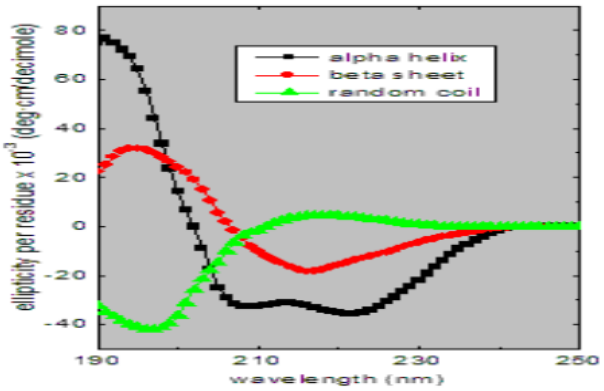
PROTEIN ENGINEERING

LECTURE 03: CIRCULAR DICHROISM (CD) SPECTROSCOPY

CIRCULAR DICHORISM SPECTROSCOPY

Circular dichroism (CD) spectroscopy measures differences in the absorption of left-handed polarized light versus right-handed polarized light which arise due to structural asymmetry. The absence of regular structure results in zero CD intensity, while an ordered structure results in a spectrum which can contain both positive and negative signals.



Circular dichroism spectroscopy is particularly good for:

- determining whether a protein is folded, and if so characterizing its secondary structure, tertiary structure, and the structural family to which it belongs
- comparing the structures of a protein obtained from different sources (*e.g.* species or expression systems) or comparing structures for different mutants of the same protein
- demonstrating comparability of solution conformation and/or thermal stability after changes in manufacturing processes or formulation
- studying the conformational stability of a protein under stress -- thermal stability, pH stability, and stability to denaturants -- and how this stability is altered by buffer composition or addition of stabilizers and excipients
  - CD is excellent for finding solvent conditions that increase the melting temperature and/or the reversibility of thermal unfolding, conditions which generally enhance shelf life
- determining whether protein-protein or protein-ligand interactions alter the conformation of protein.
  - If there are any conformational changes, this will result in a spectrum which will differ from the sum of the individual components. Small conformational changes have been seen, for example, upon formation of several different receptor/ligand complexes.

PROTEIN ENGINEERING

Determination of Protein Secondary Structure by Circular Dichroism:

Secondary structure can be determined by CD spectroscopy in the "far-UV" spectral region (190-250 nm). At these wavelengths the chromophore is the peptide bond, and the signal arises when it is located in a regular, folded environment.

Alpha-helix, beta-sheet, and random coil structures each give rise to a characteristic shape and magnitude of CD spectrum. This is illustrated by the graph to the right, which shows spectra for poly-lysine in these three different conformations. The approximate fraction of each secondary structure type that is present in any protein can thus be determined by analyzing its far-UV CD spectrum as a sum of fractional multiples of such reference spectra for each structural type.

Like all spectroscopic techniques, the CD signal reflects an average of the entire molecular population. Thus, while CD can determine that a protein contains about 50% alpha-helix, it cannot determine which specific residues are involved in the alpha-helical portion.

Far-UV CD spectra require 20 to 200 µl of solution containing 1 mg/ml to 50 µg/ml protein, in any buffer which does not have a high absorbance in this region of the spectrum. (High concentrations of DTT, histidine, or imidazole, for example, cannot be used in the far-UV region.) Note that for many formulated protein samples the absorbance due to the excipients prevents collecting spectra below 200 nm (and even 200 nm is often not possible). When that is true the accuracy reliability of secondary structure calculations (the actual percentages of different structures) is compromised, but the validity of spectral comparisons is not.

Information About Protein Tertiary Structure from Circular Dichroism:

The CD spectrum of a protein in the "near-UV" spectral region (250-350 nm) can be sensitive to certain aspects of tertiary structure. At these wavelengths the chromophores are the aromatic amino acids and disulfide bonds, and the CD signals they produce are sensitive to the overall tertiary structure of the protein.

Signals in the region from 250-270 nm are attributable to phenylalanine residues, signals from 270-290 nm are attributable to tyrosine, and those from 280-300 nm are attributable to tryptophan. Disulfide bonds give rise to broad weak signals throughout the near-UV spectrum.

If a protein retains secondary structure but no defined three-dimensional structure (*e.g.* an incorrectly folded or "molten-globule" structure), the signals in the near-UV region will be nearly zero. On the other hand, the presence of significant near-UV signals is a good indication that the protein is folded into a well-defined structure.

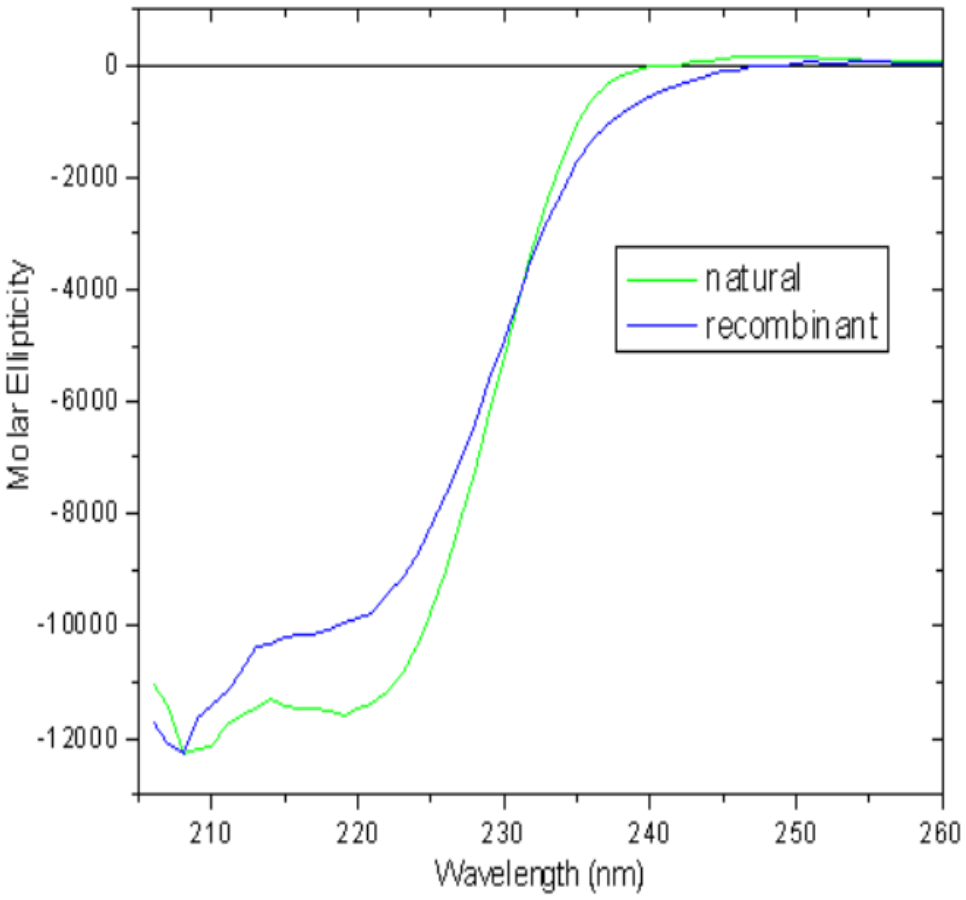
The near-UV CD spectrum can be sensitive to small changes in tertiary structure due to protein-protein interactions and/or changes in solvent conditions.

The signal strength in the near-UV CD region is much weaker than that in the far-UV CD region. Near-UV CD spectra require about 1 ml of protein solution with an OD at 280 nm of 0.5 to 1 (which corresponds to 0.25 to 2 mg/ml for most proteins).

Demonstrating Comparability of Conformation

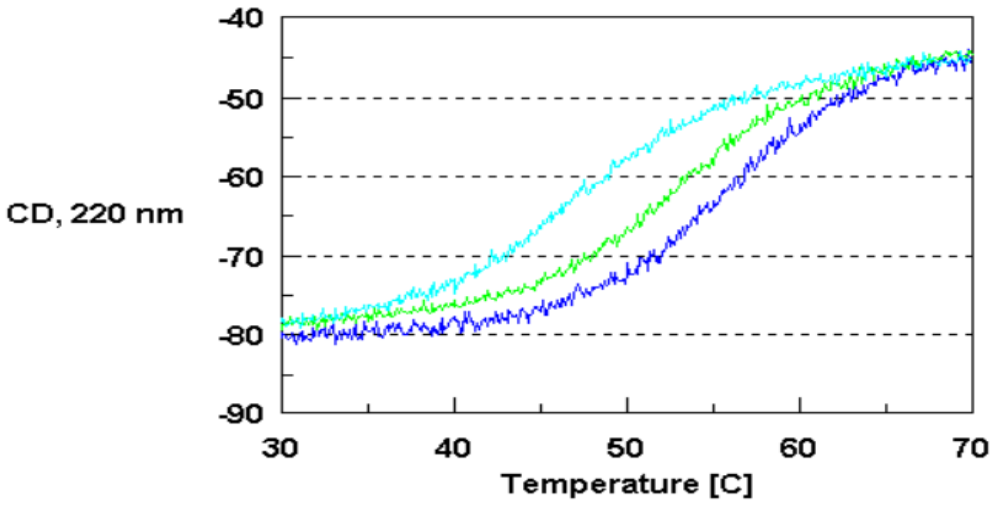
Often it is necessary to demonstrate that different lots of a protein have equivalent conformations, for example after a scale-up in the purification process or to qualify a new manufacturing site, and CD can be a good tool for this.

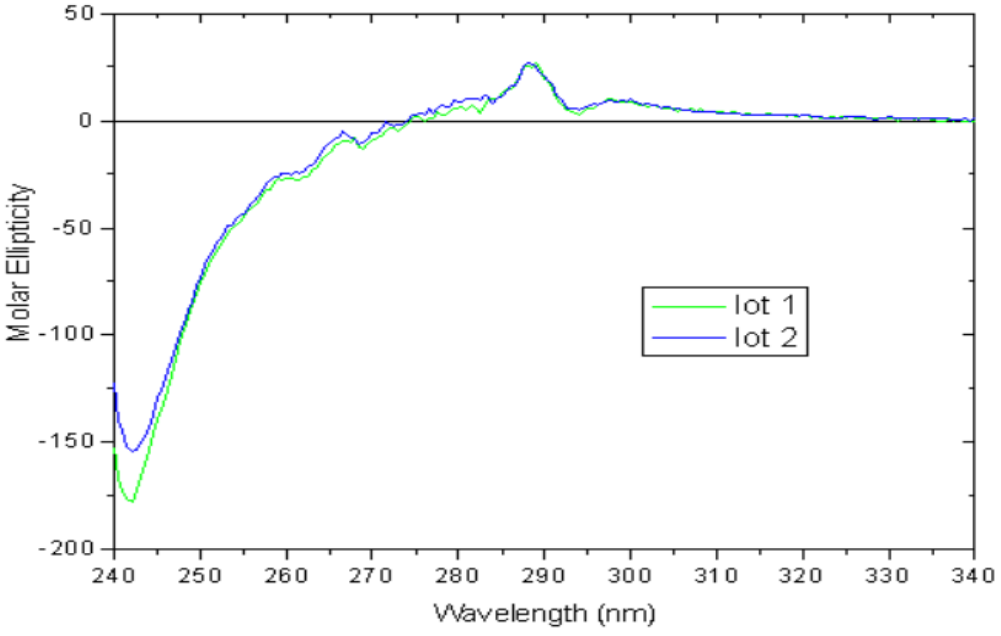
The data below show a case where the far-UV spectra show that the recombinant form of an enzyme clearly does <u>not</u> have the same secondary structure as the natural protein (*i.e.* the recombinant protein is not properly folded).



Such cases of significant differences in secondary structure are, however, unusual. More typically subtle differences in conformation do not produce a detectable difference in far-UV CD, but may produce a difference in near-UV CD. One such example, for different lots of a monoclonal antibody, is shown below. This small but reproducible difference at ~240 nm correlates with differences in the stability of different lots of this antibody.

Thermal Stability by Circular Dichroism





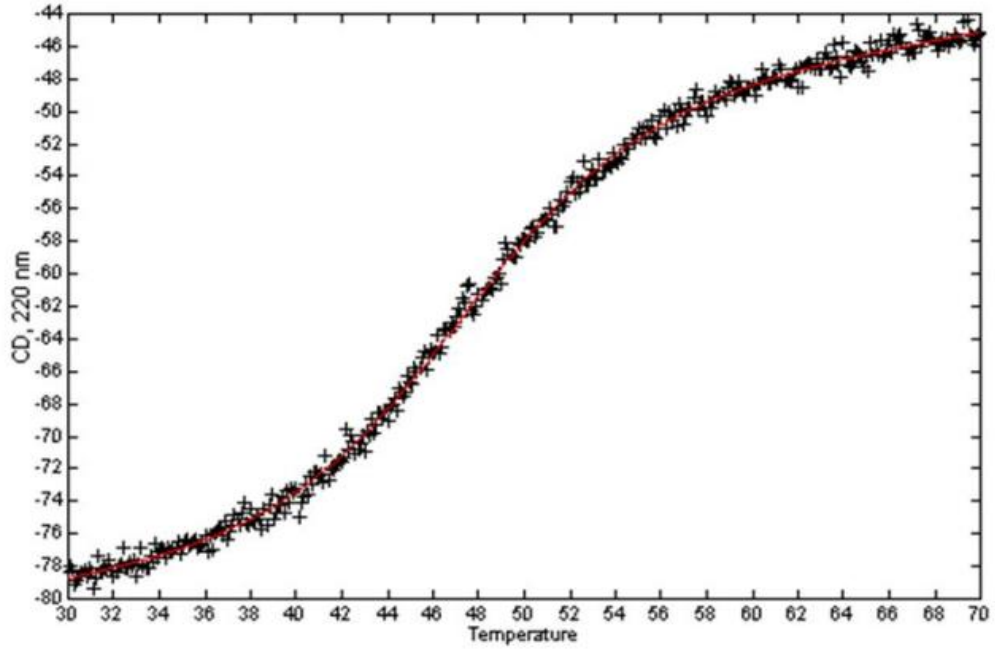
PROTEIN ENGINEERING

Thermal stability is assessed using CD by following changes in the spectrum with increasing temperature. In some cases the entire spectrum in the far- or near-UV CD region can be followed at a number of temperatures. Alternatively, a single wavelength can be chosen which monitors some specific feature of the protein structure, and the signal at that wavelength is then recorded continuously as the temperature is raised. CD is often used to assess the degree to which solution pH, buffers, and additives such as sugars, amino acids or salts alter the thermal stability.

This graph illustrates thermal scans done in our lab for the same recombinant protein in 3 different buffers. While unfolding is completely reversible under all these conditions, clearly there are quite significant differences in thermal stability.

Many proteins aggregate or precipitate quickly after they are unfolded ("melted"), making unfolding irreversible. The reversibility of the unfolding reaction can be assessed by cooling the sample and then heating again to see if the unfolding reaction is duplicated. Finding solvent conditions that make unfolding reversible may be actually be more important for long-term stability (shelf life) than raising the melting temperature.

If (and only if) the melting is fully reversible, the melting temperature is directly related to conformational stability, and the thermodynamics of protein folding can be extracted from the data. The fact that thermal unfolding can generally be measured by CD at much lower concentrations than by DSC increases the probability of reversible reactions and of thermodynamically interpretable data.



This graph illustrates a detailed analysis of one of the data sets shown above, using custom software developed in our lab. The data (+) were fitted to a simple thermodynamic unfolding model (solid line). The fit returns the melting temperature (midpoint of the transition) as 47.3 +/- 0.1 °C. The width of the transition region is related to the enthalpy of unfolding, □*H*, which the fit returns as 52 +/- 2 kcal/mol. Fitting the data also allows a more reproducible measurement of the onset of unfolding, a temperature which is often more relevant for formulation and shelf-life considerations than the midpoint. The onset (defined as the temperature at which 5% of the protein is unfolded) occurs at 36.1 +/- 0.3 °C in this case.

If the protein precipitates or aggregates as it is unfolded, the melting reaction will be irreversible, and the melting temperature will reflect the kinetics of aggregation and the solubility of the unfolded form of the molecule as well as the intrinsic conformational stability.

The cooperativity of the unfolding reaction is measured qualitatively by the width and shape of the unfolding transition. A highly cooperative unfolding reaction indicates that the protein existed initially as a compact, well-folded structure, while a very gradual, non-cooperative melting reaction indicates that the protein existed initially as a very flexible, partially unfolded protein or as a heterogeneous population of folded structures.

Melting of Secondary Structure

Changes in secondary structure, monitored in the far-UV CD region, can be determined with as little as 50 µg of protein, at concentrations of 0.2 mg/ml. By following changes over the entire far-UV CD region we can determine whether at high temperatures the protein is losing all of its secondary structure, loses only a portion of its secondary structure, or simply undergoes conformational change involving a change in secondary structure. Occasionally the unfolded form of a protein will possess a defined but totally different secondary structure than the native form (*e.g.*, TNF-alpha contains beta-sheet when folded, but alpha-helix when melted, and many proteins form amyloid-like aggregates following a transition from alpha-helix to beta-strand).

Melting of Tertiary Structure

Changes in tertiary structure can be followed by monitoring changes in the near-UV CD region. Due to the weaker signal in this region this requires 1-3 mg of protein. Such studies will reveal whether the melting of a protein occurs in a single step (with concurrent loss of both secondary and tertiary structure), or in a two-step reaction.

Melting of Protein Complexes

The effect of forming a protein-protein complex (*e.g.* ligand/receptor or antigen-antibody) on the thermal stability of the individual proteins in the complex can also be determined. This works best if the individual proteins have CD spectra which are quite different from each

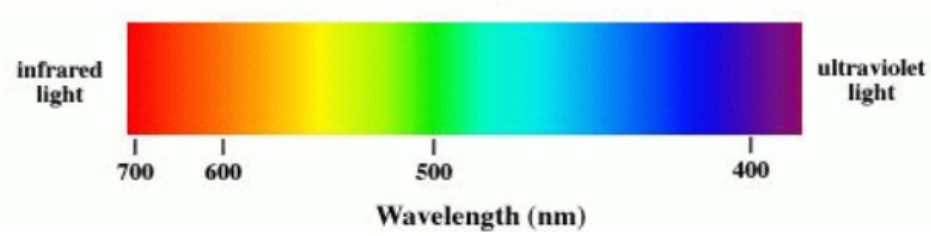
other, such that changes at specific wavelengths can be monitored to follow changes in the corresponding protein. In such cases it is possible to determine whether there is an increase in stability of one or both of the proteins following complex formation.

Infrared spectroscopy of proteins

During the last years the use of Fourier Transform Infrared spectroscopy (FTIR) to determine the structure of biologicalmacromolecules has dramatically expanded. The complete three-dimensional structure of a protein at high resolution can be determined by X-ray crystallography. This technique requires the molecule to form a well ordered crystal which is not possible for all proteins. An alternative to X-ray crystallography is multidimensional nuclear magnetic resonance (NMR) spectroscopy. Using NMR spectroscopy structures of the proteins can be determined in solution. The interpretation of the NMR spectra of large proteins is very complex, so its present application is limited to small proteins (~15-25 kDa). These limitations have led to the development of alternative methods that are not able to generate structures at atomic resolution but provide also structural information on proteins (especially on secondary structure). These methods include circular dichroism (CD) and vibrational (infrared and RAMAN) spectroscopy. The new technique of FTIR spectroscopy requires only small amounts of proteins (1mM) in a variety of environments. Therefore, high quality spectra can be obtained relatively easy without problems of background fluorescence, light scattering and problems related to the size of the proteins. The omnipresent water absorption can be subtracted by mathematical approaches. Methods are now available that can separate subcomponents that overlap in the spectra of proteins. These facts have made practical biological systems amenable to studies by FTIR spectroscopy.

Basic principles of infrared (IR) absorption

We will focus on very few aspects here, because many textbooks present excellent descriptions of the basis of IR spectroscopy (see for example *Campbell &Dwek, in Biological Spectroscoy, Benjamin Cummings, Menlo Park, CA 1984* and *Brey, Physical Chemistry and its Biological Applications, Academic Press, New York, 1984, p.133*). IR spectroscopy is the measurement of the wavelength and intensity of the absorption of infrared light by a sample. Infrared light is energetic enough to excite molecular vibrations to higher energy levels.



Electromagnetic spectrum

frequencyrange (Hz)	wavelength range	type of radiation	type of transition
10^{20} - 10^{24}	10^{-12} - 10^{-16} m	gamma rays	nuclear
10^{17} - 10^{20}	1 nm - 1 pm	x-rays	inner electrons
10^{15} - 10^{17}	400 - 1 nm	ultraviolet light	outer electrons
$4.3x10^{14}$ - $7.5x10^{14}$	700 - 400 nm	visible light	outer electrons
10^{12}-10^{14}	2.5 um - 700 nm	infrared light	vibrations
10^{8} - 10^{12}	1 mm - 2.5 um	microwaves	rotations
10^{0} - 10^{8}	10^{8} - 1 m	radio waves	spin flips

The infrared spectra usually have sharp features that are characteristic of specific types of molecular vibrations, making the spectra useful for sample identification.

Table of characteristic IR bands

X-H vibrations	***bond***	***wavenumbers (cm^{-1})***
hydroxyl	O-H	3610-3640
amines	N-H	3300-3500
aromatic rings	C-H	3000-3100
alkenes	C-H	3020-3080
alkanes	C-H	2850-2960
triple bonds		2500-1900
double bonds		1900-1500
deformation/heavy atoms		1500-

For a molecule of N atoms, 3N-6 fundamental vibrations (or normal modes) exist (3N-5 if the molecule is linear). Therefore, for the linear CO2 molecule 4 normal modes have to be expected.

Normal modes for CO_2

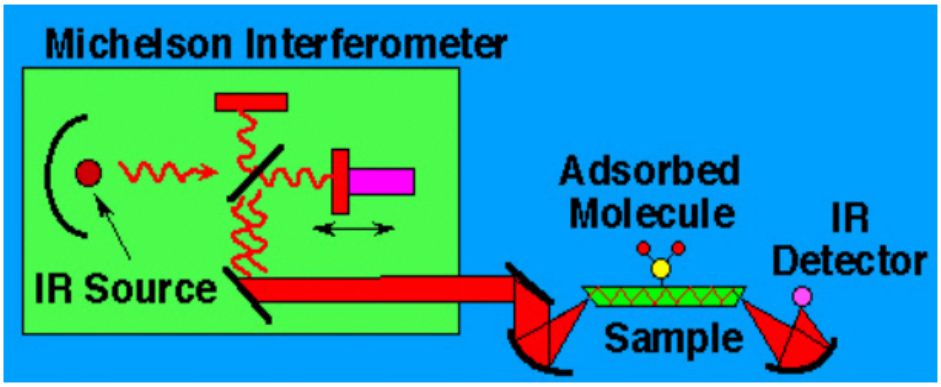
		cm^{-1}	*IR*	*RAMAN*
stretching (sym.)	-> <- O==C==O	1340	-	+
stretching (asym.)	-> <- <- O==C==O	2349	+	-
deformation	/| /| O==C==O |	667	+	-
deformation	+ - + O==C==O	667	+	-

Fourier Transform Infrared (FTIR) spectroscopy

To use the Fourier Transform Infrared Spectroscopy, a continuum source of light (such as a Nernst Globar) is used to produce light over a broad range of infrared wavelengths. Light coming from this continuum source is split into two paths using a half-silvered mirror; this light is then reflected from two mirrors back onto the beamsplitter, where it is recombined. One of these mirrors is fixed, and the second is movable. If the distance from the beamsplitter to the fixed mirror is not exactly the same as the distance from the beamsplitter to the second mirror, then when the two beams are recombined, there will be a small difference in the phase of the light between these two paths. Because of the "superposition principle" constructive and destructive interference exist for different wavelengths depending of the relative distances of the two mirrors from the beamsplitter.

It can be shown that if the intensity of light is measured and plotted as a function of the position of the movable mirror, the resultant graph is the Fourier Transform of the intensity of light as a function of wavenumber . In FTIR spectroscopy , the light is directed onto the

sample of interest, and the intensity is measured using an infrared detector. The intensity of light striking the detector is measured as a function of the mirror position, and this is then Fourier-transformed to produce a plot of intensity vs. wavenumber. As radiation source a Michelson Interferometer is used (see the drawing below).

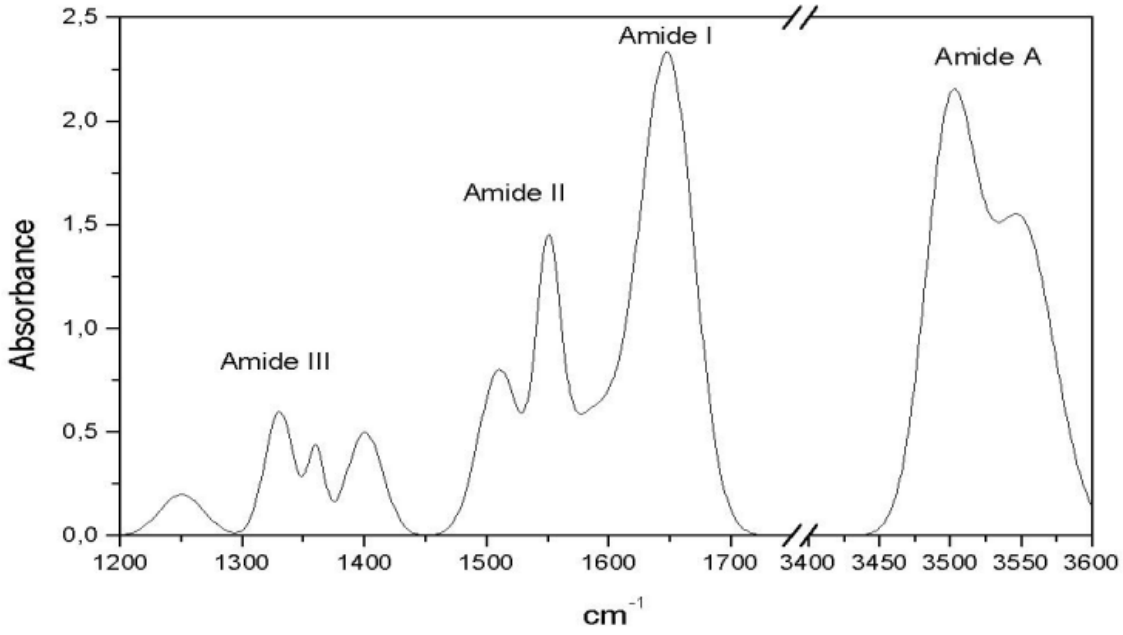


It is necessary to increase the sensitivity somehow, because the absorption due to one monolayer of molecules typically results in a change in intensity of only about one part in 10^5. For semiconductors, one way of increasing the sensitivity is to use multiple internal reflection. In this technique, the edges of the sample are polished, and the light is sent in at an angle. The light bounces around inside the sample, making about 30-50 bounces. This increases the sensitivity by about a factor of 30-50, making it possible to measure the absorption of less than one monolayer of molecules on a surface.

Band assignments

Amide vibrations

The peptide group, the structural repeat unit of proteins, gives up to 9 characteristic bands named amide A, B, I, II ... VII. The amide A band (about 3500 cm^{-1}) and amide B (about 3100 cm^{-1}) originate from a Fermi resonance between the first overtone of amide II and and the N-H stretching vibration. Amide I and amide II bands are two major bands of the protein infrared spectrum. The amide I band (between 1600 and 1700 cm^{-1}) is mainly associated with the C=O stretching vibration (70-85%) and is directly related to the backbone conformation. Amide II results from the N-H bending vibration (40-60%) and from the C-N stretching vibration (18-40%). This band is conformationally sensitive. Amide III and IV are very complex bands resulting from a mixture of several coordinate displacements. The out-of-plane motions are found in amide V, VI andVIII.



Amide A is with more than 95% due to the the N-H stretching vibration. This mode of vibration does not depend on the backbone conformation but is very sensitive to the strength of a hydrogen bond. It has wavenumbers between 3225 and 3280 cm^{-1} for hydrogen bond lengths between 2.69 to 2.85 Å,
Amide I is the most intense absorption band in proteins. It is primilarygoverend by the stretching vibrations of the C=O (70-85%) and C-N groups (10-20%). Its frequency is found in the range between 1600 and 1700 cm^{-1}. The exact band position is determined by the backbone conformation and the hydrogen bonding pattern.
Amide II is found in the 1510 and 1580 cm^{-1} region and it is more complex than amide I. Amide II derives mainly from in-plane N-H bending (40-60% of the potential energy). The rest of the potential energy arises from the C-N (18-40%) and the C-C (about 10%) stretching vibrations.
Amide III, V are very complex bands dependent on the details of the force field, the nature of side chains and hydrogen bonding. Therefore these bands are only of limited use for the extraction of structural information.

Amino acid side chain vibrations

The presence of bands arising from amino acid side chains must be recognized before attempting to extract structural information from the shapes of amide I and amide II bands. The contribution of the side chain vibrations in the region between 1800 and 1400 cm^{-1} (amide I and amide II region) has been thor. Among the 20 proteinogenous amino acids only 9 (Asp, Asn, Glu, Gln, Lys, Arg, Tyr, Phe, His) show a significant absorbance in the region discussed above. The contribution of the different amino acid side chains were fitted by a sum of Gaussian and Lorentzian components.

AS	vibration		cm^{-1}	A_0 (l/mol/cm)	FWHH (cm^{-1})	surface ($x10^{-4}$ l/mol/cm)
Asp	-COO st as	pH>pK (~4.5)	1574	380	44	5.5
	-COOH st	pH<pK (~4.5)	1716	280	50	4.1
Glu	-COO st as	pH>pK (~4.4)	1560	470	48	7.1
	-COOH st	pH<pK (~4.4)	1712	220	56	3.6
Arg	$-CN_3H_5^+$ st as		1673	420	40	4.3
	st s		1633	300	40	3.6
Lys	$-NH_3^+$ bd as		1629	130	46	1.8
	bd s		1526	100	48	1.3
Asn	-C=O st		1678	310	32	2.7
	$-NH_2$ bd		1622	160	44	2.5
Gln	-C=O st		1670	360	32	3.1
	$-NH_2$ bd		1610	220	44	3.5
Tyr	ring-OH	pH<pK (~10)	1518	430	8	1.0
	ring-O	pH>pK (~10)	1602	160	14	0.7
			1498	700	10	2.5
His	ring		1596	70	14	0.3
Phe	ring		1494	80	6	0.2
terminal						
	-COO st as		1598	240	47	3.5
	-COOH st		1740	170	50	2.1
	$-NH_3^+$ bd as		1631	210	54	3.8

	bd s		1515	200	60	4.3
	-NH_2 bd		1560	450	46	7.5

frequency, absorbance at the maximum (Ao), full width at half height (FWHH), surface of Gaussian band
st=stretching vibration
bd=bending
s=symetrical
as=asymetrical

Secondary structure of peptide model compounds

A large number of synthetic polypeptides has been used for the characterization of infrared spectra for proteins with a defined secondary structure content. For example, polylysine may adopt both beta-sheet or alpha-helical structures in dependence on temperature and pH of the solution. Experimental and theoretical work on a large number of synthetic polypeptides has provided insights into the variability of the frequencies for particular secondary structure conformations

Beta sheet structures (beta strand)

The frequencies of the main absorption bands from synthetic polypeptides adopting an antiparallel chain structure have been compiled by Chirgadze&Nevskaya . From these data it this follows, that the amide I absorption is primarily determined by the backbone conformation and independent of the amino acid sequence, its hydrophilic or hydrophobic properties and charge. The average frequency of the main component is about 1629 cm^{-1} with a minimum of 1615 cm^{-1} and a maximum of 1637 cm^{-1}. The average value for the second frequency is 1696 cm^{-1} (lowest value 1685 cm^{-1}). The parallel beta sheet structure that is not common in synthetic polypeptides leads to an amide I abosrption near 1640 cm^{-1}

Helical structures

The alpha-helix: For alpha-helical structures the mean frequency was found to be 1652 cm-1 for the amide I and 1548 cm^{-1} for the amid II absorptions. The half width of the alpha-helix band depends on the stability of the helix. For the most stable helices, the half-width of about 15 cm^{-1} corresponds to a helix-coil transition free energy of more than 300 cal/mole. Other helices display half-widths of 38 cm^{-1} and helix-coil transition free energies of about 90 cal/mole.
The 3_{10}-helix differs from the alpha-helix in that the internal hydrogen bonding occurs between residues i and i+3 instead of i and i+4 in alpha helices.

Turn structures

The beta turn structure involves 4 amino acid residues which form a loop so that the two chain segments separated by the turn adopt an antiparallel orientation and form an i to i+3

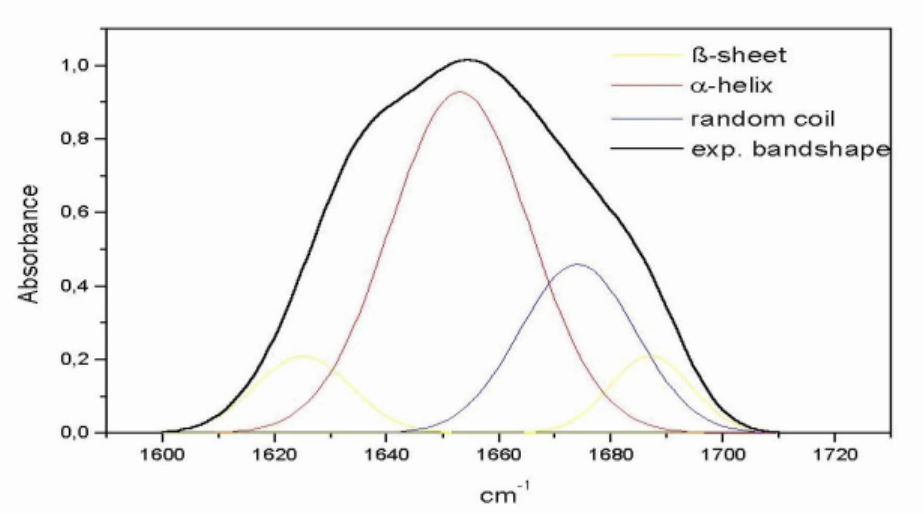
hydrogen bond. A number of turn structures have been identified from protein structures: type I (42%, non-helical), type II (15%, non-helical, requires Gly in position 3) and type III (18%, corresponds to one turn of 3_{10} helix). Assignment of beta turns by means of a normal mode analysis for insulin demonstrates a strong overlapping of the different types of beta turns with the alpha-helical absorption. However, an absorption near 1680 cm^{-1} is now clearly assigned to beta turns.

Secondary structure in proteins

The shape of the amide I band of globular proteins is characteristic of their secondary structure. With a publication by Byler & Susi The determination of secondary structures in proteins from FTIR spectra actually started. This had become possible by the availability of high signal-to-noise ratio digitalised spectra obtained by the FTIR spectrometer and by the access to computers and software able to perform many operations on the spectra in a short time.

Deconvolution of the amide I band

The concept of Fourier self deconvolution is based on the assumption, that a spectrum of single bands (each narrow band is characteristic for a secondary structure) is broadened in the liquid or solid state. Therefore, the bands overlap and can not be distinguished in the amide envelope. A curve fitting procedure can be applied to estimate quantitatively the area of each component representing a type of secondary structure. In the pioneering work by Susi & Byler the amide I was deconvoluted with a Lorentzian line shape function and a resolution enhancement factor of 2.4 was applied. The deconvoluted spectrum was fitted with Gaussian band shapes by an iterative curve fitting procedure. The results are in good agreement with with the secondary structure information obtained from X-ray crystallographic structures of the proteins under study.



	a)			*b)*			
sec. structure	**Mean (cm^{-1})**	**RMS (cm^{-1})**	**Max (cm^{-1})**	**Mean (cm^{-1})**	**RMS (cm^{-1})**	**Max (cm^{-1})**	**Region (cm^{-1})**
turns	1694	1.7	2	-	-	-	
	1688	1.1	2	-	-	-	
	1683	1.5	2	1678	2.1	5	**1682-1662**
	1670	1.4	2	1670	2.9	5	
	1663	2.2	4	1664	1.0	3	
alpha-helix	1654	1.5	3	1656	1.5	3	
				1648	1.6	3	**1662-1645**
unordered	1645	1.6	4	1641	2.0	3	**1645-1637**
beta sheet	1624	2.4	4	1624	2.5	5	
	1631	2.5	3	1633	2.1	4	**1637-1613**
	1637	1.4	3	-	-	-	
	1675	2.6	4	1685	2.1	4	**1689-1682**

PROTEIN ENGINEERING

Protein structure determination by electron cryo-microscopy

Cryo-Electron Microscopy specializes in interpreting and visualizing unstained biological complexes such as viruses, small organelle, and macromolecular biological complexes of 200 kDa or larger preserved in vitreous (*i.e.* glassy or non-crystalline) ice. The basic goal is to compare other electron microscopy techniques to use cryo-fixation to rapidly freeze the biological sample so as not to destroy its aqueous enviornment. This avoids ultrastructural changes, redistribution of elements, and washing away of substances. Specimens frozen in vitreous ice show a structure similar to the liquid state, or the native state. The near native imaging conditions allows three dimensional reconstruction of the cellular machinery. Using state of the art computer controlled, automated microscopes, image reconstruction software, and visualization tools, sub-nanometer resolution structures of large biological complexes can be achieved. In Cryo-Electron Microscopy, an electron beam, a stream of high energy particles bombards the sample. The image that is viewed is a result of the interaction of the sample with this beam. Most of the electrons that form the high resolution image appear due to elastic scattering, where only their trajectory has been changed, but their energy is unaffected. However, a small fraction of the electrons transfer some of their energy to the sample. This energy accumulates and can break apart molecular bonds, destroying the sample after some time. Therefore, for high-resolution imaging, low dose parameters require that the area to be imaged is not exposed until the picture is actually taken.

Cryo-electron microscopy can be performed by various methods of specimen preparation, two popular methods use thin film and vitreous sections of biological material. The thin film method requires biological material to be placed on an electron microscopy grid and is rapidly frozen close to liquid nitrogen temperatures. Larger samples (vitreous sections) can be vitrified by different methods including high pressure freezing. These samples can then be cut thinly and placed on the electron microscopy grid, similar to the thin film. These samples must remain at liquid nitrogen temperature to undergo the high vacuum and are exposed to the electrons.

One branch of Cryo-electron microscopy is Cyro-Electron tomography (CET). Cyro-electron tomography is performed at cryogenic temperatures as is cryo-electron microscopy; CET constructs a 3D sample from 2D images.

Uses for Cryo-Electron Microscopy

Cryo-electron microscopy is used in a variety of fields. Nanoparticle research relies heavily on electron microscopy for the visualization of small particles. Pharmaceutical companies doing drug research utilize electron microscopy to help predict the behavior of drugs and biological matter. In the case of pharmaceutical research a 3D visualization is extremely useful and at cyro-electron microscopy proposes the least damage to the sample to obtain a usable image.

PROTEIN ENGINEERING

Advantages in using Cryo-Electron Microscopy

1. Allows the examination of native and hydrated structural features of the biological sample. The sample is always in solution and never comes into contact with an adhering surface. Therefore, the shape that is observed is the true shape of the hydrated molecule in solution and has not been distorted by attaching itself and flattening against the supporting film.
2. Provides good preservation of biological structure in the microscope vacuum.
3. There are no stains or chemical fixatives to distort the sample. When stained, the sample can be damaged in many ways, such as flattening and twisting.
4. When the sample adheres to the carbon grid, it could stick in a preferential orientation. If this happens, then information will be missing from the final image set (a missing cone), and the resolution of the calculated model in that direction will be absent.
5. Provides a 2-5 fold reduction in radiation damage compared to similar sugar-embedded or freeze-dried samples at room temperature. The reason behind this is thought to be from decreasing the temperature-dependent rearrangement or diffusion of fragments resulting from bond-fracture. In the solid frozen state, rearrangement or diffusion is decreased and the protein conformation is more likely to be maintained up to higher levels of irradiation.
6. Can observe contrast between nucleic acids, proteins, and lipids to be distinguished.
7. Enables one to control the chemical environment so that examination of different functional states of molecules is possible.

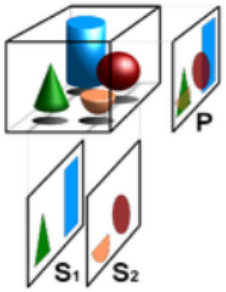
Disadvantages in using Cryo-Electron Microscopy

1. Very low signal to noise ratio. Biological macromolecules are normally made up of carbon, hydrogen, oxygen, and nitrogen. The electron absorption of such molecules is very low. As a result, image contrast is also very low and it is hard to detect features when dealing with just a few images.
2. Difficult to obtain images from tilted specimen. The ice cross section of a tilted frozen sample is too thick to yield good images.
3. Charging is more widespread when imaging a tilted frozen sample.
4. More time consuming to generate samples. However, this is generally not a big problem, especially once a working protocol is designed and good samples are readily available.
5. If vitreous ice cannot be easily formed, the resulting cubic ice absorbs electrons very easily and the frozen sample is basically worthless.
6. Sample must be maintained at less than 135 degrees Celsius.

Preparation of Frozen-hydrated Biological Specimens

The following are some general procedures for preparing frozen-hydrated biological specimens: 1)Development of a thin layer of the biological specimen. 2)Rapid cooling of the specimen to the vitreous state. 3)Transfer the specimen to the electron microscope without rewarming above the devitrification temperature. 4)Observe the specimen below the devitrification temperature with an electron dose that is low enough to preserve the structure of the sample.

Cryo-Electron Tomography



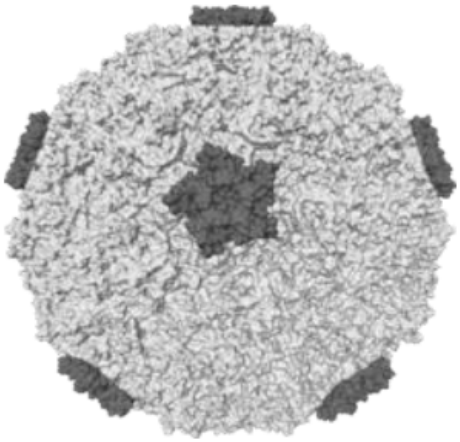
Tomography uses the effects and differences that waves of energy have on a solid object to produce a three dimensional image of the internal structure. Cryo-Electron Tomography is a branch of Cryo-Electron Microscopy in which two dimensional projections of a frozen sample, at cryogenic temperatures, are recorded and used to reconstruct a three-dimensional structure by computed back projection. This is done using a transmission electron microscope to take successive images of a sample while tilting the sample around an axis. The "projection theorem" states that a 3D object can be retrieved from its projections along different directions. So to obtain a 3D description of an object, it must be projected along different directions; this is achieved by incrementally tilting the specimen. Due to the limitations with the transmission electron microscope (TEM), the specimen can only be tilted to +- 60-70 degrees, and not to 90 degrees which would be necessary to retrieve all the 3D information about the specimen.

Cryo-ET is a very accurate way to determine the three-dimensional structure of a specimen because the rapid freezing of the object and cryogenic temperatures gives a good preservation of the structure and good time resolution of certain processes. For example, the rapid freezing of cells and tissues at a certain point in cellular processes can give a good understanding of the structure and activity of those cells and tissues at a certain point in time during that particular cellular process. This type of tomography aids in the learning of cells and their organelles at a more dynamic level. Each organelle of a cell is produced in a different color, in order to facilitate the viewing process. The cells are frozen in order for the cell to retain its original structure. Freezing such specimens is done by placing them on a grid, blotting them in a thin layer of water and emerging them into ethane before storing them into liquid nitrogen. The use of cyro-electron tomography involves the study of almost all specimens,

such as viruses. This tool can be helpful in understanding the replication states of viruses, as well as, the individual structures that viruses can become. A recent study has been done on the spikes of the viruses and the various structures the spikes affect the virus. Today, cyro-electron tomography is used to help find a cure for cancer by assessing the building blocks of the protein, cadherins, which aid in blocking cancerous tumors for spreading throughout the body. The information obtained through Cryo-ET can aid in comprehending and understanding the structural basis, and therefore, the function of many cellular processes.

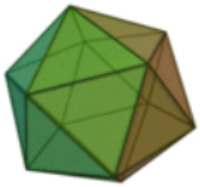
There are limitations with Cryo-ET. The main limitation is the thickness of the specimen. The specimen must been thin enough for it to freeze well and so that it can be properly collected with the TEM. If the specimen is too thick, it must be cut into thinner slices while the temperature is still very low, so that re-crystallization does not occur. There are a couple ways to obtain the images, one of the is by fixed tilt increments, and the other is by graduated tilt increments. Graduated tilt increments are more favorable, this is when the tilt increment is proportional to the cosine of the tilt angle. Another issue with Cryo-ET is radiation damage. To prevent radiation damage, the specimen should be imaged under low electron dose conditions, leading to a more limited resolution in the 3D image obtained, and it also limits the specimen thickness needed for Cryo-ET.

Icosahedral Reconstruction



Icosahedral Reconstruction refers to the application of cryo-electron microscopy in elucidating the structure of particles with appropriate (icosahedral) symmetry. The high internal symmetry of icosahedral specimens makes it easier to determine the positions of symmetry elements, thereby decreasing the amount of images required to determine the 3D structure of the specimen.[1] This may seem like an arbitrary and irrelevant solid to apply to microscopy, but in fact there is an enormous number of particles that contain such symmetry. Examples of icosahedral particles include the majority of human viruses, as well as some molecules such as dodecahydro-closo-dodecaborate ion ($B_{12}H_{12}^{2-}$) and the buckminsterfullerene.

PROTEIN ENGINEERING



A number of virus structures have been predicted and subsequently experimentally determined through the use of icosahedral reconstruction. An icosahedron belongs to the high symmetry group I_h which contains 120 symmetry operations, possibly most unique being the six five-fold symmetry axes.The properties of this symmetry group are essential for the application of cryo-electron microscopy.

Helical Reconstruction

Helical Reconstruction is a method that takes advantage of Cryo-Electron Microscopy in order to develop a three-dimensional structure for certain "filamentous" biological structures. This method used the 2-D projection images from Cryo-Electron microscopy to produce these 3-d Images as long as there is helical symmetry. This method however cannot be applied to structures that contain "seams" or "pertubations". There is a new method known as asymmetrical helical reconstruction that can be applied to helical structures that contain "seams". Similar to conventional Helical reconstruction methods, Fourier transform images are used to produce the layer line data which are then used to produce the 3-d structures.

Helical reconstruction allows the formation of large groups by regular contacts of a single type of protein molecule. Helical symmetry can be found in filamentous viruses (e.g., Pf1), in the proteins of the actin, tubulin, or other cytoskeleton, or in the proteins that form 2-D crystals folded onto the surface of a cylinder, such as the acetyocholine receptor or CopA.

Electron Crystallography

Electron crystallography is a form of microscopy that uses a beam of electrons to construct images of small solids such as proteins. This process is used to determine and predict the structure and arrangement of a protein from secondary structure crystals such as alpha helices or beta sheets based on electron scattering. It can be used to study both organic and inorganic matters, and also protein structures. Electron Crystallography complements X-ray crystallography in many ways but also succeeds where X-ray crystallography fails. For example, X-ray crystallography study requires the quaternary structure of proteins which is often hard to attain than secondary structures. Electron Crystallography presents a problem in that it can cause radiation damage to the proteins under analysis. This hinders the range and function of the microscopy process. In order to reduce radiation damage, cryofixation, in which the imaging takes place in very low temperatures such as that of liquid nitrogen, is implemented. This resource is especially valuable when a specific protein is easily denatured or damaged by the electrons from the microscope.

PROTEIN ENGINEERING

A crystal structure determination includes two steps: ‘solving’ which finds a model of the heaviest atoms within about 0.25 Å using EM-images; and ‘refine’, which finds all atoms within about 0.02 Å using Selected Area Electron Diffraction or Convergent Beam Electron Diffraction data.

The use of electron diffraction in order to study the structures of crystals began in Moscow in 1937-1938 among a group of scientist led by Pinsker and Vainshtein. Their study used their own electron diffraction cameras that had relatively low acceleration to record electron diffraction data of different materials. From this data, they were able to locate hydrogen atoms in crystal structures which can not be done using X-ray diffraction. In order to solve unknown structures, phase information is needed which was first introduced by Hauptmann and Karle in 1953 called the "direct methods". Combining the use of direct phasing methods with modern day computers, electron crystallography has made significance advances in structure determination of crystals and other molecules.

There are two different electron diffraction techniques: 1)Selected Area Electron Diffraction (SAED)which requires near kinematic condition and applies for unit cell dimension >10 Å and for thin specimens <200 Å; 2)Convergent Beam Electron Diffraction (CBED)which makes use of dynamical effects and applies for unit cell dimensions <10 Å and for thick specimens >200 Å

Why electrons? Electrons are used in favor of X-rays because it is 10^4-10^5 times stronger interaction with matter compared with X-ray; and their phases are present in high resolution electron microscopy images.

There are some key advantages of electron crystallography compared to X-ray crystallography. One of these advantages is that electron crystallography can analyze much smaller crystals. This is because electrons interact more dominantly with matter than X-rays do. Another advantage is that electron beams can be focused by magnetic lenses to create an image while X-rays cannot form an image. Because the mechanism by which electrons interact with matter is based on the electrons detecting potential distributions in crystals compared to the mechanism of X-rays which depends on the X-rays detecting electron density distribution, electron crystallography can be used in certain situations that X-ray crystallography cannot. For example, the oxidation states of atoms in a crystal.

Using electron crystallography to determine structure is important due to the ability for a protein to be observed in its natural form. By utilizing electron crystallography, one can observe a protein in a lipid-protein bilayer in the structure that it is found in, thus allowing for better determination of function.

Single-particle electron microscopy

The techniques used to reconstruct the 3 Dimensional images of the molecule from a collections of 2-D images is called electron microscopy. It presents to structural biochemists insight views in term of structural information of many biological molecules because of its

easy-to-access features. In order to acquire 3D structure from this method, two requirement must met.

1. reasonable size of proteins to large macro-molecular assemblies without need to use crystals 2. the molecules must exist in many identical copies

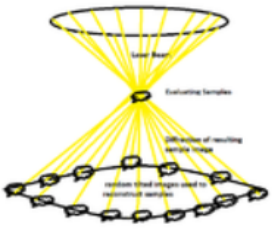
The resolution produced by electron tomography have a low resolution and high noise. The main goal of the single particle electron microscopy is to determine the geometric relation between the collected projection images. In the year of 2008, scientist were able to made it possible to trace the backbone of the polypetide chains and build atomic models. Single Particles EM indeed have the capability to deliver structural information at near atomic resolution.

The 3D structure would be the result of these following steps:

1) Sample preparation. This is the step where sample is collected and place on (metal) plate to generate the best contrast. There are three technique used to prepare the samples for single particle EM. a) Negative staining : molecules are adsorbed to a continuous carbon film in which molecules are put into a metal plate by drying b) vitrification: sample is plunged into liquid ethane to preserve molecules in a native environment, it produced low contrast. (preserving them in a fully hydrated state)Vitrification is the **best** speciment preparation method, but not applicable to heterogeneous samples. c) Cryo negative staining: high contrast image immersed in high ionic strength of saturated ammonium molybdate solution. It is good to study the small and heterogeneous samples.

2) Particle picking. This is one of the most tedious processes of all because electron microscopists have to classify and separate particles according to their similarities in orientations; all these works are done by hand in order to achieve maximum efficiency. Having said that automated picking programs have been developed, but they failed to perform the task thank to the low signal-to-noise ratio. Result of this step is collection of small individual images of particles.

3) Generate initial model. Individual images collected from previous step are used to build a preliminary model. RCT is the primary method that is being used to generate initial models. **Random Conial Tilt Reconstruction** is being use throughout the process. Random Conical Tilt are the reconstruction of the 3 dimensional image one at high tilt angle and the other at untilt angle. The tilt angles allowd the testing samples to align in unique orientations. Please see the image at the right for the figure. Random Conical Tilt is common used in negatively or cryo-negatively stained speciments, in which works well with heterogenous particle solutions.



4) Refinement. The preliminary model is used to calculate for better alignment using Euler angles, in-plane rotated/unrotated shifts of particles. From refined data, a new 3D structure of a molecule is reconstructed. The risk of overrefinementhappeneds when the testing negative temperature is applied.

Single-particle electron microscopy is having advantages ahead of electron crystallographic because its economic features. Unlike crystallographic, it does not require crystals – meaning samples don't have to be pure - which take a lot more works to achieve. Another advantage is that single particle electron microscopy takes very little sample which always makes researchers happy.

One of the disadvantages of single particle electron microscopy, however, is that it is difficult to determine the resolution of density maps as well as their accuracy. Since there is no evidence of a method to check for accuracy, often the only thing that can be done is to repeat the process and compare results to previous results. Therefore, results can only be assessed on the basis of consistency rather than accuracy.

Protein-Lipid Arrays

Electron crystallography has been used in the study of membrane proteins by analyzing arrays of samples called protein-lipid arrays. These arrays can be arranged in many different ways and offer many advantages and disadvantages. Two forms that have given the most useful density maps of membrane proteins are the two dimensional sheet like crystal arrays and the tubular like crystal arrays. These are so precise that they can reveal information about individual lipid molecules and the protein side chains due to averaging the many unit cells in the image of a sheet or tube like array enhancing the poor signal-to-noise ratio.

Over the last few decades there has been much advancement in the methods used to create these protein-lipid arrays both in tubes and sheets from detergent-solubilized purified proteins but no advancements in screens have been made like an easily manipulated robotic screen. Despite much research in the field they still lack the ability to quickly and reliably check the quality of the samples. Several laboratories have advanced the methods of analysis of these crystals making the process more efficient and more user friendly by enhancing the existing software.

The tubular crystals have not seen as extensive use as the sheet crystals even though their helical array symmetry allows for substantial advantages in determining structure. One image of a tube contains many different views of the same molecule which is enough to reconstruct

it in three dimensions without the need for tilting. To correct for distortions tubes are processed in a similar way to sheets, where two repeat lengths are divided into shorter pieces and are then compared to a reference structure to determine the parameters needed to help identify the structure completely. This procedure traditionally uses Fourier-Bessel methods to assess the data, enabling them to analyze the extent of helical symmetry preservation and twofold symmetry perpendicular to the tube axis, which can correct for the focus changes at different levels of the structure. Another method has been developed that doesn't employ the Fourier-Bessel methods and instead treats segments as strings of single particles. This alternative is becoming more popularly used for extracting structural information from poorly ordered helical polymers such as tubular protein-lipid crystals. This shows great potential for determining structures from tubes at the near atomic level of resolution.

Methods involved in electron crystallography include free-trapping to create different conformational states. To freeze-trap the specimen, the electron microscope grid is placed into liquid nitrogen-cooled ethane, which cools the specimen rapidly enough that thus allows for the trapping of a structure of a lipid-protein array which has a life-time of a millisecond or longer. The freeze-trapped protein can be activated through light or an appropriate ligand. Recent developments of helium-cooled top-entry freeze-trapping has resulted in a more clear image for data collection, and hopefully would allow for the gating mechanism of the protein-lipid bilayer to be described in more detail.

Additionally, molecular tomography is used to explore proteins in their functional context. A three-dimensional picture of an entire scene is possible to create though taking images from a series of tilt views, therefore creating a better three-dimensional image.

Single-molecule methods

It is a method that observing dynamic behavior of single molecule to determine mechanism of action at level of an individual molecule, and to identify, sort and compare subpopulation and substructure within cell. In order to characterize the dynamics of molecular structures, scientists look to real-time trajectories of individual molecule; and by observing many of them, a histogram of the dynamical properties over the population could be figured.

X-ray crystallography or NMR, in comparison to single-molecule methods, provides detail structural view but limit by static molecular view and ensemble average.

1. Single-molecule manipulation: In this method, molecules are attached to an external probe which exerts defined forces or torques on molecule in order to characterize their mechanical properties. This method is also called atomic force microscopy (AMF). Because cell is seen as a factory in which many processes are carried out by specialized machinery which converts the chemical energy into force, torque and mechanical work - of which the attached probe will now come in to detect the dynamics and mechanism. This method is recently used to study the folding and unfolding of RNA molecules and the enzymes that catalyze these reactions, and to study RNA polymerase.

2. Single-molecule detection: The molecule is tagged with a fluorescent label in two locations in the form of a “donor” and “acceptor” that can undergo fluorescence resonance energy transfer (FRET). The trajectory of molecules then can be watched regarding to a change in the intensity of the florescence of the probe or regarding to the change in FRET. Another name of this method is fluorescence method. . This is a powerful method to study dynamic behavior of molecules, their stability and track particles’ movements in and outside cell. This method is used to study, for example, the multiple interactions during translation by the ribosome.

Between the two mentioned above, the fluorescence detection method is preferable to researchers because it requires less elaborate and complex instrumentations, but the down side is the photons collected by instrument is limited.