

Synchrotron radiation circular-dichroism spectroscopy as a tool for investigating protein structures

B. A. Wallace

Copyright © International Union of Crystallography

Author(s) of this paper may load this reprint on their own web site provided that this cover page is retained. Republication of this article or its storage in electronic databases or the like is not permitted without prior permission in writing from the IUCr.

Lead Article

J. Synchrotron Rad. (2000). **7**, 289–295

Synchrotron radiation circular-dichroism spectroscopy as a tool for investigating protein structures

B. A. Wallace

Department of Crystallography, Birkbeck College, University of London, London WC1E 7HX, UK, and Centre for Protein and Membrane Structure and Dynamics, Daresbury Laboratory, Warrington WA4 4AD, UK. E-mail: ubcg91c@mail.cryst.bbk.ac.uk

(Received 12 May 2000; accepted 26 June 2000)

This paper reviews the use of synchrotron radiation circular dichroism (SRCD) spectroscopy for examining protein structures, discussing the differences between conventional circular dichroism and SRCD, as well as examples of what SRCD studies have revealed about protein structures to date. It further discusses the future potential of the technique, including roles in structural genomics, membrane protein structure elucidation, relationships to crystallographic studies, and protein folding and dynamics.

Keywords: circular dichroism spectroscopy; protein secondary structures; membrane proteins; structural genomics.

1. Introduction: circular dichroism spectroscopy

Circular dichroism (CD) is a measure of the difference in absorption of left- and right-circularly polarized light as it passes through an optically active or chiral sample. Optical activity arises due to the influence of neighbouring groups on the electronic structure of a chromophore involved in an electronic transition.

Electronic transitions span a large wavelength range, including the near UV (from 300 to 250 nm), the far UV (from 250 to 190 nm) and the VUV (below 190 nm). Both the near- and far-UV wavelength ranges are readily measurable on conventional (commercial laboratory-based) CD spectrophotometers (spectropolarimeters).

For proteins, the most important electronic transitions are the $n \rightarrow \pi^*$, the $\pi \rightarrow \pi_{||}^*$, and the $\pi \rightarrow \pi_{\perp}^*$ of the polypeptide backbone amide group. The resulting peaks fall in the far-UV region of the spectrum, centred at ~ 220 nm, 210 nm and 195 nm, respectively. Other polypeptide backbone transitions (including the $n \rightarrow \sigma^*$ and the $n' \rightarrow \pi^*$) exist at lower wavelengths in the VUV region, but have generally not been detectable by conventional CD instruments. Additional transitions resulting from aromatic and disulphide side chains are found in the near-UV region of the spectra. These can be useful for monitoring environmental effects, but are generally of considerably lower amplitude than the backbone transitions, and the underlying molecular bases of these transitions are not as well established. In coloured proteins, chromophores with transitions in the visible region of the spectrum may be valuable for monitoring tertiary and quaternary structural

features. This review will concentrate on studies of the spectroscopic characteristics and structural interpretations that can be made based on the far-UV and VUV spectra of polypeptide backbone transitions obtained by conventional CD and synchrotron radiation circular dichroism (SRCD) spectroscopy.

CD has found its principal application in studies of protein secondary structures (for reviews, see Johnson, 1990; Woody, 1995). This is because the shape (waveform) of a CD spectrum is related to the geometry (effectively the ϕ/ψ angles) of the polypeptide backbone. Thus, each type of secondary structure has a unique spectrum. The spectral characteristics of α -helices include a large-magnitude positive peak at ~ 195 nm, and two almost equivalent negative peaks of roughly half the magnitude, centred around 208 and 222 nm. For β -sheets, the positive peak at ~ 192 nm is approximately half the magnitude of the helical peak in this region of the spectrum, and there is a small negative peak at ~ 212 nm. Various types of turns, polyproline helices and what was originally termed 'random coil', but is better described as 'other' or aperiodic structures, also exhibit characteristic far-UV peaks. Because these individual spectra are independent and additive, the spectrum measured for a protein is a linear combination of the characteristic spectra of the various structural types present weighted by their fractional representation in the protein. As a result, it is possible to empirically determine secondary structural compositions from experimental CD spectra by decomposition processes. A variety of different types of algorithms for deconvoluting spectra into their component contributions have been developed. These

include least squares (both linear and non-linear) (Chen & Yang, 1971; Wallace & Teeters, 1987), singular-value deconvolution (Hennessey & Johnson, 1982), variable selection (Manavalan & Johnson, 1987), parameterized fit (Provencher & Glockner, 1981), convex constraint (Perzel *et al.*, 1992), neural networks (Bohm *et al.*, 1992), and self-consistent methods (Sreerama & Woody, 1993). Most methods quantitate the percentages of helix, sheet, turn and random (or aperiodic) structures present. All the empirical methods in use to date are most accurate at defining helical structures from CD spectra because of the dominance of the characteristic peaks from helices (they are much larger than the β -sheet contributions) and because the structural nature of helices mean their φ/ψ angles fall in a much narrower range than do those of other types of secondary structures and as a result they are much more well defined.

CD spectroscopy has also been very valuable for detecting conformational changes in proteins, as the spectra are sensitive to small alterations in polypeptide backbone structures. Consequently, the technique has been used extensively for both qualitatively and quantitatively examining the binding of ligands to proteins, as well as for deciphering the nature of interactions between proteins and other macromolecules. Again due to the spectral sensitivity to conformation, the technique has found utility in studies examining environmental effects on the structure of proteins. These types of studies have included the effects of solvents, pH, concentration, ionic strength and metal ions, amongst others. For membrane proteins, the structural consequences of binding to and inserting into phospholipid bilayers have also been examined. The advantages that CD spectroscopy has over other techniques for monitoring conformational changes include the small quantities and low concentrations of protein necessary, the absence of 'probe' molecules which could distort the effects, and the ease and speed at which the experiments can be conducted. Furthermore, such changes can be interpreted on a molecular level in terms of specific types of alterations in the polypeptide backbone.

More recently, CD has played a major role in defining protein folding pathways and rates of folding, especially through the use of stopped-flow techniques (for a review, see Kelly & Price, 1997). It has been possible to monitor the acquisition of secondary structure as a function of time, enabling the definition of rates of formation of different secondary structural types and, in conjunction with other techniques, the sequence of events in the folding pathway (including molten globule formation). Recently, such studies have been extended to membrane proteins (Riley *et al.*, 1997) to permit examination of their folding in the environment of a lipid bilayer. Most CD folding studies to date have tended to monitor the peaks near 220 nm that are primarily associated with helix formation. Measurements at low wavelengths in the far-UV region have been limited because of the flux available in conventional CD instruments. However, the data in the region around 190 nm are

potentially very valuable as they can provide information on sheet and aperiodic structures, and can complement the higher-wavelength data, which tend to be dominated by helix effects.

2. Spectropolarimeters – CD and SRCD instruments

Commercial instruments for measuring CD have been available since the 1960s. The light sources in modern instruments, xenon arc lamps, have typical fluxes at 240 nm of $\sim 10^{11}$ photons s^{-1} , but their intensity drops off rapidly at lower wavelengths (to $\sim 10^9$ photons s^{-1} at 180 nm), thereby limiting the effective wavelength range over which spectra can be measured. The light becomes linearly polarized by passing through an MgF_2 or biotite polarizer. It is converted to left- and right-circularly polarized light using a piezoelectric crystal modulator (PEM) with a frequency of 50 kHz. A photomultiplier detector is locked onto the modulator frequency to alternately measure the left- and right-handed signals. Because of the absorption of oxygen in the low-wavelength region, all the chambers of the instrument are flushed with dry nitrogen. To obtain data in the VUV region, the modulator is placed in a vacuum environment to remove all gases from this part of the optical path. Restrictions due to absorption within the optical elements of the instruments, as well as the lamp flux, limit the wavelength range that can be attainable on a conventional CD instrument. The photomultiplier windows can therefore be made of fused silica (quartz), as components other than these tend to be the limiting factors on the low-wavelength end.

A number of significant differences in design exist between conventional CD and SRCD instruments (Sutherland, 1996). The synchrotron radiation light is inherently linearly polarized, thereby eliminating the need for a polarizer. The wavelengths of interest are at the low-energy end of synchrotron radiation. The flux obtainable in the UV region from second-generation sources is $\geq 10^{11}$ photons s^{-1} at 240 nm, and is almost constant from 240 nm down to about 140 nm. On a third-generation source the use of a long-period undulator on a long straight could improve the flux (and the polarization) even more over the range of wavelengths of interest. In the future, free electron lasers, being bright coherent sources, may have potential as alternative light sources, although wavelengths in the VUV region below 200 nm have yet to be achieved. The choice of optical elements is important in SRCD, as the flux is, in general, not the limiting factor on the wavelengths measurable. Using a photoelectric modulator made of CaF_2 limits the low-wavelength range to ~ 130 nm, but with an LiF crystal this can be extended to ~ 110 nm. The use of LiF or MgF_2 for windows of the photomultiplier, or its replacement altogether with a fluorescent screen, means that this too will not be limiting on the light transmission at low wavelengths. The principal restriction to the wavelength range in SRCD is associated with the sample: the material

Table 1
Characteristics of existing SRCD facilities.

Location	SRS, Daresbury, UK	NSLS, Brookhaven, USA
Stations	3.1, 13.1b	U9B, U11
Website	http://www.dl.ac.uk/SRS/VUV-IR/CD/cpmsd.html	http://bnlslb.bio.bnl.gov/biodocs/nsls/u9b/u9b.html
Wavelength range (nm)	130–350 (3.1); 190–1000 (13.1b)	160–1500 (U9B); 120–300 (U11)
Windows	LiF and CaF ₂ PEM (3.1); Spectrosil (13.1b)	CaF ₂ and CaF ₂ PEM (U9B); LiF or CaF ₂ PEM (U11)
Maximum flux (photons s ⁻¹)	4 × 10 ¹¹ @ 180 nm; 3 × 10 ¹¹ @ 240 nm (3.1) 7 × 10 ¹¹ @ 200 nm; 5 × 10 ¹² @ 240 nm (13.1b)	2 × 10 ¹² @ 160 nm; 9 × 10 ¹² @ 240 nm (U9B) 3 × 10 ¹² @ 160 nm; 2.5 × 10 ¹² @ 240 nm (U11)
Degree of polarization	0.8 (3.1); 0.9 (13.1b)	
Minimum sample volume	<10 µl	≤ 4 µl
Temperature range	258–363 K	263–363 K
Stopped flow	Yes, 400 × 10 ⁻⁶ s (13.1b)	No
Operation	Since 1995 (3.1); 1996 (13.1b)	Since 1980

Table 2
Characteristics of planned SRCD facilities.

Location	SRS, Daresbury, UK	ASTRID, Aarhus, Denmark
Stations	12.1 (CPMSD)	UV1
Website	http://www.dl.ac.uk/SRS/VUV-IR/CD/cpmsd.html	http://www.isa.au.dk/SR/sr-facilities.html
Wavelength range (nm)	110–1000	130–450
Windows	LiF and CaF ₂ PEM	CaF ₂
Maximum flux	5 × 10 ¹³ photons s ⁻¹ @ 180 nm	10 ¹⁰ photons s ⁻¹ @ 180 nm
Degree of polarization	0.7–1.0	0.9
Minimum sample	<10 µl	~5 µl
Temperature range	258–363 K	273–393 K
Stopped flow	Yes	Planned
Operation	Planned for March 2001	Planned for August 2000 to operate for six months per year

that the sample cell (quartz) is made of, and the buffers, salts and solvents (including water) present in the solution. An important feature is that the spot size of the illumination on the sample in an SRCD instrument is only ~2 mm², thereby making the sample size requirement considerably smaller than that of a conventional CD instrument.

3. SRCD – present facilities/facilities in development

Two sites, the SRS at Daresbury Laboratory and the NSLS at Brookhaven National Laboratory, each currently have two operational SRCD instruments for studying proteins (Table 1). One station at each location is principally for steady-state (wavelength scan) measurements (3.1 at SRS, U9B at NSLS) and one station at each site is primarily designed for time-resolved (stopped-flow) measurements (13.1b at SRS, U11 at NSLS); the lower wavelength ranges of the latter stations may be more restricted than the former, depending on the optical elements and the materials necessary to produce mechanically strong windows for the mixing chambers.

Two new SRCD instruments are planned to come on-line in the next year or so; one, station 12.1 (Clarke *et al.*, 2000), will replace the existing instruments at the SRS, and the other will be a new site at ASTRID (Table 2). Both will permit steady-state measurements and stopped-flow studies and will have high flux, especially at low wavelengths. The site at the SRS is part of the new Centre for Protein and Membrane Structure and Dynamics (CPMSD), funded by the BBSRC Structural Biology Centres programme.

4. SRCD – advances/advantages over CD

4.1. Wavelength range

Published far-UV CD spectra from conventional instruments generally include the wavelength range from ~240 nm down to ~190 nm (Fig. 1). With a well tuned and maintained commercial CD instrument, it is possible to collect data in aqueous solutions in the VUV region to as low as 178 nm (Johnson, 1996), provided the buffer and concentration conditions are suitable and the instrument is well flushed with nitrogen, but this is rarely achieved in practice. On the SRCD at Daresbury (station 3.1), data can be readily obtained down to 160 nm for aqueous solutions of proteins (Fig. 1). Following the suggestion of Sutherland *et al.* (1986) to use D₂O instead of H₂O, it has been possible

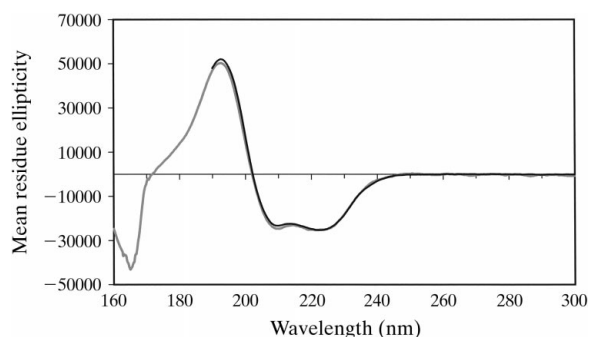


Figure 1
SRCD (grey line) and conventional CD (black line) spectra of horse myoglobin in aqueous solution. In both cases a sample concentration of 7 mg ml⁻¹ in a cell of pathlength 0.001 cm was used. The SRCD spectrum was obtained on station 3.1 at the Daresbury Laboratory over the wavelength range from 160 to 300 nm, and clearly shows the additional information content present at low wavelengths.

to obtain data as low as 158 nm for aqueous protein solutions at station 3.1 (data not shown). The significance of being able to obtain very low wavelength data is that there are at least two additional transitions which exist below 190 nm [one shoulder at ~ 175 nm that has been attributed to an $n \rightarrow \sigma^*$ transition of the amide carbonyl group (Kallenbach *et al.*, 1996) and the other a negative peak centred at ~ 165 nm]. Although the transition at the higher wavelength has been seen in non-aqueous solutions and films (where solvent absorbance is less) with conventional instruments, neither of the two low-wavelength peaks can be detected accurately in aqueous solution with a conventional instrument. Both of these low-wavelength transitions are highly sensitive to the polypeptide backbone conformation, and could be very valuable in more accurately distinguishing between different types of secondary structures. For example, helices appear to exhibit a pronounced positive shoulder at ~ 175 nm, with a crossover at ~ 170 nm and a negative peak well below 170 nm. In contrast, β -sheets produce a crossover nearer 185 nm and a relatively intense negative peak in the region between 170 and 180 nm. These waveform patterns are much more clearly different than are the higher-wavelength peptide transitions for these types of secondary structures. Inclusion of this wavelength region in reference databases could therefore provide for better accuracy in determination of sheet and aperiodic structural content.

4.2. Sensitivity

The increased intensity of a synchrotron radiation light source (especially at low wavelengths) relative to that of a xenon arc lamp means that the signal-to-noise (S/N) ratios for SRCD measurements will be considerably higher in the far-UV and VUV region where the peptide transitions occur. For example, although at high wavelengths (~ 240 nm) the fluxes of SRCD and CD instruments may be roughly comparable, at 180 nm the flux for an SRCD may be as much as 10^4 times that of a conventional instrument (Tables 1 and 2). A consequence of this is that lower concentrations of protein can be used in order to obtain comparable spectral qualities in SRCD. Alternatively, shorter-pathlength sample cells may be used (0.001 cm or less), thereby decreasing the problem of non-chiral absorption due to the solvent or buffer, and enabling the collection of lower-wavelength data.

4.3. Information content

The acquisition of data to lower wavelengths means that the SRCD spectra will be richer, containing more eigenvectors of information than conventional CD spectra owing to the additional electronic transitions measured. Consequently, it should be possible to analyse SRCD spectra for a larger number of distinct components or folding motifs. This means that instead of just helix, sheet, turn and random (or aperiodic) structures, such features as helical type (α , 3_{10} and π), parallel and antiparallel β -sheets, and different types of turns *etc.* may be distinguished. Perhaps

as many as seven different types of structure may be identified from each spectrum.

The CD reference data basis sets that are currently in use for the analyses of protein secondary structures (Sreerama & Woody, 2000) include many derived from spectra collected 20 or more years ago (Chen *et al.*, 1972; Chang *et al.*, 1978; Bolotina *et al.*, 1980; Brahms & Brahms, 1980; Hennessey & Johnson, 1982). At that time, not only was the range of spectroscopic data that could be collected more limited, but also the variety of proteins whose X-ray structures were then known was much more restricted. In many cases the reference databases include only ~ 20 proteins, and in some cases the number is as small as 5. Clearly these small samplings cannot cover the range of secondary structures and structural motifs present in all the possible protein folds. Now that the number of entries in the Protein Data Bank has exceeded 12000 (as of 1 April 2000), there are many more examples of different structural types that could be used to produce new spectral basis sets. Since empirical analyses of this sort can only accurately define structural types that were present in proteins that were used to create the original basis sets, the more features present in the proteins used to create the basis sets, the better defined will be the calculated structures of new proteins examined by CD spectroscopy. Combined with the larger spectral range now possible with SRCD, this means the analyses should be able to produce more detailed secondary structural information than has been possible to date with conventional CD spectroscopy.

5. SRCD – examples of use to study proteins

Although SRCD instruments have existed since 1980 (Sutherland *et al.*, 1980; Snyder & Rowe, 1980), it has not been until relatively recently that they have provided useful information on protein structure and folding. In the past, the method has been more effectively exploited in studies of nucleic acids and small molecules (Sutherland & Griffin, 1983; Sutherland *et al.*, 1986; Pulm *et al.*, 1997), where the VUV transitions exhibit less overlap than those in the far-UV region.

In the early 1990s it was reported that SRCD could provide information down to wavelengths of ~ 175 – 178 nm for aqueous samples (Garone *et al.*, 1990; France *et al.*, 1992; Sutherland *et al.*, 1992) and, although this was an improvement on commercial instruments at the time, it did not necessarily provide sufficient amounts of extra data to be worth the extra effort and expense, except in a few special cases. By the late 1990s, however, it became clear that data to as low as 170 nm was readily (routinely) obtainable and that the additional information included in this data could be very useful for the analysis of secondary structure (Qi *et al.*, 1997; Symmons *et al.*, 1997; Clarke & Jones, 1999), for stopped-flow kinetic studies of protein folding (Clarke *et al.*, 1999) and for binding studies (Chang *et al.*, 1999).

The SRCD studies undertaken to date which have shown the 'added value' of the low-wavelength data most clearly are those that have involved investigations of proteins with high contents of β -sheet, polyproline II and/or aperiodic structures. Because these types of secondary structures have distinct spectra at low wavelengths where the measured dichroism is not dominated by helical contributions, SRCD has provided important information that could not have been obtained by conventional CD. For example, in a study of the outer surface protein of the Lyme disease spirochete (France *et al.*, 1992), which included data down to 175 nm and used a very short pathlength cell so that high concentrations of denaturants could be used, the secondary structure of this primarily β -sheet protein could be defined. In addition, conformational changes associated with pH and thermal and chemical denaturation were much better defined as a result of the higher S/N. Similarly, improved results were seen in studies examining the unfolding of β -lactoglobulin, another protein dominated by β -sheet structures (Qi *et al.*, 1997). Temperature-dependence studies, which included SRCD data to 170 nm, clearly showed that independent processes are involved in the loss of helical, sheet and turn conformations. Using conventional CD, the sheet and turn components would have been effectively masked by the helical changes. As a final example, in a paper examining the formation of a ternary complex between DNA, a DNA-binding protein and an activator protein (Chang *et al.*, 1999), SRCD provided important additional information beyond what would have been obtainable by conventional CD: the activator protein is largely composed of polyproline II helix and β -sheets, structures not well defined at wavelengths above 190 nm. The additional data available from SRCD between 170 and 190 nm permitted more accurate secondary structure determination and, most importantly, the ability to monitor complex formation. In addition to these studies on β -sheet-rich proteins, it is notable that, by using SRCD to re-collect data for standard proteins in the reference database over the wavelength range from 168 to 260 nm, the accuracy of the analyses for all types of secondary structures have been significantly improved (Clarke & Jones, 1999).

6. SRCD – future prospects

6.1. Structural genomics – identification of novel folds

SRCD has the potential to play an important role in structural genomics programmes associated with the human genome project. One of the principal goals of structural genomics is to identify new types of protein folds that will be found in the open reading frames (ORFs) whose sequences have been determined. SRCD can aid in this pursuit. With the increased sensitivity possible in SRCD measurements, only very small amounts of protein (in the microgram or tens of microgram quantities) will be needed to define the secondary structure. Additionally, SRCD does not require the crystal formation needed for

crystallography, nor the high concentrations necessary for NMR studies, and information can be obtained facily on a time scale of minutes or hours rather than the months or years (or in the future, perhaps, weeks) that crystallography requires. While SRCD information cannot replace a complete X-ray structure determination, it should be useful in defining structural features present in a 'new' protein and also be of value as a test for modelling procedures.

Furthermore, because analyses of CD data are empirical techniques, when the 'fit' of a spectrum of an 'unknown' protein by the existing reference basis sets is found to be poor, this may be a very good indication that a new type of structure not represented in the reference database is being observed. This could either be a new motif or a new variation on a known secondary structural type. The more structural types used to produce the basis sets in the future, the better we will be able to identify truly 'new' structures, which could then be targeted within a structural genomics programme for examination in detail by X-ray crystallography. Thus, SRCD may be particularly useful for screening potential candidates for new types of protein folds.

6.2. Membrane proteins – a new CD structural database

It has been estimated that roughly one-third of all ORFs in the human genome may be membrane proteins. This is a category of protein for which there is a dearth of structural information. This paucity of information has been due to the difficulty in both crystallizing and solving the X-ray structures of hydrophobic and amphipathic membrane proteins, and has meant that, to date, only about 20 high-resolution structures of membrane proteins have been determined. In addition, a number of the structures determined thus far are from the same family, or are closely related, so few examples of unique membrane protein folds exist. CD (and in particular SRCD) could play an important role in providing information about the secondary structures of membrane proteins whose structures have not yet been determined, and in this way it could act as a guide for developing improved methods for prediction of membrane protein structures. To achieve accurate CD analyses of membrane proteins, however, will require the availability of a reference database specifically designed for membrane proteins, since it appears that their spectral properties differ considerably from those of soluble proteins (B. A. Wallace, A. J. W. Orry & R. W. Jones, unpublished results). The differences seen between spectra of membrane proteins and aqueous-soluble proteins may, in part, be due to solvent-dependent spectral shifts (Cascio & Wallace, 1995; Chen & Wallace, 1997) which arise because the interior of a membrane lipid bilayer has a considerably different dielectric constant than water. Collection and compilation of such a reference database, consisting of spectra of membrane proteins whose structures are known from X-ray crystallography, is underway at the CPMSD at the SRS (B. A. Wallace, A. J. W. Orry & R. W. Jones, unpublished results).

The additional information content obtainable by SRCD means that more subtle features of the secondary structures of membrane proteins may be analysable. For instance, in β -barrels, which are an important structural motif found in many membrane proteins, one of the important differences is the number of β -strands per barrel, which affects the twist of the strands. It may be possible to differentiate these types of features with the additional eigenvectors of information present in the SRCD data.

SRCD has a further advantage over conventional CD for membrane protein spectra data collection: the detector geometry. Membrane samples tend to scatter light achirally (Schneider & Harmatz, 1976; Wallace & Mao, 1984). In many cases this results in scattered light falling outside the detector face in the instrument geometry found in most commercial CD machines. This results in a distortion of the measured dichroism. In the SRCD the detector geometry is very flexible and can be changed to accommodate differential scattering (Clarke & Jones, 1999), thereby making it more suitable for studies of membrane proteins.

6.3. Dynamics, binding and relationship to crystallographic studies

The information which can be produced by CD and SRCD is complementary to that provided by X-ray crystallography (Wallace, 1999) in that it can lend insight into dynamics and environmental effects, can be useful for determining conditions for crystallization, and can permit the observation of protein folding pathways.

SRCD is valuable for studies monitoring conformational changes with ligand binding and under different environments, as have been performed with conventional CD. In the case of SRCD, however, the higher-intensity beam and subsequent enhanced S/N means that more accurate definitions of the changes will be possible, and the resultant smaller error bars on the measurements mean that it will be possible to detect smaller significant changes (Wallace, 2000).

By monitoring binding constants and conformational effects of salt, pH and additives, CD can be used for assaying conditions for crystallization. This is especially possible using SRCD because, with its high flux at low wavelengths, higher concentrations of salts and buffers (which reflect the conditions often used in crystallization) can be tolerated. As an example of this, the membrane protein database which is currently under construction (B. A. Wallace, A. J. W. Orry & R. W. Jones, unpublished results) is being collected using solution conditions under which the proteins have been crystallized, in order to permit more valid comparisons with the X-ray structures. Furthermore, the 'physiological relevance' of crystallization conditions (always an issue in crystallography) can also be assessed by comparison of the spectra taken under these conditions with spectra taken under pH, salt, concentration and other conditions that more nearly approximate those found *in vivo*.

Protein folding studies using CD can be undertaken for two very different purposes: as a means of monitoring the refolding of cloned expressed proteins (especially those produced as inclusion bodies) for use in other biophysical studies, and for examining the process of protein folding *in vitro*. For the former studies, the high sensitivity of SRCD means that smaller quantities of protein can be used, and more subtle differences between the refolded and native proteins can be detected. For the latter studies, faster kinetic measurements can be made with SRCD, even using a conventional mixing apparatus, because the improved S/N will decrease the averaging time needed to obtain comparable signals. This will mean that measurements can be extended into the sub-millisecond time range, with smaller samples of proteins. The small spot size of illumination of the sample in an SRCD relative to that in a conventional CD means that in stopped-flow mode much smaller sample volumes will be needed, a considerable practical advantage. In addition, a consequence of the higher flux at low wavelengths is that, in addition to the 208–240 nm range commonly monitored by conventional instruments, the information-rich 190–200 nm range can also be sampled by stopped-flow SRCD.

Hence, the advantages of SRCD will not be limited to secondary structure determinations, but will also be evident in a wide range of dynamic studies of proteins.

7. Conclusions

The future of SRCD looks particularly bright owing to the convergence of technical instrumental developments and advances in our knowledge of protein structures. Virtually all types of spectroscopic studies undertaken by conventional CD on proteins should be improved by the use of SRCD. Thus it is anticipated that this technique will find important new usage in modern structural molecular biology and in the developing field of structural genomics.

I thank D. T. Clarke, A. J. W. Orry and R. W. Janes for help with data collection at station 3.1, Daresbury Laboratory, and R. W. Janes for help with data processing. I also thank John Sutherland and John Kenney for providing me with information on the SRCD facilities at Brookhaven and Aarhus, respectively, and Gareth Jones for helpful discussions. This work was supported, in part, by the BBSRC Structural Biology Centre grant to the Centre for Protein and Membrane Structure and Dynamics (the CPMSD), and beam time grants from the BBSRC.

References

- Bohm, G., Muhr, R. & Jaenicke, R. (1992). *Protein Eng.* **5**, 191–195.
- Bolotina, I. A., Chekhov, V. O., Lugauskas, V. Y., Finkelstein, A. V. & Ptitsyn, O. S. (1980). *Mol. Biol.* **14**, 701–709.
- Brahms, S. & Brahms, J. (1980). *J. Mol. Biol.* **138**, 149–178.
- Cascio, M. & Wallace, B. A. (1995). *Analyt. Biochem.* **227**, 90–100.

- Chang, C. T., Wu, C. S. & Yang, J. T. (1978). *Analyt. Biochem.* **91**, 13–31.
- Chang, J. F., Phillips, K., Lundback, T., Gstaiger, M., Ladbury, J. E. & Luisi, B. (1999). *J. Mol. Biol.* **288**, 941–952.
- Chen, Y. & Wallace, B. A. (1997). *Biophys. Chem.* **65**, 65–74.
- Chen, Y.-H. & Yang, J. T. (1971). *Biochem. Biophys. Res. Commun.* **44**, 1285–1291.
- Chen, Y.-H., Yang, J. T. & Martinez, H. M. (1972). *Biochemistry*, **13**, 3350–3359.
- Clarke, D. T., Bowler, M. A., Fell, B. D., Flaherty, J. V., Grant, A. F., Jones, G. R., Martin-Fernandez, M. L., Shaw, D. A., Todd, B., Wallace, B. A. & Towns-Andrews, E. (2000). *Synchrotron Rad. News*, **13**, 21–27.
- Clarke, D. T., Doig, A. J., Stapeley, B. J. & Jones, G. R. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 7232–7237.
- Clarke, D. T. & Jones, G. R. (1999). *Biochemistry*, **38**, 10457–10462.
- France, L. L., Kieleczawa, J., Dunn, J. J., Hind, G. & Sutherland, J. C. (1992). *Biochim. Biophys. Acta*, **1120**, 59–68.
- Garone, L., Albaugh, S. & Steiner, R. F. (1990). *Biopolymers*, **30**, 1139–1149.
- Hennessey, J. P. & Johnson, W. C. (1982). *Analyt. Biochem.* **125**, 177–188.
- Johnson, W. C. Jr (1990). *Proteins Struct. Funct. Genet.* **7**, 205–214.
- Johnson, W. C. Jr (1996). *Circular Dichroism and the Conformational Analysis of Biomolecules*, edited by G. D. Fasman, pp. 635–652. New York: Plenum Press.
- Kallenbach, N. R., Lyu, P. & Zhou, H. (1996). *Circular Dichroism and the Conformational Analysis of Biomolecules*, edited by G. D. Fasman, pp. 201–259. New York: Plenum Press.
- Kelly, S. M. & Price, N. C. (1997). *Biochim. Biophys. Acta*, **1338**, 161–185.
- Manavalan, P. & Johnson, W. C. (1987). *Analyt. Biochem.* **167**, 76–85.
- Perczel, A., Park, K. & Fasman, G. D. (1992). *Analyt. Biochem.* **203**, 83–93.
- Provencher, S. W. & Glockner, J. (1981). *Biochemistry*, **20**, 33–37.
- Pulm, F., Schramm, J., Hormes, J., Grimme, S. & Peyerimhoff, S. D. (1997). *Chem. Phys.* **224**, 143–155.
- Qi, X. L., Holt, C., McNulty, D., Clarke, D. T., Brownlow, S. & Jones, G. R. (1997). *Biochem. J.* **324**, 341–346.
- Riley, M. L., Wallace, B. A., Flitsch, S. L. & Booth, P. J. (1997). *Biochemistry*, **36**, 192–196.
- Schneider, A. S. & Harmatz, D. (1976). *Biochemistry*, **15**, 4158–4162.
- Snyder, P. A. & Rowe, E. M. (1980). *Nucl. Instrum. Methods*, **172**, 345–349.
- Sreerama, N. & Woody, R. W. (1993). *Analyt. Biochem.* **209**, 32–44.
- Sreerama, N. & Woody, R. W. (2000). *Biophysical J.* **78**, 334a.
- Sutherland, J. C. (1996). *Circular Dichroism and the Conformational Analysis of Biomolecules*, edited by G. D. Fasman, pp. 599–633. New York: Plenum Press.
- Sutherland, J. C., Desmond, E. J. & Takacs, P. Z. (1980). *Nucl. Instrum. Methods*, **172**, 195–199.
- Sutherland, J. C., Emrick, A., France, L. L., Monteleone, D. C. & Trunk, J. (1992). *Biotechniques*, **13**, 588–590.
- Sutherland, J. C. & Griffin, K. P. (1983). *Biopolymers* **22**, 1445–1448.
- Sutherland, J. C., Lin, B., Mugavero, J., Trunk, J., Tomasz, M., Santella, R., Marky, L. & Breslauer, K. J. (1986). *Photochem. Photobiol.* **44**, 295–301.
- Symmons, M. F., Buchanan, S. G. S., Clarke, D. T., Jones, G. & Gay, N. J. (1997). *FEBS Lett.* **412**, 397–403.
- Wallace, B. A. (1999). *CCP4 Newsl.* **37**, 29–30.
- Wallace, B. A. (2000). *Nature Struct. Biol.* In the press.
- Wallace, B. A. & Mao, D. (1984). *Analyt. Biochem.* **142**, 317–328.
- Wallace, B. A. & Teeters, C. L. (1987). *Biochemistry*, **26**, 65–70.
- Woody, R. W. (1995). *Methods Enzymol.* **246**, 34–71.