62.

Sometimes when you mix two protein fragments the changes can be very dramatic. Figure 12 illustrates what happens when you mix two fragments of thioredoxin. They are mainly disordered but when mixed form a complex which has a spectrum similar to that of the intact protein. The spectrum of the fragments and reconstituted protein is shown in figure 25, while the secondary structure of the fragments, analyzed by several different analysis programs is given in table 3.

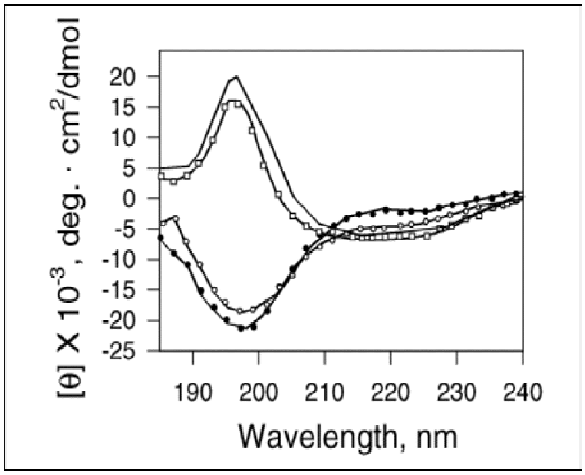


Fig. 25. Spectra of the isolated N- and C-fragments of thioredoxin (N, filled circle; C, empty circle), cleaved (filled triangle) and uncleaved Trx (line). The protein concentration was 20 μM inKPi unless otherwise indicated. All the spectra represent average of five scans corrected by subtracting the spectrum of KPi.
Georgescu, R. E., Braswell, E. H., Zhu, D., and Tasayco, M. L. (1999) Energetics of assembling an artificial heterodimer with an alpha/beta motif:cleaved versus uncleaved Escherichia coli thioredoxin. Biochemistry 38,13355-66.

.

Table 3.

Secondary Structure Analysis of Fragments and Cleaved and Uncleaved Thioredoxin

protein	method	% α-helix	% β–sheet	% turns	others
N	K2D	6.0	29	*a*	65.0
	Contin	4.0	14.0	10.0	72.0
	Mlr (Brahms)*b*	2.5	28.7	10.3	58.5
C	K2D	2.0	16.0	*a*	82.0
	Contin	3.0	11.0	1.0	86.0
	Mlr (Brahms)*b*	0.0	15.7	0.0	84.3
NC	K2D	30.0	20.0	*a*	50.0
	Contin	20.0	40.0	7.0	34.0
	Mlr (Fasman)*b*	22.8	23.7	9.2	21.8; 22.5*c*
	Selcon	24.66	21.53	16.86	22.5*c*
Trx	K2D	27.0	25.0	*a*	48.0
	Contin	18.0	37.0	8.0	37.0
	Mlr (Fasman)*b*	31.8	22.9	7.2 1	6.9; 21.2*c*
	Selcon	28.35	24.8	24.37	21.44
X-ray data*d*		40.0	28.25	14.0	17.75

a Undetermined. *b* Selected Database. *c* Estimate of aromatic and disulfide bond contributions.
d Based on molecule A of the PDB file 2trx.pdb.

C. Determination of the thermodynamics of folding and unfolding of proteins and nucleic acids

CD can be used to analyze the thermodynamics of proteins and nucleic acids. In the simplest case a molecule can exit is two states: folded, F, and folded, U. At any temperature, T, the constant of folding is:

$$K=[F]/[U] = [F]/(1-[F]) \quad (4)$$

The free energy of folding is

$$\Delta G=nRT\ln K \quad (5)$$

The fraction folded at any temperature is α.

$$\alpha=[F]/([F]+[U]) =K/(1+K) \quad (6)$$

$$\alpha = (\theta_{obs} - \theta_U)/(\theta_F - \theta_U) \quad (7)$$

θ_{obs} is the observe ellipticity at any temperature, θ_F is the ellipticity of the fully folded form and θ_U is the ellipticity of the unfolded form.

To fit the change of CD as a function of temperature, T, one uses the Gibbs-Helmholtz equation that describes the folding as a function of temperature. The following equations are fit:

$$\Delta G = \Delta H(1 - T/T_M) - \Delta Cp((T_M - T) + T\ln(T/T_M)) \quad (8)$$

$$K = \exp(-\Delta G/(RT)) \quad (9)$$

$$\alpha = K/(1+K) \quad (10)$$

$$[\theta]_{obs} = \alpha([\theta]_F - [\theta]_U) + [\theta]_U \quad (11)$$

T_M is the temperature where $\alpha = 0.5$ and ΔCp is the change in heat capacity going from the folded to the unfolded state.

Initial values of ΔH, ΔCp, T_M, $[\theta]_F$ and $[\theta]_U$ are estimated, and the values giving the best fit to the raw data is found using non-linear least squares procedures. One can fit the unfolding of homo and heterodimers and trimers using similar equations

Comparison of Circular Dichroism and Scanning Calorimetry

If a protein unfolds in a two-state transition the enthalpy and entropy determined using CD are in excellent agreement with that calculated from scanning calorimetry (see Figure 26 below).

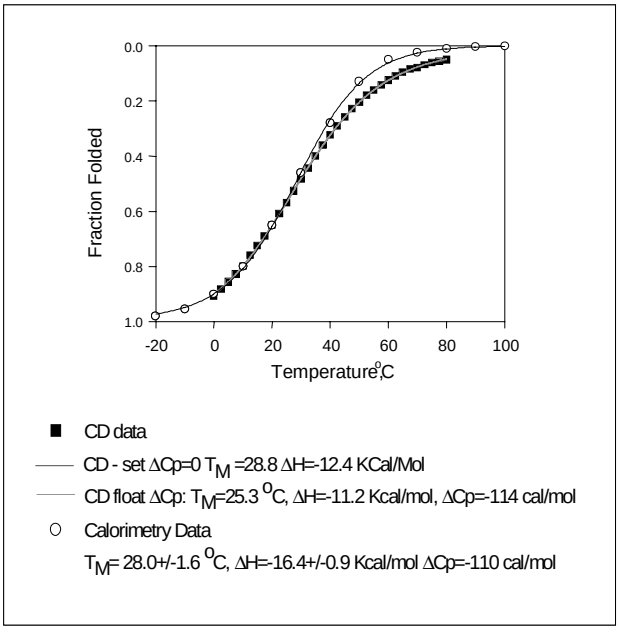


Fig. 26. Comparison of the thermodynamics of unfolding of a helical peptide using differential scanning calorimetry and circular dichroism. The goodness of the fits of the CD curve fitting the ΔCp term as a variable or setting $\Delta Cp = 0$ are virtually identical.
Data from Taylor, J.W., Greenfield, N.J., Wu, B. and Privalov, P.L. A calorimetric study of the folding-unfolding of an alpha-helix with covalently closed N and C-terminal loops. 1999; J. Mol. Biol. 291:965-976

It is difficult to determine the ΔCp of folding from a single CD measurement of the unfolding as a function of temperature because there is really not enough information to determine this value. If the enthalpy of unfolding/folding, ΔH, changes when the pH or salt is changed, however, one can determine ΔCp by plotting ΔH as a function of T_M. The slope of the line yields ΔCp. If one looks at the unfolding of an oligomer, it is possible to directly determine the ΔCp from CD measurements by looking at the unfolding as a function of concentration and temperature and doing a global fit to the Gibbs-Helmholtz equation. Figure 27 illustrates the unfolding of a designed dimeric coil coil as a function of temperature at four different concentration. The data were fit globally to equations (8-11).

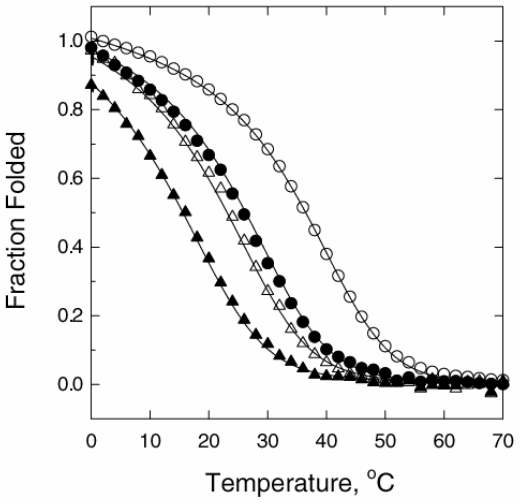


Fig 27. The unfolding of a chimeric coiled coil containing the first 14 residues of rat striated muscle a-tropomyosin and the last 18 residues of the yeast transcription factor, GCN4, as a function of temperature and concentration followed by the changes in ellipticity at 222 nm. (▲)1.5 mM, (D) 7.5 mM, (●) 15 mM, (o), 150 mM. The points were fit globally to the Gibb-Helmholtz equation for a two-state transition between a folded dimer and unfolded monomers. The pre and post transition ellipticities were corrected for small linear changes as a function of temperature. The values for T_M of folding at K=1, the enthalpy, ΔH, of unfolding at ΔG = 0 and the ΔCp of folding were 72.4 °C, -54.3 Kcal/mol and -384 cal/mol·deg, respectively. Data from Greenfield, N.J., Montelione, G.T., Farid, R. and Hitchcock-DeGregori, S.E., The structure of the N-terminus of tropomyosin in a chimeric peptide. Nuclear magnetic resonance structure and circular dichroism studies. Biochemistry 1998; 37: 7834-43.

D. Using Circular Dichroism to Determine Binding Constants

Binding constants can be determined by:

1. Direct titrations
2. Serial dilutions
3. Changes in susceptibility to denaturants.
4. Changes in thermal stability

1. Direct titrations

If one titrates a protein, A, with another protein or a ligand, B, the concentration of product [AB] formed is in equilibrium with free [A] and [B].

The dissociation constant is of the complex is defined as K_D.

$$\mathbf{K_D = [A][B]/[AB] = [AB]/([A_o]-[AB])([B_o]-[AB]} \qquad \mathbf{(12)}$$

A_o and B_o are the initial concentration of A and B and AB is the amount of complex formed. The saturation fraction, s, is defined as $[AB]/[A_o]$.

If there is a change is the intrinsic or extrinsic CD accompanying binding, the change, $[\theta]$, is proportional to the amount of complex formed.

$$\mathbf{[\theta]=\varepsilon\,[AB]} \qquad \mathbf{(13)}$$

where ε is the proportionality constant.

When all of protein A is bound by protein B

$$\mathbf{[\theta]max=\varepsilon[A_o]\ and\ s=[\theta]/[\theta]max=[\theta]/\varepsilon\,[A_o]} \qquad \mathbf{(14)}$$

Case 1. Weak Binding

If one titrates protein A with B and the dissociation constant >> protein concentration then $[B]\sim[B_o]$ and the data can be fit by simple Scatchard or Hill equations.

Scatchard Plot

$$\Delta[\theta]/[A_o][B_o] = k(\varepsilon - (\Delta[\theta]/Ao) \qquad (15)$$

Hill Plot

$$\Delta[\theta]=(\Delta[\theta]_{max}k^h[B_o]^h)/(1+(k^h[B_o]^h))+ C \qquad (16)$$

References: Hill, A. V. (1910) The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves *J. Physiol. (Lond.) 40*, iv-vii.
Scatchard, G. (1949) The attractions of proteins for small molecules and ions *Ann NY Acad Sci 51*, 660-72
Deranleau, D.A., (1969)Theory of the Measurement of Weak Molecular Complexes. I. General Considerations. J. Am. Chem. Soc.; 91:4044-4049.
Deranleau, D.A., (1969)Theory of the measurement of weak molecular complexes. II. Consequences of multiple equilibria. J. Am. Chem. Soc. 91:4050-4054.

Case 2 -Tight Binding

When binding is tight, where $K_D \leq Ao$ one must correct the concentration of protein B for the amount bound.

$$\Delta[\theta]=\Delta[\theta]_{max}((1+(k[B_o]+(k[A_o])/(2k[A_o]) - sqrt((1+(k[B_o]+(k[A_o])/(2k[A_o])^2)-([B_o]/[A_o]))) \qquad (17)$$

In this treatment $k=1/K_D$. The equation can be solved by estimating initial values of k and $[\theta]max$ and finding the best fits using non-linear least squares curve fitting procedures.

Reference: Engel, G. (1974) Estimation of binding parameters of enzyme-ligand complex from fluorometric data by a curve fitting procedure: seryl-tRNA synthetase- tRNA Ser complex. Anal Biochem 61, 184-91

Figure 28 and 29 illustrate how the change in elllipticity as a function of concentration can be used to determine binding constants. Figure 28 shows the interaction of model peptides containing the N-termini of tropomodulin and tropomyosin. As well as binding to each other, these two proteins bind to actin and regulate actin filament dynamics among other functions.

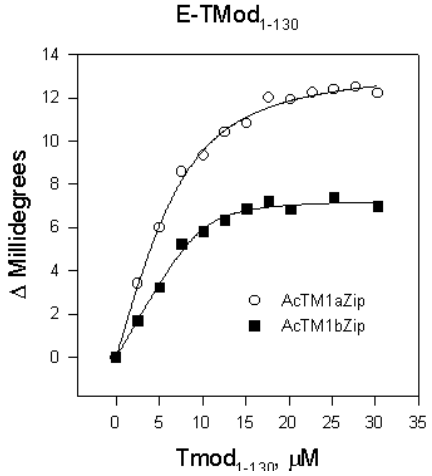


Figure 28. Protein- Protein interactions. Increase in ellipticity at 222 nm, 30° C, when a fragment of tropomodulin 1, $Tmod1_{1-130}$ binds to chimeric peptides containing the N-termini of long (1a) and short (1b) tropomyosins, 10 nmol of the TMZip peptides in 0.5 ml of 100 mM NaCl, 10 mM sodium phosphate, pH 6.5 in 2 mm pathlength cells, were titrated with the Tmod fragment. The ellipticity changes were corrected for the ellipticity of the Tmod peptide alone and for dilution. The data were fitted to equation (17) to give dissociation constants of E-$Tmod1_{1-130}$ for AcTM1aZip and AcTM1bZip of 2.8±1.2 and 0.5±0.3 μM respectively. Data after Greenfield, N. J., and Fowler, V. M. (2002) Tropomyosin requires an intact N-terminal coiled coil to interact with tropomodulin. Biophys J 82, 2580-91.

Figure 29 shows the interaction of dihydrofolate reductase with folates and an anti-cancer drug, Methotrexate, and the titration of the enzyme with dihydrofolate.

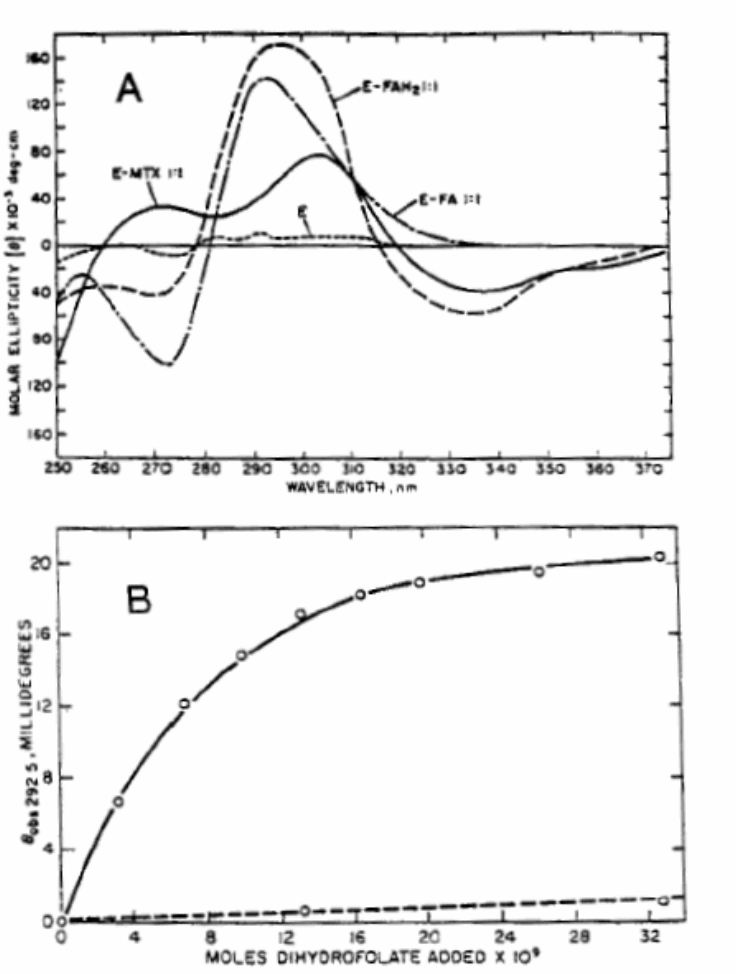


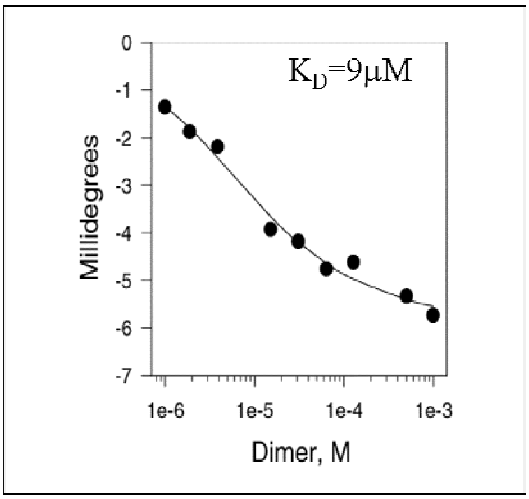
Fig. 29. Titration of an enzyme with a ligand A. The molar CD of dihydrofolate reductase(DHFR) and 1:1 complexes with dihydrofolate (long dashed line) folate (dot-dash) and methrotrexate (solid) The enzyme alone is short dashes. B the increase in ellipticity at 292.5 nm as a function of dihyrdrofolate concentration.. (solid line) DHFR + dihydrofolate, (dash) dihydrofolate alone from Greenfield, N.J., Williams, M.N, Poe M. and Hoogsteen, K. Circular dichroism studies of dihydrofolate reductase from a methotrexate resistant strain of E. Coli. Biochemistry 1972;11: 4706-4711. The dissociation constant was approx. 1 μM.

2. Binding Constants from Serial Dilutions

If binding of two proteins is accompanied by folding, then one can determine the dissociation constant of the complex by determining the molar CD as a function of concentration.

$$[\theta]_{obs}=[\theta]_F + ([\theta]_U - [\theta]_F)(((-k_D + \mathrm{sqrt}(k_D^2 + 4C_o k_D))/2C_o)) \qquad (18)$$

Here [θ]obs is the ellipticity of the mixture of the proteins, at any total concentration, C_O, $[\theta]_F$ is the ellipticity of the fully folded complex and $[\theta]_U$ is the sum of the ellipticity of the fully unfolded individual proteins. Figure 30 illustrates the effect of dilution on the CD of the complex of two fragments of the immunoglobulin binding domain B1 of streptococcal protein G, PGB1(1-40) and PGB1(41-56) The data were fitted using equation (18)



Fig, 30 Association of two complementary fragments of the immunoglobulin binding domain B1 of streptococcal protein G, PGB1(1-40) and PGB1(41-56) evaluated by molecular ellipticityat 222 nm. Equimolar concentrations of the fragments were dissolved in 50 mM phosphate buffer (pH 5.5) at 298 K. The closed circles and solid line show observed data and best curve fitted to the equation (18) to determine the apparent Kd. Honda, S., Kobayashi, N., Munekata, E., and Uedaira, H. (1999) Fragment reconstitution of a small protein: folding energetics of the reconstituted immunoglobulin binding domain B1 of streptococcal protein G. Biochemistry38, 1203-13..

3. Chemical Denaturation

Binding constants can also be determine from the change in stability to chemical denaturation when a macromolecule interacts with another macromolecule or ligand.
When A binds to B,

$$\mathbf{K_D = [A][B]/[AB] = Ptf_u2/(2(1-f_u))} \qquad \mathbf{(19)}$$

$[P_t]$ is the total concentration of [A]+[B]+[2AB] and f_u is the fraction of the unfolded/dissociated complex

$$\mathbf{\Delta G_u = - nRTlnK_D} \qquad \mathbf{(20)}$$

In the present of a denaturant, D,

$$\mathbf{\Delta G_D = \Delta G_o + m[D]} \qquad \mathbf{(21)}$$

The unfolded fraction can be calculated from the CD

$$\mathbf{f_u = 1-([\theta]_{obs}-[\theta]_u)/([\theta]_f-[\theta]_u)} \qquad \mathbf{(22)}$$

K_D and ΔG_D can be calculate at any conc. of D using eq. (19) and (20). Plotting ΔG_D againts [D], the y intercept $=\Delta G_o$ and $K_D = -\exp(\Delta G_o)/RT$

Figure 31 illustrates the use of chemical denaturation to determine the association constant of two chains of a designed coiled coil.

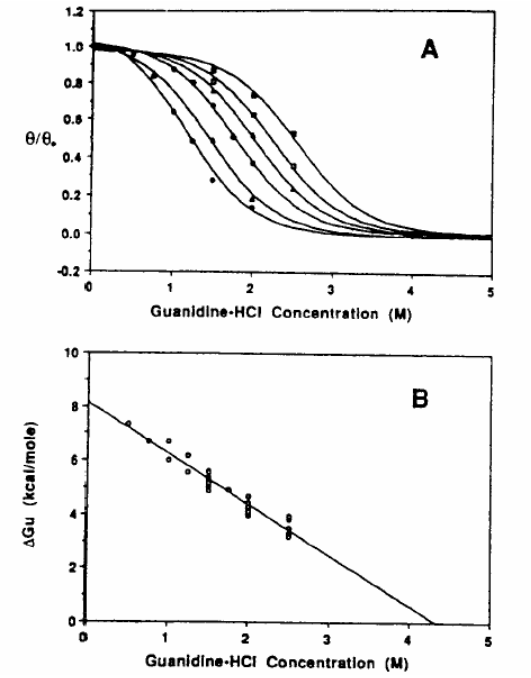


Fig. 31. A. Concentration dependence of the guanidine-HCl denaturation of a coiled-coil model peptide. Concentrations of the peptide in μM were: •, 48;Δ, 131;o, 213;▲, 592; □, 888; ■, 1345. B. Plot of the linear dependence of ΔG of unfolding on the concentration of guanidine-HCl at different concs. of the peptide, extrapolated to zero to determine the free energy of unfolding in the absence of denaturant. The correlation coefficient for the regression line was 0.94.
N.E. Zhou, C.M. Kay and R.S. Hodges, *J. Biol. Chem.* 267 (1992) 2664.

4. The Use of Changes in Thermal Stability to determine binding constants.

I

If two proteins bind to each other only when they are folded and the protein complex unfolds cooperatively and reversibly one can use the change in stability upon binding to measure the binding constant.

In the simplest case both proteins (eg X and Y) are monomers.

The constant of folding of protein X is $k_{FX}=[X_F]/[X_U]$.

The constant of folding of protein Y is $k_{FY}=[Y_F]/[Y_u]$.

The constant of folding of the dimer is $k_{FXY}=[XY_F]/[X_U][Y_U]$.

The association constant $K_A=[XY_F]/[X_F][Y_F]$.

Making the substitutions: $K_A=k_{FXY}/k_{FX}k_{FY}$

To determine the association constant:

1. Determine the ΔH and ΔS of folding of the individual monomers X and Y and the stability of the dimer XY by following their ellipticity as a function of temperature or denaturant.

2. Calculate the free energy of folding at any temperature using:
$\Delta G=\Delta H-T\Delta S$.

3. Determine the folding constant from the relationship:
$kF = \exp(-\Delta G/RT)$.

4. Divide the kF of folding of the dimer by the kF values of folding of the two monomers to get the binding constant for dimer formation.

4. Protein-Ligand Binding Constants from Unfolding Studies

If a ligand binds to a protein only when it is folded, the increase in stability of the complex compared to the free protein can be used to estimate the binding constant. If the changes in enthalpy of folding upon binding are small, the binding constant can be estimated using the equation:

$$\Delta Tm=(TT_oR/\Delta H)\ln(1+K[L]) \qquad (23)$$

where ΔTm is the change in the midpoint of the thermal denaturation curve, T and T_o are the Tm values in the present and absence of ligand, ΔH is the enthalpy of unfolding of the protein, K is the equilibrium for the binding of the ligand to the folded protein and [L] is the free concentration of ligand.
References: Pace, C. N., T. McGrath. 1980. Substrate stabilization of lysozyme to thermal and guanidine hydrochloride denaturation. J Biol Chem 255:3862-3865.,1980.
Schellman, J. A. 1975. Macromolecular binding. Biopolymers 14:999-1018
Schellman, J. A. 1976. The effect of binding on the melting temperature of biopolymers. Biopolymers 15:999-1000

The use of changes in stability as a function of temperature to measure protein-protein interactions is given in figure 32.

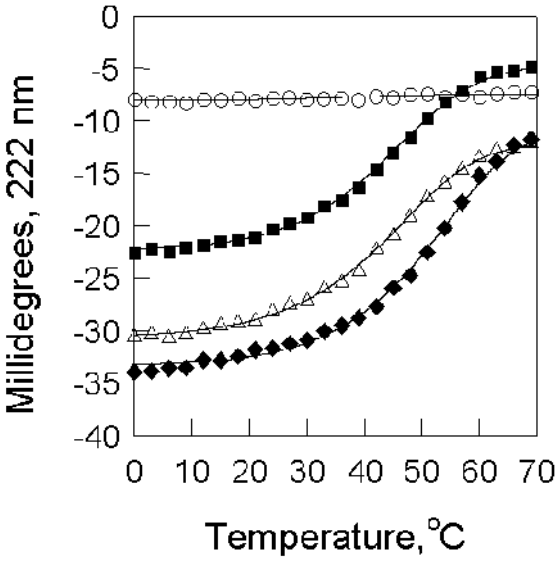
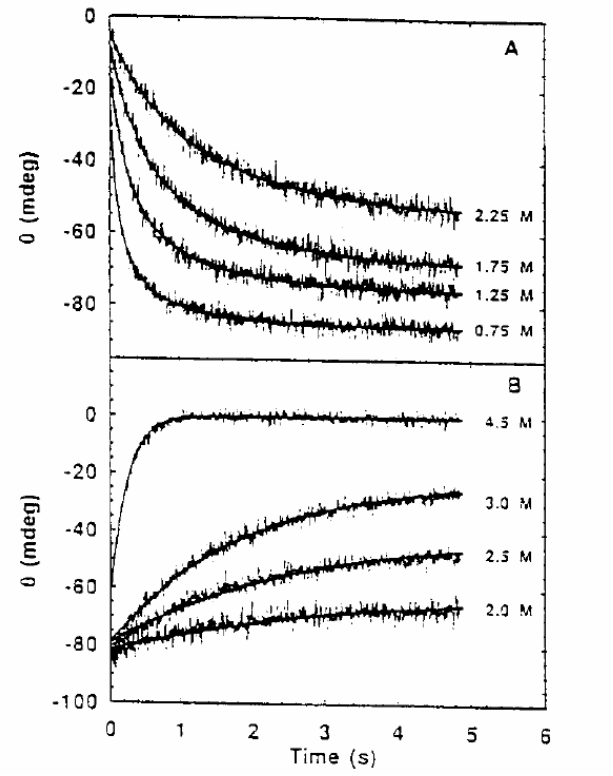


Fig. 32. Thermal denaturation curves for (o) E-Tmod$_{1\text{-}130}$, (■) AcTM1bZip, (Δ) the sum of the spectra of E-Tmod$_{1\text{-}130}$ and AcTM1bZip and (♦) the spectrum of the mixture E-Tmod$_{1\text{-}130}$ andAcTM1bZip. The lines are the best fits of the data assuming two-state transitions between folded and unfolded conformations.
The data were used to estimate dissociation constants for the complex of the peptides of 0.23 ± 0.15 μM.
Greenfield, N. J., and Fowler, V. M. (2002) Biophys J 82, 2580-91.

E. Changes in circular dichrosm as a function of time can be used to estimate the kinetics of folding and unfolding..

Figure 33 illustrates the unfolding and refolding of the leucine zipper region of GCN4 as a function of time.



Fig, 33 Kinetics of refolding (A) and unfolding (B) traces at various final Gnd-HCl concs. and constant peptide cons. of the GCN4-P1 monitored by the change in ellipticity at 222 nm at pH 7.0 and 5 o C. Solid lines represent global fits to the model of a native dimer dissociating to give two unfolded chains. J.A. Zitzewitz, O. Bilsel, J. Luo, B.E. Jones and C.R. Matthews, *Biochemistr*y, 34 (1995) 12812..55

Summary

1. CD can be used to determine the secondary structure of proteins and nucleic acids.

2. CD can be used to determine conformational changes accompanying folding, mutations, protein-ligand, protein-nucleic acid and protein-protein interactions

3. Binding constants for the interacting proteins can be estimated from CD data using:

   A. Direct Titrations
   B. Serial Dilutions
   C. Chemical Denaturation
   D. Thermal Denaturation

4. CD can be used to follow kinetics of folding, unfolding and binding.

Review Articles:

1. Greenfield, N.J. and Fasman, G.D. Polypeptides, characterization. Encyclopedia of polymer science and technology. J. Wiley and Sons, N.Y. 1971;15:410 429.
2. Adler, A.J., Greenfield, N.J. and Fasman, G.D. Circular dichroism and optical rotatory dispersion of proteins and polypeptides. Methods. Enzymol.1973;27 part D:675 735.
3. Greenfield, N.J. Enzyme ligand complexes: spectroscopic studies. CRC Crit. Rev. Biochem. 1975;3:71-110.
4. Greenfield, N.J. Methods to estimate the conformation of proteins and polypeptides from circular dichroism data, a review. Analytical Biochemistry 1996;235:1-10.
5. Greenfield, N.J. Applications of circular dichroism in protein and peptide analysis. Trends in Analytical Chemistry. 1999;18:236-244.
6. Greenfield, N.J. Biomacromolecular applications of circular dichroism and ORD, in Encylopeidea of Spectroscopy and Spectrometry, Lindon, J.C., Tranter, G.E. and Holmes, J.L, eds. Academic Press, 2000, pp 117-130.
7. Circular dichroism studies of protein-protein interactions in Protein Protein Interactions: Methods and Protocols, H. Fu, ed. Methods Mol Biol. 2004;261:55-78.
8. Greenfield, N.J. Analysis of Circular Dichroism Data, in Methods in Enzymology, Numerical Computer Methods, Part D., Johnson, M.L. and Brand, L, eds. Academic Press Methods Enzymol. 2004;383:282-317.