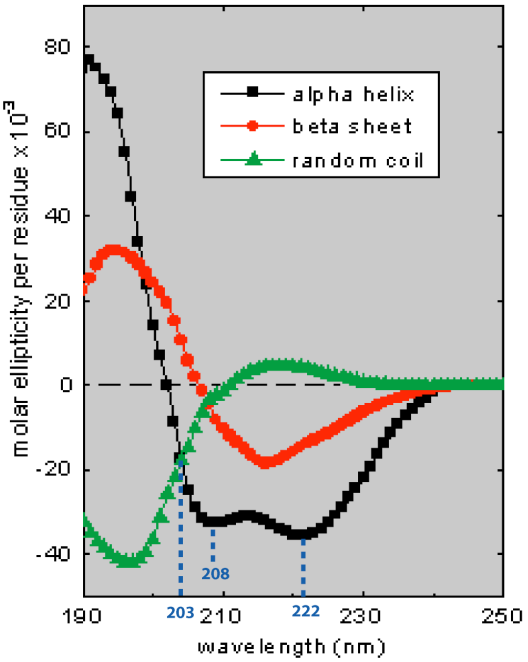
Some General Information on CD of Proteins

Sources include: http://www.ap-lab.com/circular_dichroism.htm

Far-UV range (190-250nm)

Secondary structure can be determined by CD spectroscopy in the far-UV region. 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 below, which shows spectra for poly-lysine in these three different conformations.



- **Alpha helix** has negative bands at 222nm and 208nm and a positive one at 190nm.
- **Beta sheet** shows a negative band at 218 nm and a positive one at 196 nm.
- **Random coil** has a positive band at 212 nm and a negative one around 195 nm.

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. (e.g. For an alpha helical protein with increasing amounts of random coil present, the 222 nm minimum becomes shallower and the 208 nm minimum moves to lower wavelengths ⇒ black spectrum + increasing contributions from green spectrum.) 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 helical portion.

There may also be significant contributions from the aromatic groups in the far-UV region. In some cases the positive contributions from the aromatic groups offset completely the negative ellipticity from the protein, so that only a very weak signal is obtained.

Solvent transparency is extremely important in scanning down to lower wavelengths. You need to avoid compounds that absorb in the region of interest. See my webpage on CD Sample Preparation for further details: http://structbio.vanderbilt.edu/wetlab/cd.sample.prep.php

For far-UV, people use smaller pathlength (1mm) cells because they decrease solvent absorbance when scanning down to lower wavelengths. A typical protein concentration for a 1mm cell is ~0.1 mg/mL. Minimum volume for the CSB's cuvette is ~150uL.

Near-UV range (250-350nm)

At these wavelengths the chromophores are the aromatic amino acid sidechains and disulfide bonds, and the CD signals they produce are sensitive to the overall tertiary structure of the protein. 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.

Signals in the near-UV region are coming from:

- Phenylalanine: 250-270 nm
- Tyrosine: 270-290 nm
- Tryptophan: 280-300 nm
- Disulfide bonds give rise to broad weak signals throughout the near-UV spectrum.

Usually tryptophan will give the biggest signal, followed by tyrosine and then phenylalanine. The CD bands from individual residues may be positive or negative and may vary widely in intensity, so it is often difficult to separate out the contributions of individual aromatic residues. The signals may also cancel each other out, so no near UV CD signal does not necessarily mean no tertiary structure.

The signal strength in the near-UV CD region is much weaker than that in the far-UV CD region, so larger pathlength cells (1cm) and higher sample concentrations are needed. Near-UV CD spectra require ~0.5 ml of protein solution with an A_{280nm} = 0.5 to 1 (which corresponds to 0.25 to 2 mg/ml for most proteins).

Visible range

Prosthetic groups such as metal ions, heme, and retinal exhibit CD signals in the visible region.

Converting your CD signal to mean residue ellipticity

(ref: Myers et. al. Biochemistry (1997) **36**:36 p.10923-10929)

The units used in CD spectroscopy can be confusing! The Jasco J-810 gives its raw data output in ellipticity (measured in millidegrees). For proteins and peptides, CD data are usually reported in units of mean residue ellipticity (degrees squared centimeters per decimole). To process your data:

1. Subtract the blank signal of cuvette and buffer. (Note: the blank data set should have been collected under identical conditions as sample spectrum.)
2. Apply the following equation:

$$[\theta] = \frac{100(\text{signal})}{Cnl}$$

where:

$[\theta]$ = mean residue ellipticity in deg cm^2 $dmol^{-1}$
signal = raw output in mdeg
C = protein or peptide concentration in mM
n = # amino acid residues
l = cell pathlength in cm

With these units,

Far-UV data should give $[\theta]$ ~ 10,000-20,000 deg cm^2 $dmol^{-1}$
Near UV data should give $[\theta]$ ~ 100-1,000 deg cm^2 $dmol^{-1}$

Determining secondary structure from CD data

Accurate protein concentrations are essential for determining the distribution of secondary structures in a sample. This is usually done by measuring A_{280nm} and calculating the concentration using known extinction coefficients for the aromatic residues. Quantitative amino acid analysis is an alternative method. Lowry or Bradford analyses may be in error by a factor of two and are not sufficiently accurate for use with CD measurements unless they have been calibrated against a careful amino acid analysis of the protein of interest.

Here are a couple of links to websites that have different algorithms for estimating secondary structure using CD data. Dichroweb (the first link) requires registration but has several different programs (selcon3, contin, varslc, cdsstr, and k2d) to choose from. The users guide on their website gives more details on the various programs. Most of the programs require CD data down to 190nm or lower with the exception of k2d which only requires data from 200nm- 240nm.

http://www.cryst.bbk.ac.uk/cdweb/html
http://www.embl-heidelberg.de/~andrade/k2d.html

Equations for calculating fractional helicity using the signal at 222nm and assuming only helix and random coil are present can be found in: Rohl & Baldwin Biochemistry (1997) **36**, p.8435-8442. There are no equivalent equations for calculating fraction of beta sheet, just various fitting programs on the web and usually they require data down to 180nm to obtain reasonable fits.

Coiled Coils

The (222/208) nm ratio ≥ 1 for coiled coils and ≤ 0.86 for isolated helices. For further details, see:

Lau SY, Taneja AK, Hodges RS, “Synthesis of a model protein of defined secondary and quaternary structure. Effect of chain length on the stabilization and formation of two-stranded alpha-helical coiled-coils.” J Biol Chem. 1984 Nov 10;259(21):13253-61.

Zhou NE, Kay CM, Hodges RS. “Synthetic model proteins: positional effects of interchain hydrophobic interactions on stability of two-stranded alpha-helical coiled-coils.” J Biol Chem. 1992;267: 2664-2670

Measuring protein stability by CD

Thermal and denaturant melts can be used to determine the stability of a protein. For an excellent review and protocol see:

Pace, C.N. and J.M. Scholtz, *Measuring the conformational stability of a protein* in "Protein Structure: A Practical Approach" (Creighton, T. E.) pp. 299-321, IRL Press, Oxford, 1996.

A copy of this book chapter is in the red binder in the CD room (MRB3 5141)—please Xerox and return. For denaturant melts, you will need to measure the concentration of urea/guanidine by refractive index. The CSB owns a refractometer that you can use—see Laura.