Circular Dichroism: Studies of Proteins

Sharon M Kelly, University of Glasgow, Scotland, UK
Nicholas C Price, University of Glasgow, Scotland, UK

Based in part on the previous version of this Encyclopedia of Life Sciences (ELS) article, Circular Dichroism: Studies of Proteins by Nicholas C Price.

Circular dichroism (CD) is the differential absorption of the left- and right-circularly polarized components of plane-polarized electromagnetic radiation. It can provide structural and dynamic information about biological macromolecules, particularly proteins. The CD spectra in the far-UV (typically 180–240 nm) can give reliable quantitative estimates of the proportions of secondary structural features (helix, sheet, turn, etc.) present in proteins. The spectra in the near-UV (260–320 nm) can be used to explore the environments of aromatic amino acid side-chains and hence to give a measure of tertiary structure. Although CD cannot provide the high-resolution structural data available from X-ray crystallography or nuclear magnetic resonance, its convenience and applicability under a wide variety of experimental conditions make it the technique of choice in many applications, including exploring protein–ligand interactions, conformational changes and protein folding.

Basic Principles of Circular Dichroism

Circular dichroism (CD) refers to the differential absorption of the left- and right-circularly polarized components of plane-polarized electromagnetic radiation. This effect will occur when an absorbing group (or chromophore) is chiral (i.e. optically active) for one or more of the following reasons:

- it contains one or more chiral centres (e.g. carbon atoms);
- it is covalently linked to a chiral centre and
- it is placed in an asymmetric environment.

This article is confined to the CD of electronic transitions in molecules; for more specialized aspects such as vibrational CD (see Polianichko and Wieser, 2005). In the CD experiment, plane-polarized radiation is split into its circularly polarized components by passage through a modulator. This is usually a piezoelectric crystal such as quartz subjected to an alternating electric field (usually at 50 kHz). The modulator transmits each of the components in turn. If, after passage through the sample of interest, the components are not absorbed or are absorbed to equal extents, their recombination will generate radiation polarized in the original plane. However, if one of the components is absorbed by the sample to a greater extent than the other, the resultant recombined radiation will be elliptically polarized (Figure 1).

ELS subject area: Structural Biology

How to cite:
DOI: 10.1002/9780470015902.a0003043.pub2

Figure 1 Origin of the CD effect. The left- (L) and right- (R) circularly polarized components of plane-polarized radiation: (a) the two components have the same amplitude and when combined generate plane-polarized radiation and (b) the components are of different magnitude and the resultant (dashed line) is elliptically polarized.
In practice, the CD instrument, or spectropolarimeter, does not recombine the components, but detects them separately and displays the dichroism at a given wavelength of radiation expressed as either the difference in absorbance of the two components \((\Delta A = A_L - A_R)\) or as the ellipticity \((\theta)\) in degrees (where \(\theta = \tan^{-1} (b/a)\) and \(b\) and \(a\) refer to the minor and major axes of the resultant ellipse). There is a simple numerical relationship between \(\Delta A\) and \(\theta\) (in degrees), namely \(\theta = 32.98 \Delta A\). In terms of molar units, the CD signal can be expressed as either a residual absorbance \(\left(\Delta A; \text{units cm}^{-1} \text{L mol}^{-1}\right)\) or a difference in molar absorbance \(\Delta a^\perp\) (in molar units, \(\Delta a^\perp\) or as the ellipticity \(\left[A^\perp; \text{units cm}^2 \text{dmol}^{-1}\right]\)). The two units are related by \(\theta = 3298 \Delta a^\perp\) (Martin, 1996; Kelly et al., 2005). A CD spectrum is obtained when the dichroism is plotted as a function of wavelength; it should be noted that the CD signal can be positive or negative, depending on whether the left or right component is absorbed more at that particular wavelength. See also: Spectroscopic Techniques

In most biological work, the observed ellipticities are of the order of 10 mdeg; that is, the difference between \(A_L\) and \(A_R\) is of the order of \(3 \times 10^{-4}\) absorbance units. Given the small magnitude of this effect, it is essential that careful attention is paid to the experimental conditions to ensure that meaningful data are obtained. The CD spectra of proteins can give valuable insights into their secondary and tertiary structure and conformational changes that may accompany binding of ligands or denaturing agents. CD can be used in a time-resolved mode to elucidate a number of aspects of protein folding and unfolding. In addition, CD can be used to study other biological chromophores, such as cofactors or nucleic acids. Reliable CD data obtained by paying careful attention to experimental details are being increasingly used in characterizing therapeutic proteins and other products of commercial interest (Lees et al., 2004), and a Protein Circular Dichroism Data Bank (PCDD) for the deposition of validated CD data is being established along the lines of the Protein Data Bank (PDB). See also: Proteins: Fundamental Chemical Properties

The Far-UV CD Spectra of Proteins

The CD spectrum in the far-UV region, typically from 240 nm to 180 nm (or even lower; see below) can be used to provide quantitative estimates of the secondary structure content of a protein (Johnson, 1990). In this region, the absorbing chromophore is principally the peptide bond. There is a weak but broad \(n \to \pi^*\) transition centred around 220 nm and an intense \(\pi \to \pi^*\) transition about 190 nm. It should be noted, however, that aromatic amino acids (especially tryptophan) and disulfide bonds can also contribute to the CD spectrum in this region. The different forms of regular secondary structure found in peptides and proteins exhibit distinct far-UV CD spectra (Figure 2).

Assessing the contribution of the different structural types to a given spectrum thus consists essentially of the solution of a number of simultaneous equations using data over an appropriate range of the far-UV spectrum. The various algorithms available have been reviewed extensively (Greenfield, 1996; Srerama and Woody, 2004a, b). The online server DICHROWEB has been developed, hosted at Birkbeck College, London, UK, which allows CD data to be entered in a number of formats and to be analysed by the various algorithms available with a choice of reference databases (Lobley et al., 2002; Whitmore and Wallace, 2004). In general, the various methods provide rather more reliable estimates of the contents of \(\alpha\) helix compared with \(\beta\)-sheet or \(\beta\)-turn. If data are only available to approximately 200 nm, it is possible to make estimates of the \(\alpha\) helical content from the ellipticities at 208 and 222 nm, which correspond to the double minima characteristic of the \(\alpha\) helix (Figure 2). See also: Peptide Bonds, Disulfide Bonds and Properties of Small Peptides

The light source of a conventional spectropolarimeter is a Xenon arc which provides radiation down to approximately 180 nm. However, using the more intense radiation available from synchrotron sources (synchrotron radiation circular dichroism, SRCD) data down to approximately 160 nm can be collected. The analysis of far-UV CD spectra can be undertaken with much greater confidence when spectral data are available over this wider range of wavelengths (Wallace and Janes, 2001).

The Near-UV CD Spectra of Proteins

The near-UV CD of proteins arises from the environments of each aromatic amino acid side-chain as well as possible contributions from disulfide bonds or other nonprotein cofactors that might absorb in this region of the spectrum. Small model compounds of the aromatic amino acids exhibit CD spectra because the chromophore is linked to the nearby chiral \(\alpha\)-carbon atom. In proteins in their native states, the side-chains of these amino acids will be placed in a variety of asymmetric environments characteristic of the tertiary structure of the protein. Each of the aromatic amino acids tends to have a characteristic wavelength profile: tryptophan has a peak close to 290 nm with fine
structure between 290 and 305 nm; tyrosine has a peak between 275 and 282 nm (the fine structure at longer wavelengths may be obscured by that from tryptophan); phenylalanine has sharp fine structure between 255 and 270 nm (Figure 3). The following factors tend to influence the intensities of aromatic CD bands:

- The number of aromatic amino acids. Proteins with large numbers of aromatic amino acids can often have CD bands smaller than expected because of cancelling effects of positive and negative contributions.
- The rigidity of the protein, with the more highly mobile side-chains having lower intensities. See also: Protein Structural Flexibility: Molecular Motions
- Interactions between aromatic amino acids, which are particularly significant if the distance between them is less than 1 nm.

The contributions of individual aromatic amino acids to the near-UV CD spectrum of a protein can be assessed by examining the CD spectra of suitable mutant proteins in which they have been replaced in turn by other amino acids with different spectral properties. Particularly detailed studies have been made of the contribution of each of the seven tryptophan (Trp) side-chains to the spectra of human carbonic anhydrase II (Freskgärd et al., 1994) and of the six tyrosine (Tyr) side-chains in bovine ribonuclease A (Woody and Woody, 2003). Although near-UV CD spectra of proteins are not readily amenable to detailed interpretation in terms of tertiary structural features (Kahn, 1979), progress is being made in calculations of such spectra from known three-dimensional structures of proteins (Alden et al., 1997; Kurapkat et al., 1997; Rogers and Hirst, 2004). At an empirical level, near-UV CD spectra can provide very useful fingerprints for comparisons of the tertiary structures of related proteins, such as wild-type and mutant proteins. They have also given valuable insights into the partially folded, ‘molten globule’ state of proteins. See also: Biological Macromolecules: UV-visible Spectrophotometry

Experimental Aspects of Circular Dichroism

It is important to emphasize a number of the experimental parameters involved in collecting CD data, since the quality of the data depends critically on these. Because of the relatively small size of the CD signal, parameters such as the time constant, scan rate, number of scans and bandwidth should be optimized (Kelly et al., 2005). In addition, the spectropolarimeter should be properly calibrated and the pathlengths of cells used should be carefully checked (Kelly et al., 2005). It is possible to study events on the stopped-flow time scale (10 ms or less) by accumulating data from a number of replicate experiments to improve the signal-to-noise ratio.

To keep the noise within acceptable levels, it is important to ensure that the total absorbance of the solution (arising from the protein and the solvent, etc.) should be kept below about one unit (in theory the measurements are most accurately made when the total absorbance is 0.863). Careful selection of solvents, buffers and any added electrolytes should be made to ensure that these components do not lead to a high overall absorbance; in addition the sample concentration and cell pathlength should be chosen to give an appropriate absorbance. It should be noted that high concentrations of imidazole (which is used to elute His-tagged recombinant proteins from immobilised metal affinity chromatography columns) will give rise to high absorbance below 220 nm. In addition, high concentrations of chloride ions (often added to increase the ionic strength of a solution) will result in high absorbance below 195 nm. The widely used solvent dimethyl sulfoxide (DMSO) absorbs strongly below 240 nm. In such cases, the concentrations of the absorbing species should be lowered by dialysis or gel filtration of the samples. To be able to estimate the secondary structure content of a protein from the far-UV CD, it is essential to know the protein concentration accurately (to within at least ±5%). The choice of method to determine the protein concentration has been discussed (Price, 1996; Kelly et al., 2005). The amount of protein required for CD measurements can be gauged from the need to keep the absorbance less than one unit. Depending on the design of the short-pathlength cuvette used (normally between 0.1 and 1 mm), an acceptable far-UV CD spectrum can be obtained using as little as 10 µg of protein, but generally between 100 and 500 µg of sample is required to determine the optimal conditions for recording spectra. For measurements in the near-UV or visible regions, several milligrams of protein are usually required. In favourable circumstances, good quality far-UV or near-UV CD spectra can be obtained within approximately 30 min. It is also worth noting that CD is essentially a nondestructive technique so that in most cases nearly all of the sample can be recovered after spectral data have been accumulated. However, in the case of synchrotron radiation CD, prolonged exposure of the sample to the intense radiation can lead to significant photo-damage.
(Kelly et al., 2005). See also: Protein Determination; Protein Secondary Structures: Prediction; Protein Structure Classification

Conformational Changes in Proteins

In solution, the overall conformation of a protein represents a time average of all the kinetically accessible states. The conformation can be influenced by a number of factors including temperature, pH, polarity of solvent, binding of ligands and others. CD is an extremely convenient technique for detecting and monitoring the extent of conformational changes that may be associated with the activity or regulation of a protein. The binding of Ca$^{2+}$ ions to calmodulin gives rise to substantial changes in both the far- and near-UV CD spectra of the protein. The changes in the far-UV indicate that binding of the metal ion leads to a significant increase in $\alpha$ helical content, estimated as being from 39% to 47% (Martin and Bayley, 1986). Addition of Ca$^{2+}$ ions to the adenylate cyclase toxin (CyaA) secreted by Bordetella pertussis leads to marked changes in the relative magnitudes of the negative minima in the far-UV CD at 208 and 222 nm. Although the secondary structure analysis programmes suggest that the overall helical content is not changed significantly by addition of metal ions, a more detailed analysis of the spectra suggests that isolated helical elements in the protein could be interacting to form coiled-coil structures in the presence of Ca$^{2+}$ (Cheung et al., 2009). The formation of chymotrypsin from the inactive precursor chymotrypsinogen is accompanied by only small changes in the far-UV CD spectrum, but much larger changes in the near-UV, which have been interpreted as reflecting changes in the environments of two of the eight Trp side-chains in the protein, that is, Trp 172 and Trp 215 (Woody and Dunker, 1996). The binding of molybdate to the molybdenum-sensing protein ModE from Escherichia coli does not lead to any marked change in the far-UV CD, but there are marked changes in the near-UV CD especially approximately 292 nm where the CD peak becomes much sharper (Figure 4). This is probably due to the changes in the orientation and mobility of the Trp 186 side-chain in each subunit of the dimeric protein on binding molybdate (Boxer et al., 2004). See also: Transition States: Substrate-induced Conformational Transitions

The Unfolding of Proteins

The most detailed studies of unfolding have been undertaken using chaotropic agents such as urea or guanidinium chloride (GdmHCl); unfolding by these agents is generally easier to control and interpret than thermal unfolding. CD measurements in different spectral regions can be used to determine the relative stabilities of the structural features of proteins. In favourable cases, such as ribonuclease T1, the unfolding curve can be analysed in terms of simple models (e.g. a two-state model involving only native and unfolded states), to arrive at a quantitative estimate of the stability in terms of the free energy required for unfolding of the native form. Studies of a number of globular proteins give surprisingly small values for $\Delta G^0$ (in the range +20–+ 50 kJ mol$^{-1}$), corresponding to the energy of only a few hydrogen bonds (Pace, 1990). The small values reflect the fine balance that exists between the enthalpic and entropic contributions to the equilibrium between native and unfolded states. See also: Cell Biophysics; Protein Denaturation and the Denatured State; Protein Stability; Protein Unfolding and Denaturants

It is now clear that a significant proportion (possibly as high as 30%) of eukaryotic proteins exist in their native state with some degree of disorder (Fink, 2005). These can be divided into two main categories according to whether the protein shows disorder along the entire length of the polypeptide chain or whether the protein has a short disordered region within a folded structure. The first group has been termed ‘natively unfolded proteins’ and can be further divided into those proteins which have significant amounts of secondary structure, but lack stable tertiary structure, and those which possess little or no secondary structure. The former subgroup have been described as ‘molten globules’ and were first characterized by Ptitsyn and his group in studies on the milk protein $\alpha$-lactalbumin at low pH (Ptitsyn, 1995). Disordered proteins are known to play a number of key roles in processes such as protein unfolding and folding, and imparting specificity to protein–ligand interactions (Fink, 2005). CD is an extremely useful technique in the study of disordered proteins, since the far-UV CD can report on the extent of secondary structure and the near-UV CD will indicate the extent of stable tertiary structure. Thus $\alpha$-synuclein which is a major component of the granular material in Lewy bodies which characterize Parkinson disease, appears to be natively unfolded at neutral pH, with little evidence for secondary structure from the far-UV CD. The protein can be induced to form fibrils with $\beta$-strands aligned perpendicular to the fibre axis by lowering the pH (Uversky et al., 2001). The
Ce3 domain of immunoglobulin E (IgE) interacts with key receptors to initiate the allergic response. When expressed as an isolated protein this domain has the characteristics of a ‘molten globule’ with significant secondary structure (from the far-UV CD), but no stable tertiary structure (from the near-UV CD) (Figure 5). Using CD together with NMR and fluorescence data, it was shown that on interaction with the high affinity receptor (sFcεRIa), the Ce3 domain adopts a folded tertiary structure (Price et al., 2005). See also: Fluorescence Spectrophotometry

The Folding of Proteins

Understanding the mechanism of protein folding represents one of the major challenges of molecular biology (Radford, 2000; Dobson, 2003, 2004; Chen et al., 2008). The folding of proteins involves specific pathways, since a random search of all possible conformational states would take an exceedingly long time, whereas in practice folding occurs on a time scale of seconds or minutes. Among the stages that occur in protein folding are: See also: Protein Folding In Vivo; Protein Folding: Overview of Pathways

- Formation of elements of recognizable secondary structure (helices, sheets, turns, etc.).
- Collapse of the extended polypeptide chain to form a more compact state in which nonpolar side-chains are largely buried from the solvent.
- Formation of long-range interactions that stabilize the native tertiary state.
- Association between subunits in multisubunit proteins.

Experimentally, the most amenable approach to gathering information on protein folding is to study the refolding of denatured proteins, using rapid dilution of the denaturing agent to initiate refolding. Stopped-flow CD measurements in both the far- and near-UV have been extremely important in obtaining information on the rate of regain of secondary and tertiary structures of proteins. Using experimental data from CD and other techniques such as NMR, fluorescence and small angle scattering alongside computer-based modelling, the in vitro folding pathways of several small proteins such as lysozyme, barnase and chymotrypsin inhibitor 2 have been defined in detail and some progress has been made in understanding how larger proteins fold (Radford, 2000; Dobson, 2004).

One fairly general finding is that the regain of native secondary structure (monitored by far-UV CD) occurs very quickly, often within the dead time of the stopped-flow CD technique, which is typically approximately 5 ms. By contrast, the regain of tertiary structure (monitored by near-UV CD) occurs more slowly, typically on the time scale of a few seconds. This rate is consistent with that of the formation of specific binding or catalytic sites in the protein. There is continuing debate about whether the formation of secondary structure precedes or follows the formation of a compact state, and indeed whether any one particular model is applicable to the folding of all proteins.

Further questions can be raised concerning the relevance of the studies on protein refolding to the mechanism of folding during biosynthesis. In the latter case, the involvement of cotranslational folding and the intervention of accessory factors such as molecular chaperones, peptidyl-proline isomerase and protein disulfide isomerase help to

![Figure 5](image.png) Figure 5  The far- and near-UV CD spectra of the Ce3 domain of immunoglobulin E. The main graph shows the far-UV CD spectrum and the inset shows (in red) the near-UV CD spectrum. The very small CD signal in the near-UV CD is indicative of the lack of stable tertiary structure.
ensure that protein folding is a rapid and highly efficient process with little or no byproduct such as aggregated protein. See also: Chaperones, Chaperonin and Heat-Shock Proteins; Proline Residues in Proteins; Protein Disulfide Isomerases

References


Further Reading


