Calibration and Standardisation of Synchrotron Radiation Circular Dichroism and Conventional Circular Dichroism Spectrophotometers

Andrew J. Miles \(^a\), Frank Wien \(^a\), Jonathan G. Lees \(^a\), A. Rodger \(^b\), Robert W. Janes \(^c\) and B.A. Wallace \(^a,\ast\)

\(^a\) Department of Crystallography, Birkbeck College, University of London, London WC1E 7HX, UK
\(^b\) Department of Chemistry, Warwick University, Coventry CV4 7AL, UK
\(^c\) School of Biological Sciences, Queen Mary, University of London, London E1 4NS, UK

Abstract. Synchrotron radiation circular dichroism (SRCD) is an emerging technique in structural biology with particular value in protein secondary structure analyses since it permits the collection of data down to much lower wavelengths than conventional circular dichroism (cCD) instruments. Reference database spectra collected on different SRCD instruments in the future as well as current reference datasets derived from cCD spectra must be compatible. Therefore there is a need for standardization of calibration methods to ensure quality control. In this study, magnitude and optical rotation measurements on four cCD and three SRCD instruments were compared at 192.5, 219, 290 and 490 nm. At high wavelengths, all gave comparable results, however, at the lower wavelengths, some variations were observable. The consequences of these differences on the spectrum, and the calculated secondary structure, of a representative protein (myoglobin) are demonstrated. A method is proposed for standardising spectra obtained on any CD instrument, conventional or synchrotron-based, with respect to existing and future databases.

1. Introduction

Synchrotron radiation circular dichroism (SRCD) spectroscopy is a method that was first developed more than 20 years ago [1,2] but which is just beginning to realise its full potential in the study of protein secondary structures. It permits the collection of much lower wavelength data than conventional circular dichroism (cCD) spectroscopy due to the high light flux available from the synchrotron source. Indeed, recently a protein spectrum in aqueous solution was obtained with a lower wavelength limit of \(\lesssim 154 \text{ nm}\) [3], as compared to a practical limit of \(\sim 185–190 \text{ nm}\) on conventional instruments. It has been suggested that the extra information from the vacuum ultraviolet (VUV) data will allow a greater discernment of structural features, including information on folds and motifs [4]. Creation of a new reference database of protein CD spectra, including the lower wavelength data which will enable fold recognition studies, is in progress [3].

There are currently a number of operational SRCD instruments world-wide [5], and several more will be coming on-line in the near future [6]. Therefore establishment of a consistent set of calibration standards and protocols for SRCD and cCD is now essential. Instrumental calibrations of magnitude,
polarisation, and wavelength are needed, as are precise measurements of optical cell pathlengths and protein concentrations, in order to obtain the correct measurements of ellipticities that are necessary for secondary structural analyses [7–9]. Since the currently available reference databases for protein secondary structural analyses were derived from measurements made on cCD instruments [10], and future reference databases will be measured on SRCD instruments but must also be usable with cCD data, it is especially important that SRCD and cCD instruments are calibrated to the same standards [6]. Establishing protocols for calibration will also be important in the context of “good practice” procedures for protein drug standardisation and characterisation in the pharmaceutical industry where CD is currently being employed to confirm batch-to-batch reproducibility of protein structure.

Conventional circular dichroism instruments are routinely calibrated for amplitude at 290 nm using camphor sulphonic acid (CSA) [11] or ammonium camphorsulphonate (ACS) [12]. However, calibrations should be done at more than one wavelength [13] in order to cover the wavelength range measured in a protein spectrum, and to demonstrate the linearity of the response. Other candidates that have been proposed for calibration standards are the lower wavelength (192.5 nm) CSA peak [11], pantolactone (PL) [14] with a peak at 219 nm, and cobalt (III) tris-ethylenediamine (Co-en), with a maximum at 490 nm [15], which also provides a good standard for the visible region.

As a preliminary attempt at standardisation, this study compares optical rotation magnitude measurements for a number of compounds on three SRCD instruments with those obtained on four cCD instruments. In addition, spectra of the protein myoglobin were also measured on all these instruments in order to examine the effects of instrument variation on the spectrum of a standard protein. The calculated secondary structures derived from these spectra were then compared in order to determine the consequences of the variations. Finally, a method is proposed for standardisation of protein spectra obtained on any instrument, which will “correct” the spectra and enable empirical analyses based on existing and future database spectra, using the values obtained for four standards on these instruments.

2. Materials and methods

2.1. Materials

(D)-(−)-pantolactone, of 99.99% purity (Sigma-Aldrich Company Ltd), was dissolved at a concentration of 0.0115 M in 18.2 MΩ deionized water (dH₂O). (+)-Camphor-10-sulphonic acid, 99% purity (Sigma-Aldrich Company Ltd), was dissolved in dH₂O at a concentration of 0.047 M, determined by the absorbance at 285 nm assuming an extinction coefficient of 34.5 M⁻¹ cm⁻¹ [16]. Co[ethylendiamine]₃Cl₃·NaCl·6H₂O, the kind gift of Peter W. Thulstrup (from Roskilde University, Denmark), was dissolved in dH₂O at a concentration of 33.4 mM, as determined by the absorbance at 464 nm, using an extinction coefficient of 84 M⁻¹ cm⁻¹ [15]. The CSA sample was diluted to concentrations ranging from 0.9× to 0.1× to assay the CD signal as a function of concentration, as a test of the linearity of the response. The samples of CSA, PL and Co-en were stored in the dark at 4°C and used within 2 weeks of preparation. In the first set of experiments, the same samples (using the same CD cells) were tested on most of the machines (4 cCDs and 1 SRCD) within that 2-week period. These samples were retested on the first machine at the end of this time to ensure that no changes had taken place. Because of constraints associated with beamtime scheduling which did not permit measurements on the other two SRCDs to be done during the initial two week period, another set of samples was prepared in the same way and was used for measurements on them, and for comparison, on one of the cCDs that had been used for the first set of measurements.
Horse skeletal myoglobin (ICN Biochemicals) was allowed to dissolve in dH$_2$O overnight at a concentration of $\sim$8 mg/ml. One ml of the solution was then dialysed against 50 ml of dH$_2$O for 2 hrs to reduce the salt content, degassed and centrifuged at 5000 $\times$ $g$ to separate any undissolved material. The final concentration of the protein was determined to be 7.43 mg/ml by replicate quantitative amino acid analyses.

2.2. CD and SRCD spectroscopy

The following parameters were set on all instruments: bandwidth 1 nm, step size 0.2 nm, temperature 25$^\circ$C. Spectra were collected on the Aviv 62ds and 215 instruments and on all the SRCD instruments in step scan mode with an averaging (dwell) time of 1 second. At least 3 measurements were made and averaged for each sample on each instrument. The Jasco J-600 and J-715 instruments were run in continuous scan mode, averaging over 8 accumulations with a response time of 0.5 s. The PL, CSA, and Co-en spectra were all collected using a 1 mm pathlength Suprasil cell (Hellma UK Ltd). A dH$_2$O baseline, collected using the same cell and the same parameters, was subtracted from all spectra. All myoglobin spectra were obtained using the same cylindrical 0.001 cm pathlength demountable cell (Hellma UK Ltd) held in a specially adapted cell holder; baselines of dialysate were collected in the same cell.

SRCD measurements were made at the SRS Daresbury (UK) on beamline CD12; at ISA (Denmark) on beamline UV1; and at BESSY2 (Germany) on beamline 3m_NIM1_C. Where possible, the parameters were set to the values used on the cCD instruments. 3m_NIM1_C as yet had no bandwidth control, so the slits were fully open, and due to limited availability of beamtime, only 2 repeats were done. Because two of the beamlines had upper wavelength limits of <400 nm, measurements on the Co-en sample were not possible on these instruments. The SRCD instrument on beamline CD12 records intensities as arbitrary counts. Therefore, the literature value for CSA at 290 nm was assumed and the other values calculated from this for the CD12 data.

The pathlengths of the optical cells used were determined using the absorbance of a 1 mM solution of potassium chromate ($\varepsilon_{372} = 4830$ M$^{-1}$ cm$^{-1}$ (Chemical Rubber Company)) and by the interference fringe method [17] for the (nominally) 0.1 and 0.001 cm cells, respectively.

2.3. Secondary structure calculations

The secondary structural analyses used the DICHROWEB (http://www.cryst.bbk.ac.uk/cdweb) interactive webserver [8] with the CONTIN [10,18] algorithm and database 6 [8]. The normalised root mean square deviation (NRMSD) parameter [19] was calculated as a means of assessing the goodness-of-fit.

2.4. Standardisation method

For each of the instruments for which measurements were available for all four standards, a standardisation procedure was applied to the myoglobin spectrum collected on that instrument. This was done as follows: for each standard, the literature delta epsilon value was divided by the delta epsilon value measured on that instrument. Then a second-order polynomial was fit to the ratios at the four different wavelengths, and the ratio values at all other wavelengths ($R_\lambda$) were interpolated from the ratio versus wavelength curve. Then for each myoglobin data point, the scaled delta epsilon value $\Delta\varepsilon_{s\lambda} = R_\lambda \cdot (\Delta\varepsilon_\lambda)$ was plotted as the “standardised” spectrum. The procedure is detailed in the Annex of this paper.
3. Results

One advantage of this study is that the comparative measurements could be done on equivalent samples (close in time and prior to degradation) using the same CD cell (and thus the same pathlength), by the same investigator, thereby eliminating many possible sources of systematic errors, and enabling a true comparison of the instrument characteristics.

3.1. Accuracy of optical rotation magnitude measurements

Pantolactone has a $\Delta\varepsilon$ at 219 nm of $-4.9$ mdeg M$^{-1}$ cm$^{-1}$ [20] therefore a signal of 186 mdeg was expected at this wavelength for the 0.0115 M solution used in this work.

CSA has a $\Delta\varepsilon$ at 290 nm of 2.37 mdeg M$^{-1}$ cm$^{-1}$ [20] therefore a signal of 367 mdeg was expected for the 0.047 M solution used in this work. $\Delta\varepsilon$ for the second CSA peak at 192.5 nm is 4.72 mdeg M$^{-1}$ cm$^{-1}$ [11]. Therefore a peak of 146 mdeg was expected from a 5 times dilution of the CSA sample. It should be noted that the CSA (and all peaks) varied in wavelength position between the different instruments: The low wavelength CSA peak was found at positions ranging from 190.5 to 194.0 nm, and the high wavelength peak was between 289.1 and 293.2 nm. Thus calibration of wavelength position is another important parameter to consider for instrument standardisation (Miles & Wallace, in prep.). In this paper, however, the ellipticity values reported are always for the peak maximum, regardless of what the peak wavelength position was.

Assuming enantiomeric purity, with a $\Delta\varepsilon$ at 490 nm of 1.89 mdeg M$^{-1}$ cm$^{-1}$ [15], the Co-en sample used in this work should give rise to a CD signal of 208 mdeg. The results obtained on all instruments are displayed graphically in Fig. 1a. The “literature value” curve is established by a simple fit (Excel) to the four data points and is for visualisation purposes only. For clarity, all the cCD measurements are shown as open symbols, and all the SRCD measurements are shown as filled symbols. A given symbol signifies data measured on a single instrument.

Since instruments are generally routinely calibrated using the CSA 290 nm peak, it is not surprising that they all correspond well to each other and to the literature value at this wavelength. However, the deviations are considerably larger at lower wavelengths.

At 490 nm, all ellipticity measurements were within 6% of the literature value. Variations for single machines across all four calibration standards range from 0 to 3% for one cCD instrument to 4 to 18% for another cCD instrument, with the SRCD instruments mostly falling in the middle range (0 to 10%). The variations include values that are both too large and too small, with no particular pattern being discernable. Also, there is no clear trend of deviation between cCDs or the SRCDs. The most significant result is that none of the instruments vary by a constant amount from the literature values across all wavelengths, so no simple scale factor can be applied. This is in complete contrast to the common practice whereby an instrument is calibrated at a single point and it is assumed the whole wavelength range is thereby calibrated.

At all three calibration wavelengths in the UV, two of the cCDs are calibrated to within 3% of the expected value. This is close to the 2% error level estimated from repeated measurements. One cCD displays significant deviation at 219 nm ($-10\%$) and another deviates by $+11\%$ at 192.5 nm and $+18\%$ at 219 nm. One SRCD instrument was found to have a deviation of $\sim8\%$ at all the wavelengths measured, whilst another was within 4% of the literature values at three wavelengths, but deviates at 192.5 nm by $>10\%$, hence its 192.5/290 peak ratio is the highest value measured, 2.20.
Fig. 1. Variations between different instruments: (a) Amplitude variations of four cCD and three SRCD instruments for three standard compounds (CSA, PL, and Co-en) at four wavelengths. The “literature value” curve is established by a simple fit to the values, and is for visualisation purposes only. cCD measurements are shown as open symbols, SRCD measurements are shown as filled symbols. A given symbol signifies data measured on a single instrument. (b) Plot of ($\Delta \varepsilon_{192.5 \text{ nm}} / \Delta \varepsilon_{290 \text{ nm}}$) ratios of CSA for the cCD and SRCD instruments, using the same symbols as in Fig. 1a. (c) Uncorrected myoglobin spectra obtained on cCD and SRCD instruments. All spectra were obtained using the same concentration of protein and the same cell (and hence the same pathlength). The spectra from the cCD instruments are shown in red/orange/yellow/purple; the spectra from the SRCD instruments are in light and dark blue and green (in each case, the same colour as used for the symbols in Figs 1a,b). The spectrum of myoglobin from the reference database [10] is shown in black, for comparison. (d–f) Myoglobin spectra from various instruments scaled to calibration points at (c) 290 nm, (d) 219 nm, and (e) 192.5 nm, respectively. It can be seen that there is still considerable variation between spectra obtained on different instruments, with no clear segregation into cCD and SRCD populations. (g) Convergence of the myoglobin spectra after standardisation by the multiwavelength (“four-point”) method proposed in this paper.
3.2. Linearity of detector response

It is important to demonstrate that the instruments produce linear responses across the range of ellipticity values measured, in order for the CSA dilution calculations to be correct. This was assayed by measuring a series of dilutions of the CSA sample, and plotting ellipticity versus concentration (data not shown). All instruments surveyed exhibited linear responses up to ellipticity values of 300 mdeg.

Another important issue to consider is the total absorbance (often measured as high tension (HT)) of the sample. When the sample absorbance is too high, the number of photons reaching the detector becomes too small for accurate measurement and the CD signals cannot be trusted. The cut-off must therefore be determined on each individual instrument. In this study, all measurements were done under conditions where the absorbance was still within the accurately measurable range.

3.3. CSA signal ratios

The ratio of the ellipticities of the two CSA peaks (Δε_{192.5 nm}/Δε_{290 nm}) is often used as a simple “two-point” means of calibrating the instrument, as opposed to the absolute measurement of the ellipticity of a single peak. Reported values in the literature range from ~1.90 to 2.20 (i.e., [11,17,21–23]). In this study, the values for the ratio ranged from 1.96 to 2.20 (Fig. 1b). However, this parameter on its own is not particularly useful, as the absolute magnitude may be incorrect, whilst the ratio is reasonable. Also, there is no clear standard for the “correct” value, with significant variations possible due to, amongst other sources, CSA purity, light-induced degradation, and instrument stray light. Hence, we consider the magnitudes of the individual peaks may be more suitable measures of the CD calibration. However, they are dependent on having accurate values for the extinction coefficients and ellipticity values for the standard materials.

3.4. Effects of instrumental variation on protein spectra

The variations between instruments for the standards are considerable in the critical far UV spectral range used for protein analyses. The obvious consequences of this is that protein spectra measured on the various instruments will differ considerably both in magnitude and shape (the latter because the deviations from true signals are different for different wavelengths) from each other (Fig. 1c) and from the whale myoglobin spectrum in the reference database (black curve, Fig. 1c). Therefore the simple scaling of all the spectra by a factor determined at one wavelength will not produce equivalent spectra. To illustrate this, all the spectra were normalised to the literature values at each of the single UV wavelengths (192.5, 219 and 290 nm), and replotted (Figs 1d–f). The most successful of these was the scaling to the 219 peak, but in this case the spectra showed some variation, especially the one outlier (green) that had very different values for the standards at all wavelengths. Nevertheless, none of the single wavelength scalings produced very good correspondences. An alternative, the four-point standardisation method, described above, resulted in very close correspondences with each other (Fig. 1g) and with the spectrum in the reference database.

3.5. Effects of instrumental variation on calculated secondary structures

Two instruments (one cCD and one SRCD) produce spectra nearly identical to the reference database spectrum of myoglobin, so it is no surprise that they produce similar calculated secondary structures.
However, the unscaled spectra (Fig. 1c) for the other instruments produce dramatically different calculated secondary structures (ranging from 0.44 to 0.83, average 0.71), where the value calculated by the DSSP algorithm from the crystal structure was 0.75 [24]. This result demonstrates that without proper cross-calibration, empirical methods for secondary structural analyses will not work for either cCD or SRCD measurements. However, after the four-point standardisation scaling, the resulting spectra produced a narrower range of secondary structures (range from 0.70 to 0.74, average = 0.72), close to the crystal structure value. On this basis, we propose that the scaling method described may be a useful one for both cCD and SRCD data.

4. Discussion

4.1. Comparisons of cCD and SRCD spectra

It can be seen that there is considerable variation between spectra obtained on various instruments, with no clear segregation into cCD and SRCD populations, and no clear trend between or within the types of instruments. There is very close correspondence between one of the cCD instruments (red) and one of the SRCD instruments (light blue). In a previous study, it was suggested that at the very low wavelength limits, conventional spectra differ slightly from SRCD spectra [25]; this was attributed to the necessity of fully opening the slits in the cCD to achieve high enough light flux, which ultimately allowed through light of other wavelengths, and seriously questions the use of data at wavelengths \( \leq 185 \text{ nm} \) from cCD instruments. The purpose of the present study, however, was to examine the correlations at wavelengths where both instruments are operating optimally, i.e. above \( \sim 190 \text{ nm} \).

4.2. Proposed standardisation method for calibration corrections

The variations between the measurements of the standards on the various instruments are reflected in the spectra of the protein sample. The overall magnitudes and the relative peak magnitudes are significantly different in spectra from the different instruments (Fig. 1c), and most experimental spectra are at variance with the corresponding spectrum from the reference database (in black, Fig. 1c). These spectral differences result in large differences and inaccuracies in the calculated secondary structures derived from them. The variations could also lead to problems in use of CD for “good practice” quality control analyses of pharmaceutical proteins if uncorrected data from different instruments were compared. As a result, it is clear that there needs to be some procedure that can be used to standardise/calibrate the spectra taken on one instrument with the spectra taken on other instruments and in the reference databases.

The established method for calibration of cCD instruments uses only the value for the CSA calibration at 290 nm to normalise the spectra on a given instrument. That this is actually not a very suitable method (see Fig. 1f) is unsurprising given that this peak is outside the far UV wavelength range used for measuring protein spectra. Normalisations using either the 192.5 nm CSA or the 219 nm PL peaks are somewhat better (Figs 1d and 1e) but still do not take into account the wavelength-dependent variations. Hence we propose a new procedure which utilises the ratios of the measured and expected delta epsilon values at each wavelength to develop a model function for corrections to be applied to spectra. This has proved to be a much more successful method of calibration than any single wavelength approach, resulting in superposable spectra from the various instruments and with the corresponding spectrum from the reference dataset (Fig. 1g).
Fig. 2. Examples of application of the standardisation method to data collected on two different CD instruments (blue and orange curves): (a) uncorrected spectra of myoglobin obtained on the two CD instruments; (b) plots of $R_\lambda$ vs. wavelength for the two instruments, derived from fits to the four standard measurements; (c) corrected (according to the method described in the Annex) spectra of myoglobin obtained on the two CD instruments.

Acknowledgements

This work was supported by grants B13586 and B14225 from the BBSRC to BAW. JL was supported by a BBSRC CASE studentship with GlaxoSmithKline, and AM was supported by an MRC studentship. Beamtime access at ISA was enabled by an EC Access to Research Infrastructure Action of the Improving Human Potential Programme to Soren Pape Moller (Aarhus University). We thank the following for technical assistance at the SRCD beamlines: Dr. David Clarke (CD12 beamline, SRS Daresbury), Dr. Soren Vronning Hoffmann (UV1 beamline, ISA), and Dr. Peter Baumgartel (beamline 3m_NIM1_C, BESSY2). We thank Dr. Peter Sharratt of the PNAC facility, University of Cambridge, for the quantitative amino acid analyses.

Annex: Protocol for the four-point standardisation method

1. Measure the CD spectrum of the protein of interest (Fig. 2a).
2. Measure the CD spectra for the three standards (CSA, PL, Co-en) at the following four wavelengths: 192.5, 219, 290, 490 nm.
3. For the above measurements on the standards, calculate the ratio ($R_\lambda$) of the literature value for the ellipticity divided by the measured value for the ellipticity.
4. Fit a second order polynomial to the four $R_\lambda$ vs. wavelength data points (Fig. 2b).
5. Determine the value of the polynomial ($R_\lambda$) for each wavelength over the whole spectral range.
6. Multiply the CD spectrum of the protein of interest by the $R_\lambda$ value at each wavelength and plot the resulting standardised spectrum (Fig. 2c).

References


Submit your manuscripts at http://www.hindawi.com