Beginners guide to circular dichroism

Alison Rodger (Macquarie University, Australia)
Doug Marshall (Applied Photophysics, UK)

Circular dichroism (CD) is used to give information about the chirality or handedness of molecular systems. It is particularly widely applied to determine the secondary structure of proteins such as biopharmaceutical products.

Introduction

CD is used to give information about the chirality or handedness of molecular systems. CD is the difference in absorption of left and right circularly polarized light:

$$CD = A_l - A_r \quad (1)$$

CD and the related phenomenon of optical rotation of linearly polarized light were studied in the 19th century by scientists including Biot, Fresnel, Pasteur, and Cotton. However, because CD signals are very small differences between two larger numbers, with the difference typically being four orders of magnitude smaller than absorbances, it is really only in the last 50 years that electronics and instrumentation have enabled CD studies of solutions of molecules. CD is used to give information about the chirality or handedness of molecular systems.

Why CD is relevant

Applications for CD range from the search for extra-terrestrial life to 3D display technology, but with most emphasis on chiral molecules. The spectra of two enantiomeric (mirror image) molecules are equal in magnitude and opposite in sign (Figure 1). Within the life sciences, CD is most used for rapid comparison and characterization of protein secondary structures. It offers a way to determine the effect a mutation or a change in the environment of a protein (temperature, pH, ionic strength) might have on the overall structure, often prior to more resource-intensive characterization. This can avoid resources going to waste in expensive analyses of protein samples that, e.g., did not retain the native fold. In addition, CD is used to study protein-folding, stability and ligand binding and for characterization studies in the development of biotherapeutics. CD is also used in the study of nucleic acids and small biomolecules.

Challenges to acquiring a good CD spectrum

CD spectra are now no more difficult to obtain than the corresponding absorbance spectrum if one has access to a CD spectropolarimeter. CD signals can only be measured where there is absorbance, though the relative magnitudes of different parts of a CD spectrum differ from those of the sample's absorbance spectrum, and some CD bands are positive and some negative (Figure 1). This article outlines what a CD spectrum actually is, how to avoid at least some of the artefacts that result in quite a lot of published CD data being nonsense and what information can be extracted from a CD spectrum. The focus is on electronic CD, particularly of proteins, though the principles largely translate to any electronic or vibrational CD experiment.

As with all analytical tools, care should be taken to ensure that the CD spectropolarimeter is calibrated (intensity and wavelength – preferably not just a single-point calibration) and well maintained and that users have sufficient training to collect good quality data. Most instruments are supplied with instructions, and manufacturers readily support new users because their reputations depend on high-quality data from their systems being published. Camphor sulphonic acid is the primary standard for CD and its ammonium salt is a less hygroscopic alternative. More recently the CoEDDS of Figure 1, for which both enantiomers are available, has been used as it has the advantage of remaining stable in solution for years. Although one concentration (e.g., 4 mM) can be used across the visible and UV regions of the spectrum, better data are obtained if a higher concentration is used in the visible region.

Although equation (1) is the definition of CD, it is not actually what most instruments measure and plot. CD machines measure the ratio of an AC current corresponding to the difference in photon intensities, I, of left (l) and right (r) circularly polarized light reaching the detector (i.e., the light that is not absorbed) and a

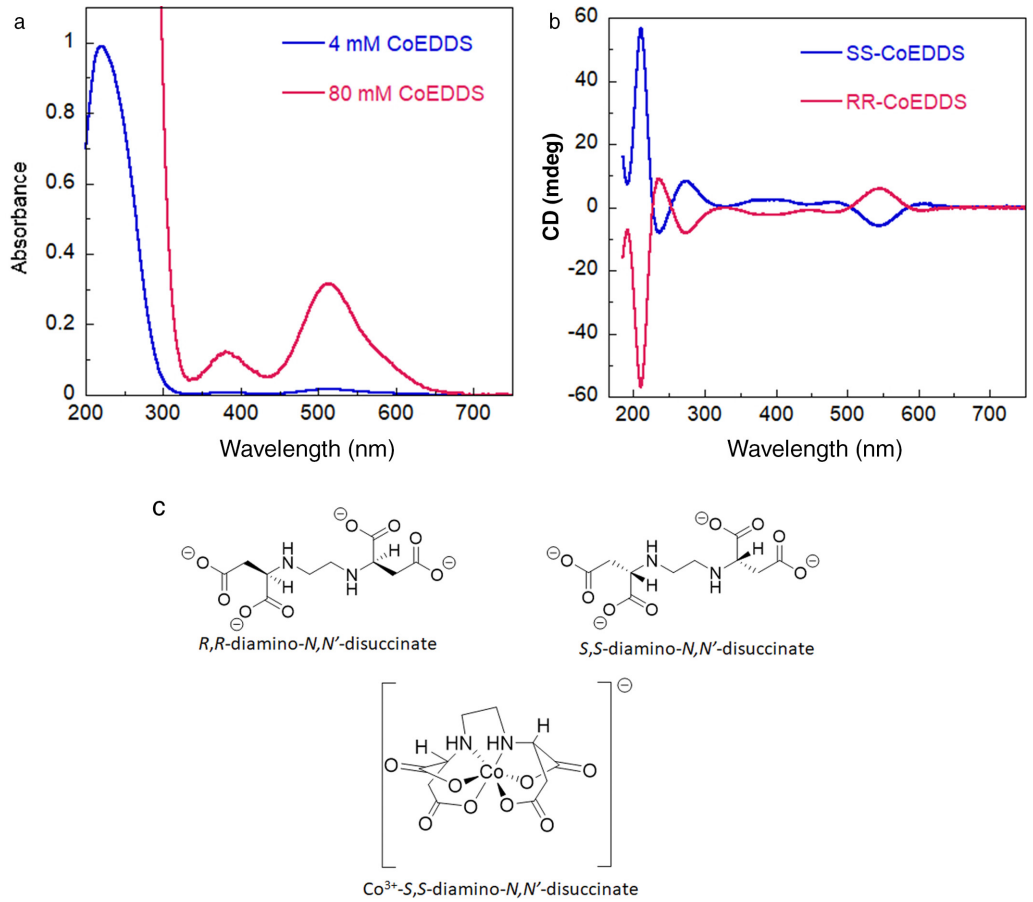


Figure 1. (a) Absorbance (4 and 80 mM in 1-cm path length cuvettes) and (b) CD spectra (0.072 mM in a 1-cm path length cuvette) of RR- and SS-CoEDDS (EDDS: ethylene diamine disuccinic acid). (c) Structures of RR- and SS-EDDS and the CoEDDS complex, SS-mer is illustrated.

DC current corresponding to the average of the two intensities.

$$\text{AC/DC} = \frac{2(I_l - I_r)}{(I_l + I_r)} \qquad (2)$$

The most common operating mode for CD machines is for a variable high-tension (HT) voltage to be applied to the photomultiplier so that the DC signal is kept constant. The software of the instrument then plots the approximation

$$\text{CD} = log_{10}\left(\frac{2+\text{AC/DC}}{2-\text{AC/DC}}\right) \approx \frac{\text{AC}}{\text{DC}} log_{10} e \qquad (3)$$

Consideration of these equations leads to some experimental precautions needed to ensure real data, not artefacts, are being measured. Instruments sometimes but not always provide an error flag for these issues.

(i) Equation (3) is an approximation, that requires AC<<DC.

(ii) The DC current must be accurately held constant (assuming one is working in a fixed-DC mode).

In practice, as long as the CD signal is less than 20% of the absorbance, the equation (3) approximation error is less than 1%. The approximation is therefore satisfactory for the CD of normal solutions of, e.g., biomolecules. More care is, however, required if one is using a CD spectropolarimeter for linear dichroism spectroscopy or to study samples such as cholesteric liquid crystals.

When the HT voltage is above about 600 V (i.e., a highly absorbing sample), too few photons are reaching the detector for the DC to be effectively held constant. This issue becomes more and more significant below 200 nm in most bench-top CD machines because the photon flux of the Xe arc lamp used as the light source only has 1% of its 500-nm flux. In older instruments, this is obvious because of high levels of noise. In newer instruments, attempts to enhance the signal:noise sometimes obscure this problem. Any stray light, including from holes in the sample compartment, of whatever frequency will simply be counted by the photomultiplier tube, thus artificially depressing the HT. If an instrument has the option of setting the CD scale, it may be advisable to use a scale orders of magnitude larger than the CD signal – at least to check that the CD signal is not being damped by the electronics of the instrument.

In this context it should be noted that achiral molecules absorb light – so, e.g., Cl^- (including that in phosphate-buffered saline) and other buffer components reduce the photons available for the analytes of interest to absorb. Proteins often need to have relatively high concentrations of buffers or ions present to maintain their stability. If the protein structure is concentration independent, an alternative to reducing buffer concentrations is to increase the protein concentration and reduce the path length to ensure the absorbance does not exceed 2. For path lengths less than 1 mm it is important to measure it for each cell and each user as path lengths of demountable cells are affected by the volume of sample used and how the cell is assembled.

Further artefacts can occur in any form of spectroscopy, including CD, due to heterogeneity of the sample. Consider an inhomogeneous sample where all the absorbing molecules are put in the left-hand half of the cuvette with an absorbance of, say, 1 ($A=log\left(I_0/I\right) = log\left(10/1\right)$). The 50% of photons passing through the right-hand side of the cuvette are all transmitted whereas those on the left are attenuated making the total transmission

$$\frac{I_0+I}{2I_0} \quad (4)$$

whose absorbance is

$$A=\log\left(\frac{2I_0}{I_0+I}\right) = \log\left(\frac{20}{11}\right) = 0.26 \quad (5)$$

whereas the absorbance of the sample after it has been thoroughly mixed is expected to be 0.5. This absorbance flattening also affects CD and can have a significant effect on, e.g., protein samples in membranes.

A final artefact that must be considered is that any light scattered by a sample is deemed by the spectrometer to have been absorbed – the instrument simply counts the photons that are incident on the photomultiplier tube and 'assumes' all others have been absorbed.

The above issues all mean that significant care must be taken when using a bench-top CD machine in the UV region of the spectrum – precisely where the CD of most biomolecules has useful data.

Interpreting a circular dichroism spectrum

Assuming that the above pitfalls have been avoided and one has measured a good CD spectrum, we ask what can be deduced from it. A non-zero spectrum tells you that the system being studied is chiral, with helical rearrangements of electrons during electronic transitions. For small molecules, fairly high-quality CD spectra can be calculated for known (or proposed) geometries and compared with experiment. Alternatively, if a molecule can be divided into achiral chromophores whose transition polarizations are known, then one of a number of models for calculating the CD spectrum may be appropriate and the handedness of the molecule may be able to be determined quite simply. For example, the coupling of electric dipole transition moments in the aromatic phenanthroline ligands of $[Ru(1,10\text{-phenanthroline})_3]^{2+}$ gives rise to the sharp in-ligand bands of Figure 2. The fact that the zero CD point corresponds with the absorbance maximum confirms that the CD is dominated by the coupled oscillator model. The d-d bands arise from a very different mechanism: the perturbation of magnetic dipole allowed d-d transitions by in-ligand transitions arranged in a helical geometry about the metal centre.

For biomacromolecules, such as proteins and DNAs, we usually know the chirality of the component amino acids or nucleosides and we are more interested in what the CD spectrum tells us about how they are arranged into secondary or even tertiary structures (Figure 3). The use one makes of a CD spectrum depends on what question is being asked: is an absolute measure of helix and sheet content of a protein required? Or does one simply want to know whether a peptide folds in a lipidic environment or a nucleic acid or protein unfolds when heated? The latter questions can often be answered by visual inspection of spectra (Figure 3). However, one does need to take care as to what information is being displayed. In Figure 3c, fluorescence tells us that the tryptophans in α-lactalbumin unfold in two phases, the first being slow and the second from 55°C to 65°C. In light of the fluorescence data, the aromatic CD (which is relatively noisy) is probably telling the same story, though it could be one broad transition. By way of contrast, the backbone protein spectrum measured at 222 nm unfolds in a single concerted step from 55°C to 65°C, at the same time as the environments of some of the tryptophans is exposed to water as shown by the fluorescence.

By far the most widespread use of CD spectroscopy is to identify protein secondary structure content. The success of CD structure-fitting is predicated on the observed spectrum being the sum of the contributions from different motifs within the protein. To get quantitative answers significantly better than an experienced spectroscopist can provide just by eye, it is essential to have a good reference set of spectra from proteins of known structure, preferably including proteins in some way related to the unknown one being studied. Various reference sets and algorithms have been collated on the Protein Circular Dichroism Data Bank and the Dichroweb site.

It has long been accepted that the further down in wavelength one measures, the more accurately more types of backbone structures can be extracted from CD spectra. However, some recent Bayesian analysis has reinforced the value of collecting data to at least 190

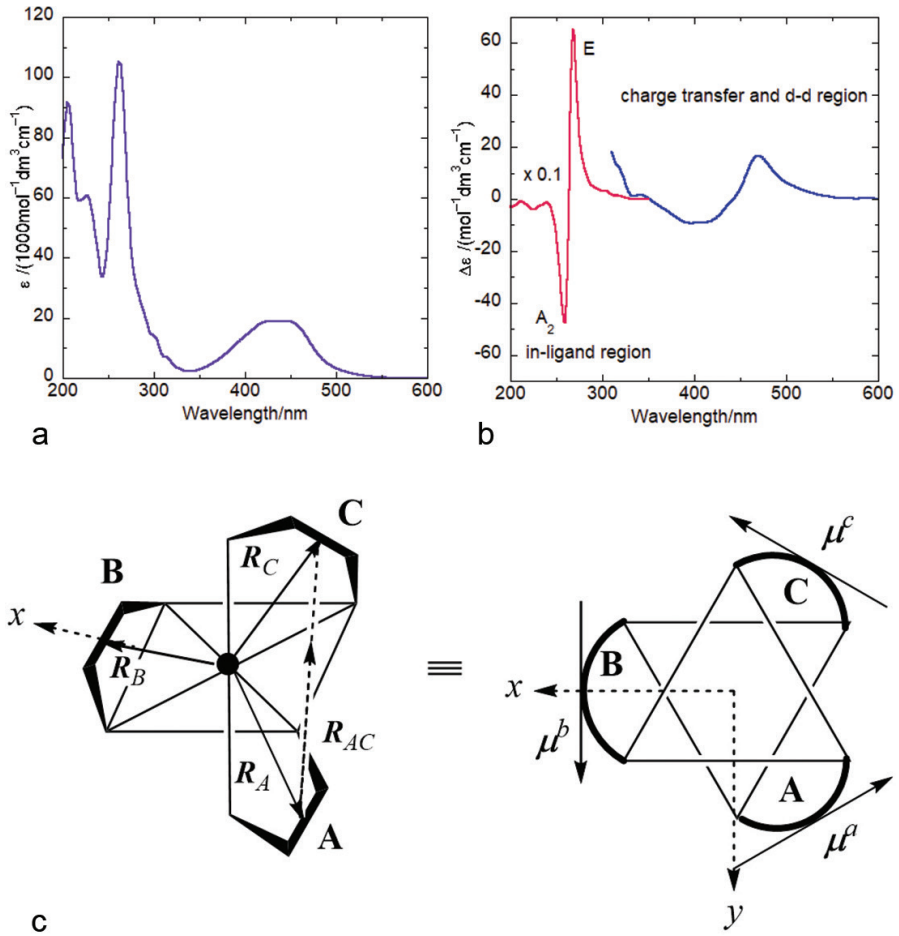


Figure 2. (a) Molar absorbance and (b) CD spectrum of Λ-[Ru(1,10-phenanthroline)$_3$]$^{2+}$ in a 1-cm path length cuvette and (c) schematic of the electric dipole transition moments that couple to give the dominant in-ligand CD signal.

nm but made it clear that only three types of structure (a collation of helical structures, a collation of sheet structures and 'other') can be extracted from a normal reference set even if it has data down to 175 nm. In addition, the similarity of unfolded, polyproline II and β_{II} spectra (Figure 3a) means that care must be taken to investigate the reference spectra selected by any algorithm for its final fit.

Our own self-organizing map structure-fitting approach (SOMSpec) has the advantage of enabling the user to interrogate the output to identify which reference proteins have been used to estimate the secondary structure of the unknown. This helps an analysis of the quality of the secondary structure estimate to be made and makes it possible to look for the source of any anomalies. For example, we know that lysozyme is significantly unfolded at 100°C (Figure 4a) but a straightforward fit with the SOMSpec algorithms (and other) suggests it increases its β-sheet content from 16% at room temperature to 34% at 100°C (with no evidence of, e.g., amyloid fibre formation). The reason proves to be that the selected best matching units include α-chymotrypsin and chymotrypsinogen, both of which Sreerama and Woody categorize as β_{II} proteins. The CD significance of this is that although the best matching proteins have β_{II} structure, the spectrum of β_{II}s look like random coil spectra so the 34% β-sheet assignment at 100°C is misleading and should be 34% unfolded. This shuffling of assigned content between β-sheet and random coil always needs to be considered.

Some recent applications of CD spectroscopy

As a general rule CD is applied as one of a number of techniques to characterize complex molecules or molecular assemblies with chiral components. This is illustrated with three examples.

Biosimilars

Biosimilar products are usually a different company's attempts to produce a biopharmaceutical that is 'highly

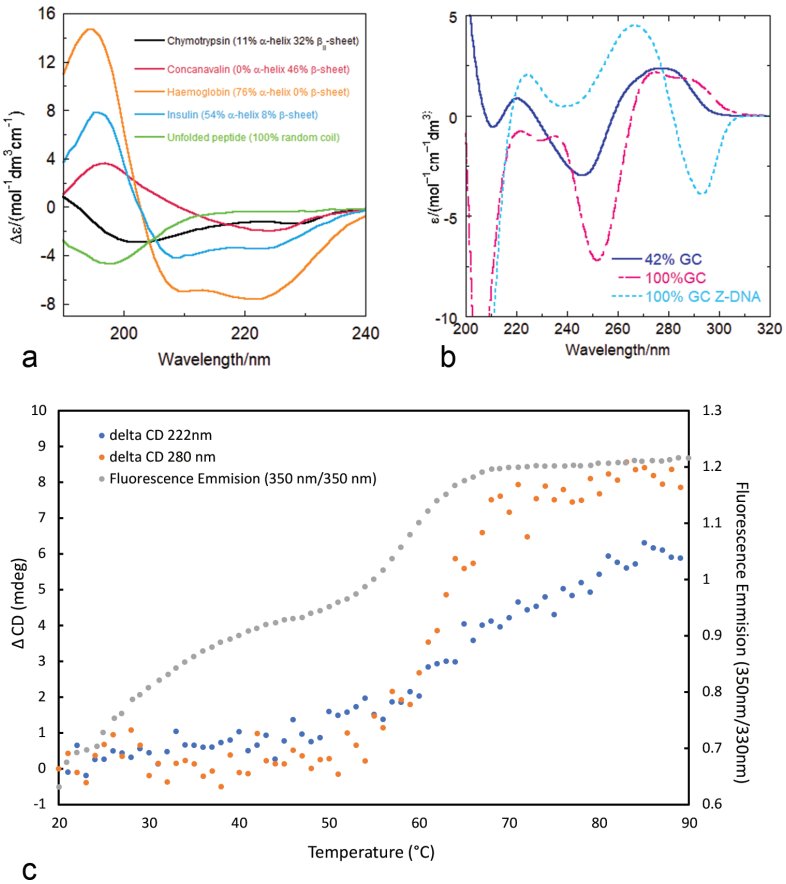


Figure 3. (a) Far UV (below 250 nm) CD spectra for a selection of proteins of different secondary structure content. (b) CD spectra for some DNAs with differing GC content. The Z-DNA was formed by adding spermine. (c) Melting curves of α-lactalbumin measured by CD (at two wavelengths) and fluorescence (presented as the ratio of 350 nm to 330 nm intensities).

similar' to another already approved biological medicine whose patent has expired. Unlike small-molecule generic drugs, biosimilars require not only the primary structure (what is bonded to what), but also secondary (local folds), tertiary and even quaternary structure similarity. Also, unlike small molecules, some differences at the atomic level, especially with glycosylation patterns, may be allowed. Comparisons of far UV CD (from 250 nm down to at least 190 nm) is generally deemed to be a good indicator of whether the secondary structure content is changing. For many proteins, the near UV region (from 300 nm to 250 nm) can be used to give an indication of

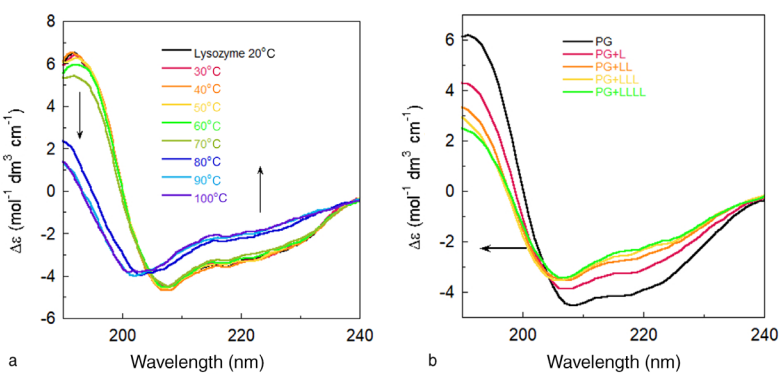


Figure 4. (a) CD spectrum of lysozyme as a function of temperature. (b) CD spectrum of the 96-residue protein G (PG) and of PG with the addition of 1, 2, 3 or 4 copies of a 21-residue linker peptide.

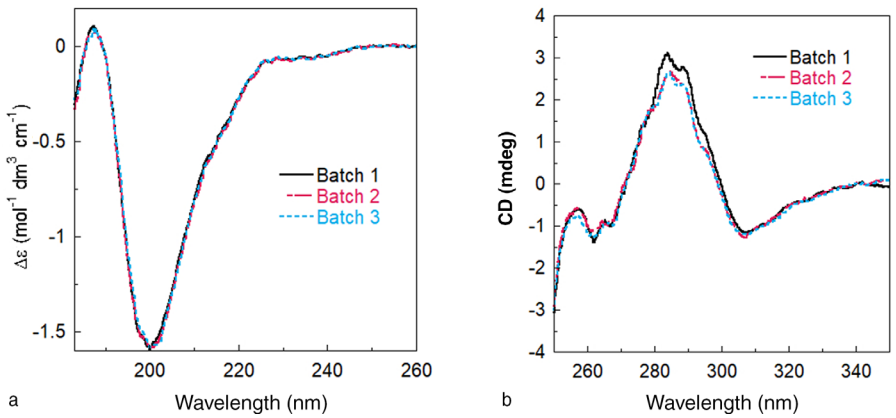


Figure 5. Baseline-corrected and zeroed CD spectra of three batches of a biopharmaceutical protein product in the far UV (a) and near UV (b).

tertiary structure – namely how the secondary structure units assemble. The near UV region is often referred to as the aromatic region – though it includes disulphide-bond contributions and seldom shows any contribution from phenylalanines. So, the near UV is dominated by tryptophan signals and shows contributions from tyrosines.

An example of batch-to-batch CD comparison of a biopharmaceutical product is given in Figure 5. Visual inspection of the spectra suggests it is largely unfolded (negative maximum at 200 nm), though CD structure-fitting suggests it is β-sheet. The resolution of this, as discussed above, is that it is probably a combination of β_{II} and unfolded structures. Close inspection of the data suggests there is a small difference in the secondary structure of three batches of the product. The aromatic CD suggests the tertiary structures of the different batches also differ.

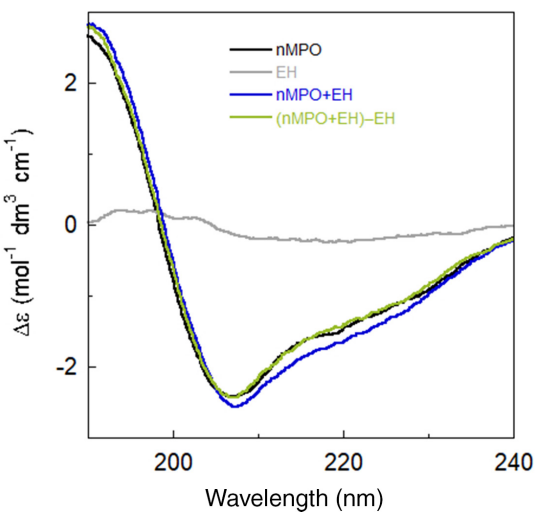


Figure 6. CD of nMPO and EH-treated nMPO, and the theoretical spectrum of EH-treated nMPO with EH subtracted. It should be noted that all spectra were baseline corrected by subtracting a buffer spectrum before EH was subtracted from EH-treated nMPO.

Identification of unfolded domains in proteins

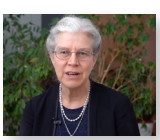
Lysozyme CD as a function of temperature (Figure 4a) illustrates a relatively sharp transition from the fully folded protein (with 39% α-helix and 16% β-sheet) to a largely unfolded protein (11% α-helix and 11% β-sheet). Similar profile differences are observed for the CD spectra of a series of proteins composed of a truncated form of a *Streptococcus* protein G linked to increasing numbers of a silica-binding peptide (Figure 4b). A more quantitative analysis indicates that the added peptides are unfolded as are a few N-terminus residues of protein G.

Neutrophil granule myeloperoxidase

Myeloperoxidase (MPO) is an important glycoprotein which plays a role in neutrophil-mediated immunity. The CD of neutrophil-derived MPO (nMPO) before and after treatment with Endoglycosidase H (EH), as judged by comparing the nMPO spectrum and the difference between EH-treated nMPO and EH, suggests there may be a small difference between the two proteins, but Figure 6 is not entirely convincing. However, the simple experiment of measuring CD as a function of temperature shows that the treated nMPO is much more stable (beginning to melt at 72°C) than untreated nMPO (beginning to melt at 60°C). Melting experiments can be done by measuring full wavelength scans at temperature intervals or by selecting a key wavelength (such as 208 nm for α-helical proteins) to measure the signal as the temperature is changed. Some CD instruments have an option to measure both in one experiment. ■

Further reading

- Bansal, R., Elgundi, Z., Goodchild, S.C., et al. (2020) The effect of oligomerization on a solid-binding peptide binding to silica-based materials. Nanomaterials, **10**, 1–15. DOI: 10.3390/nano10061070
- Berova, N., Nakanishi, K. and Woody, R.W. (eds.) (2000) *Circular dichroism principles and applications.* Wiley-VCH, New York
- Berova, N., Polavarapu, P.L., Nakanishi, K. and R.W. Woody (eds.) (2014) *Comprehensive Chiroptical Spectroscopy* Wiley-VCH, New York
- Coggan, D.Z., Haworth, I.S., Bates, P.J., et al. (1999) DNA binding of ruthenium tris-(1,10-phenanthroline): evidence for the dependence of binding mode on metal complex concentration. *Inorg. Chem*., **38**, 4486–97. DOI: 10.1021/ic990654c
- Goodchild, S.C., Jayasundera, K. and Rodger, A. (2019) 'Circular dichroism and related spectroscopic techniques.' In *Biomolecular and bioanalytical techniques: theory, methodology and applications*, Ramesh, V. (ed.) John Wiley & Sons, Hoboken
- Hall, V., Sklepari, M. and Rodger, A. (2014) Protein secondary structure prediction from circular dichroism spectra using a self-organizing map with concentration correction. *Chirality*, **26**, 471–82. DOI: 10.1002/chir.22338
- Hall, V., Nash, A. and Rodger, A. (2014) SSNN, a method for neural network protein secondary structure fitting using circular dichroism data. *Anal. Methods*, **6**, 6721–6726. DOI: 10.1039/C3AY41831F
- Hiort, C., Nordén, B. and Rodger, A. (1990) Enantioselective DNA binding of $[Ru^{II}(1,10\text{-phenanthroline})_3]^{2+}$ studied with linear dichroism. *J. Am. Chem. Soc*., **112**, 1971–1982. DOI: 10.1021/ja00161a050
- Nordén, B., Rodger, A. and Dafforn, T.R.. (2010) *Linear dichroism and circular dichroism: a textbook on polarized spectroscopy*. Royal Society of Chemistry, Cambridge.
- Ravi, J., Schiffmann, D., Tantra, R. et al. (2010) International comparability in spectroscopic measurements of protein structure by circular dichroism: CCQM-P59. *Metrologia*, **47**, 08022. DOI: 10.1088/0026-1394/47/1A/08022
- Schipper, P.E., and Rodger, A. (1986) Generalized selection rules for circular dichroism: A symmetry adapted perturbation model for magnetic dipole allowed transitions. *Chem. Phys*. **109**, 173–193. DOI: 10.1016/0301-0104(86)87050-1
- Spencer, S.E.F., and Rodger, A. (2020) Bayesian inference assessment of protein secondary structure analysis using circular dichroism data – how much structural information is contained in protein circular dichroism spectra? *Anal. Methods*. **13**, 359–368. DOI: 10.1039/d0ay01645d
- Sreerama, N., and Woody, R.W. (2003) Structural composition of β_I- and β_{II}-proteins. *Protein Sci*. **12**, 384–388. DOI: 10.1110/ps.0235003
- Sutherland, J.C. (2002) Simultaneous measurement of circular dichroism and fluorescence polarization anisotropy. *Proc. SPIE 4625, Clinical Diagnostic Systems: Technologies and Instrumentation*. **126**, 126–136. DOI: 10.1117/12.469782
- Tjondro, H.C., Ugonotti, J., Kawahara, R. et al. (2020) Hyper-truncated Asn355- and Asn391-glycans modulate the activity of neutrophil granule myeloperoxidase. *J. Biol. Chem.* DOI: 10.1074/jbc.RA120.016342
- Whitmore, L., and Wallace, B.A. (2004) DICHROWEB: an online server for protein secondary structure analyses from circular dichroism specroscopic data. *Nucleic Acids Res*., **32**, W668–673. DOI: 10.1093/nar/gkh371
- Whitmore, L., Woollett, B., Miles, A.J. et al. (2011) PCDDB: the protein circular dichroism data bank, a repository for circular dichroism spectral and metadata. *Nucleic Acids Res*. **39**, D480–D486. DOI: 10.1093/nar/gkq1026



Alison Rodger began her career at the University of Sydney and after nearly 30 years in the UK at the Universities of Cambridge, Oxford and Warwick moved to Macquarie University in 2017. Her research focuses on understanding the structure and function of biomacromolecules and their assemblies and often requires new techniques. Email: alison.rodger@mq.edu.au

Doug Marshall began his research career as a postdoc at UCL in 2007 before moving on to University of Pennsylvania and University of Essex. He specialized in understanding the structure and function of respiratory chain proteins. Doug now works as product manager for CD technology at Applied Photophysics.

