Chapter 14

CD Spectroscopy: An Essential Tool for Quality Control of Protein Folding

Giuliano Siligardi and Rohanah Hussain

Abstract

The production of diffraction quality protein crystals for X-ray crystallography has been greatly accelerated by the development of high-throughput protein (HTP) methods, which enable a large number of crystallization conditions to be rapidly investigated. Monitoring sample quality and the effect of crystallization buffers on protein behavior in solution should be considered as part of the crystallization experiment. Circular Dichroism (CD) spectroscopy is the ideal technique for these tasks as it can be operated in a high-throughput mode. Using CD to screen ligand binding interactions could show whether protein function/activity is retained, altered, or lost under different crystallization conditions. In this chapter, several methods for high-throughput CD (HTCD) applied to the preparation of proteins for crystallization will be presented. Quality control (QC) of protein batches in terms of conformational folding is often disregarded in protein production. Examples of batch-to-batch variation in the local tertiary structure of aromatic side chain residues revealed by CD will be discussed. In some of the examples, the fact that ligand binding properties were affected by changes in folding clearly shows that the characterization of folding of recombinant protein batches should not be ignored but be implemented as an important part of protein quality control.

Key words Circular dichroism (CD), Synchrotron radiation circular dichroism (SRCD), Protein secondary structure, Local tertiary structure, High-throughput CD (HTCD), SRCD UV-denaturation assay

1 Introduction

In the last few decades, drug discovery has benefited from information about the three-dimensional structure of proteins at atomic resolution determined by X-ray crystallography [1] and NMR spectroscopy [2]. However, the number of resolved human protein structures deposited in the Protein Data Bank is still only a small fraction of the approximately 20,000 protein-coding genes in the human genome [3]. Proteins that cannot be studied by either X-ray crystallography due to their failure to crystallize or NMR due to size or irregular structure can be characterized by Circular Dichroism (CD) spectroscopy. CD is the differential absorption between left and right circularly polarized light of a chiral molecule [4] and therefore CD spectroscopy is sensitive to the absolute

Raymond J. Owens (ed.), *Structural Proteomics: High-Throughput Methods*, Methods in Molecular Biology, vol. 1261, DOI 10.1007/978-1-4939-2230-7_14,

configuration and conformation of chiral molecules. With the exception of glycine, amino acids are chiral molecules adopting either l or d stereoisomers which are nonsuperimposable mirror-imaged configurations. Natural proteins only consist of l amino acids.

Although of low resolution compared to X-ray and NMR methods, CD enables the rapid and direct characterization of protein folding in solution as a function of concentration, aqueous buffer composition, ionic strength, temperature, pH, detergents, chemical agents [4], and UV irradiation, age, and ligand binding interactions [5–7]. Importantly, CD spectroscopy can be used to assess the conformational behavior of a protein in a broad range of conditions to optimize protein formulation prior to carrying out high-resolution techniques such as X-ray crystallography and NMR spectroscopy. The production of good quality protein crystals for X-ray crystallography in terms of size and diffraction patterns has been greatly accelerated by high-throughput (HTP) screening of many crystallization conditions in multi-well plate formats. Combining this with high-throughput CD (HTCD) screening of protein formulations used for crystallization experiments provides added value. It can be used to characterize the sample in terms of conformational behavior in solution and may also identify structural features of the sample in a particular crystallization buffer that favor crystallization. Protein function, such as ligand binding interaction, can also be tested by HTCD to assess whether the protein function/activity is retained, altered, or lost under different crystallization conditions.

Another application of CD spectroscopy is monitoring batch-to-batch variation in recombinant protein production. This quality control (QC) is relevant not only to structural studies but also to the use of recombinant proteins for therapeutic purposes [8]. Recombinant proteins are often produced in several batches during the lifetime of a scientific project. The QC of protein batches is an important aspect of this process and HTCD offers a way of rapidly assessing protein folding and any variation between different batches of samples. For the biotechnology and pharmaceutical industry the QC of protein folding by CD spectroscopy could be used as a fingerprint of product quality.

In this chapter the application of SRCD using the Diamond B23 module A beamline to the quality control of proteins will be described.

2 Sample Preparation for CD Measurements

Scanning proteins in the far-UV region (185–260 nm) reveals details of secondary structure content, while measurements in the near-UV region (260–330 nm) reflect the local tertiary

conformation of aromatic side chain residues such as tryptophan, tyrosine, and phenylalanine and dihedral angles of the disulfide bonds [9]. The sample specification for SRCD measurements in the far-UV spectrum at Diamond B23 module A beamline is summarized as follows. The concentration of the protein sample needs to be 0.3–0.5 mg/ml in a total cell volume of 30–200 μl for a 0.02 cm path length cell. The UV absorbance (A280 nm) of the solution should be between 0.5 and 1.5. Protein samples are often formulated with sodium chloride (NaCl) to maintain solubility. Since chloride anions are *not* optically transparent below 200 nm it is critical to keep the concentration of chloride ions (NaCl) and other buffer components (e.g., Tris–HCl, Hepes, MES) in the sample as low as possible, typically 10–20 mM. Higher salt concentration can be used provided the sample is at higher concentration allowing smaller path lengths of 2–50 μm to be used according to Beer-Lambert's Law. Only 2-mercaptoethanol can be used as a reducing agent in the sample if this is required for stabilizing the protein. SRCD beamlines can penetrate by 5–10 nm the cutoff of the solvents/buffers encountered with benchtop instruments; however, any CD band observed below 170 nm should be treated suspiciously as due to noise artifact. For measurements in the near-UV region, the sample concentration is typically 1–2 mg/ml in total volume of approximately 500 μl for a path length cell of 1 cm. Buffers containing 150 mM or higher NaCl can be used since the chloride anions are transparent in the near-UV region.

For a protein under a given environment condition, the content of protein secondary structure elements (α-helix, β-strand, β-turns, collagen type (PPII) and unordered) can be estimated quantitatively using the CONTIN, SELCON, and CDSSTR methods [10]. These methods can be applied using the suite of programs of bespoke benchtop CD instruments Chirascan (APL, UK), CD20 (Olis, USA), Jasco CD instruments (Japan), or accessing the following web sites: CD-Pro [http://lamar.colostate.edu/~sreeram/CDPro/main.html], DichroWed [http://dichroweb.cryst.bbk.ac.uk/html/links.shtml], and Diamond B23 CD-Apps [http://www.diamond.ac.uk/Beamlines/Soft-Condensed-Matter/B23/manual/Beamline-software.html, http://confluence.diamond.ac.uk/display/B23Tech/CD+Apps+documentation].

3 Instrumentation

Three screening systems are available for HTCD: the automated CD (ACD) Chirascan (Applied Photophysics, Leatherhead, UK [www.photophysics.com]) since 2012, the

high-throughput CD J1000 and J1500 Jasco spectropolarimeters [www.jasco.uk] since 2012, and HTCD-B23 module A beamline (Diamond Light Source, Chilton, Oxfordshire, UK [www.diamond.ac.uk]) since 2013. ACD, which is mainly used by Biotech and Pharma industries, uses a robot to inject the sample solution from 96-well or 384-well multiplates into a cuvette cell for CD spectral measurements. The automated cell cleaning, cell drying, sample injection, and CD measurement enable the processing of up to 200 samples per day [11]. The Jasco system that can accommodate two 96-well multiplates is based on a similar method of sample injection into a cuvette cell, CD measurement, waste collection, and cleaning cycle for the subsequent sample injection [www.jasco.de]. For both Chirascan and Jasco HTCD systems, the choice of the optimum concentration for the protein solutions to be used with the liquid transfer systems follows the same procedure described in Subheading 1, which is a protein concentration of 0.3–0.5 mg/ml for a 0.02 cm path length cell.

HTCD-B23 on the other hand is a multiplate CD technology that measures directly the synchrotron radiation CD (SRCD) from UV transparent fuse quartz 96-well or 384-well multiplates (Hellma, UK). The position of the transparent fuse quartz 96-well or 384-well multiplates containing the protein samples is controlled by a motorized X-Y stage in a vertical sample compartment (Fig. 1). The SRCD measurements are possible in this chamber due to the highly collimated micro light beam of about 1 mm (horizontal) × 0.5 mm (vertical) of the B23 beamline that passing through the flatter central area of the meniscus solution is not affected by its curvature. The solution of the vertical chamber cannot be implemented with benchtop CD instruments because of their highly divergent and larger incident beam light (8 mm × 10 mm), which will be affected by the meniscus curvature of the solution in the well and by the polarization artifacts produced by deflecting the incident light vertically. The important feature of this is that varying the volume of the sample solution in the well will vary linearly the solution height leading to reproducible measurements in the range from 1 (60 μl for 96-well and 20 μl for 384-well) to 8 mm equivalent path length (240 μl for 96-well and 80 μl for 384-well) (Fig. 1). Another important advantage of HTCD-B23 is that detergents, indispensable for membrane proteins production and viscous solutions, are not limiting or troublesome factors. The productivity of HTCD-B23 can be increased up to 300 samples per day. The use of the HTCD-B23 is open to UK scientific academia and industry with call for research proposals twice per year: in March and September [www.diamond.ac.uk].

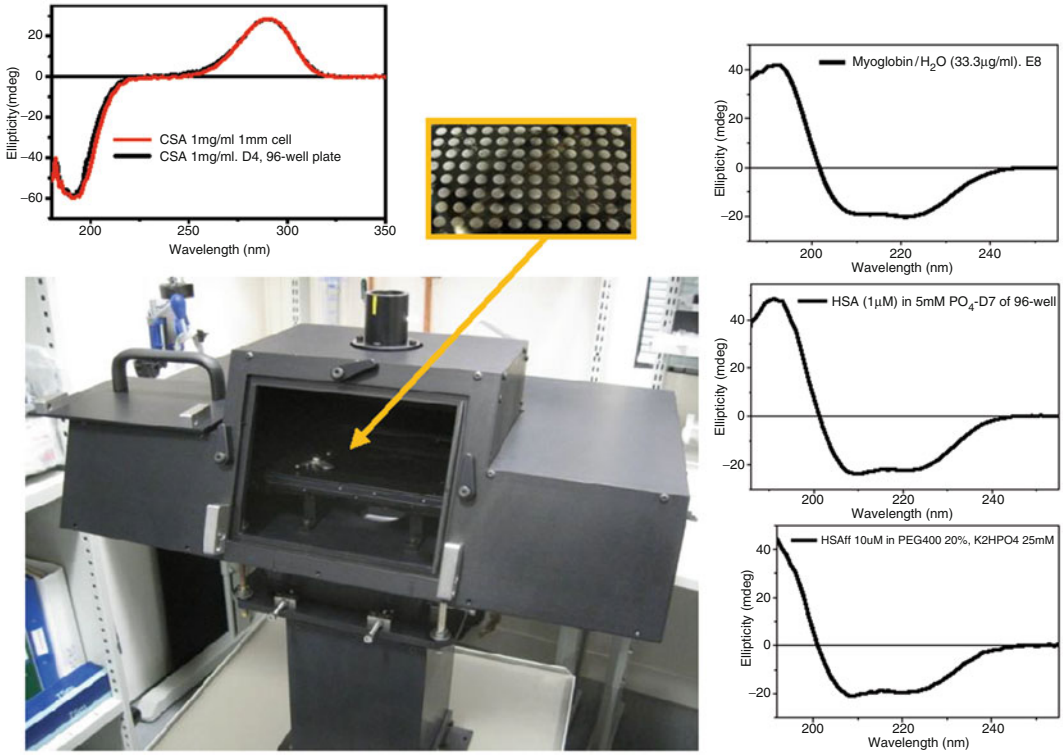

Fig. 1 Vertical sample compartment of Diamond B23 module A beamline. The chamber enables the SRCD measurements of horizontally positioned samples with respect to the incident monochromatic light. It has been designed to accommodate the 96- and 384-well multiplates made of fused quartz (Suprasil, Hellma). The *central insert* shows where the enlarged 96-well multiplate is located inside the chamber (*yellow arrow*). The *left insert* illustrates the SRCD spectrum of 60 µl of 1 mg/ml of 1*S*-(+)-Camphor-10-sulfonic acid (CSA) measured in the D4 well (*black*) of the 96-well plate that was identical to that measured for the same solution in 1 mm pathlength cell (*red*). The *right inserts* illustrate from *top* to *bottom* the SRCD spectra of myoglobin in H_2O and human serum albumin (HSA) in 5 mM phosphate buffer and 20 % PEG 400 in 25 mM phosphate buffer respectively

4 Applications of SRCD

4.1 Protein Folding in Crystallization Buffers

The example in Fig. 2 illustrates the preliminary SRCD spectra of 96 equine skeletal muscle myoglobin (Sigma M0630) dissolved in MemGold2™ HT-96 crystallization screen (Molecular Dimensions (www.moleculardimensions.com)) (Table 1) using the 384-well multiplate. The MemGold2™ consists of 96 conditions of the most recent alpha helical membrane protein crystallization conditions that contain combinations of high ionic strength (100–200 mM) cations (Li, Na, K, Mg, Cd, Zn, and ammonium), anions (chloride, acetate, citrate, sulfate, cacodylate, formate), buffers (Tris, Hepes, Mes, Mme, Bis-Tris, Mada, Mops, Choline, Glycine, and phosphate) at various pHs from 4 to 9, and PEG of

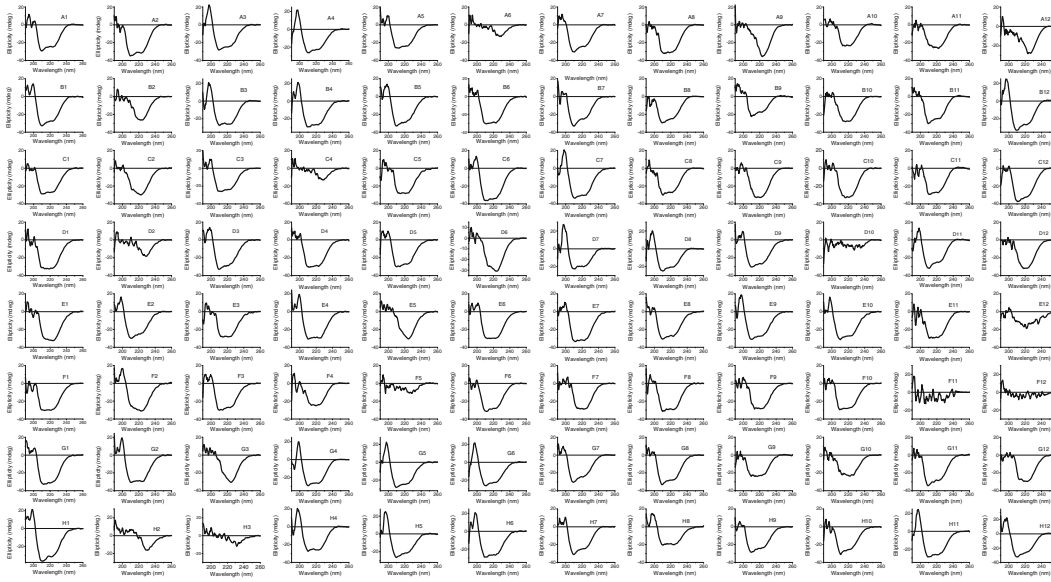

Fig. 2 SRCD spectra of 96 myoglobin solutions prepared from 96 crystallization buffer conditions of MemGold2™ solvent kit (*see* Table 1)

different molecular sizes (200–8,000) at high concentration from 9 to 66 % (Table 1). For each solution, 20 μl of myoglobin (0.030 mg/ml) was transferred into 96 wells of the 384-well multiplate (Hellma, UK) used for this experiment. When possible, it is convenient to speed up the sample preparation by diluting a stock solution of high protein concentration with the appropriate volume of buffer to reach the final concentration of 0.010 mg/ml. Filling the well with 20 μl of protein solution, the collected SRCD spectrum was identical to that measured with 1 mm path length cell (Hellma, UK). For each well, 20 μl was found to be the minimum volume retaining spectral reproducibility in terms of path length. Using a 96-well multiplate instead would have required 60 μl to achieve the same 1 mm path length due to the larger diameter of the well cell.

Figure 3 shows an example of data analysis in which the percentage of different secondary structure elements in the samples has been estimated from SRCD data using CONTIN/LL, a variant of CONTIN algorithm [10, 12, 13]. The estimation is calculated using a linear combination of CD spectra of reference soluble and membrane proteins with known secondary structure content assigned from X-ray coordinates [14]. For all CD methods, it is essential to be consistent with this method for assigning the protein secondary structure from X-ray atomic coordinates for the reference proteins.

Table 1
MemGold2™ Kit (Molecular Dimensions) for the screening of 96 crystallization conditions for alpha helical membrane proteins

A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12
14 % PEG550 MME	44 % PEG 3000	10 % PEG 450	8 % PEG 1450 0.02 M $CaCl_2$ 0.04 M $MgSO_4$	32 % PEG 400 0.05 M NiSO4 0.05 M LiCl	10 % PEG 3350	11.5 % PEG 4000	30 % PEG 400 0.1 M CdCl	20 % PEG 2000	31 % PEG 400	32 % PEG 400 0.2 M $(NH_4)_3PO_4$	14 % PEG4000
0.2 M $MgCl_2$, 0.005 M $CdCl_2$ 0.1 M Tris	0.1 M KAc	0.08 M $MgSO_4$	0.02 M MES	0.05 M Tris	0. 1 M $NaPO_4$ 0.1 M KPO_4	0.1 M NaCl	0.1 M LiCl	0.2 M $(NH_4)_2SO_4$	0.2 M $LiSO_4$	0.1 M $(NH_4)_2SO_4$ 0.1 M Na Citrate	0. 05 M Na citrate
pH 7.5	0.01 M KCl	0.02 M NaCl 0.02 M MES	pH 6.5	pH 8.5	0.1 M Bis-Tris Propane	0.1 M LiSO4	0.1 M NaAc	0.1 M NaCl	0.1 M NaCl	pH 4.5	0.12 M KCl
	0.02 M Tris	pH 6			pH 7.5	0.1 M ADA	pH 4.5	0.1 M Na Citrate	0.1 M HEPES pH 7.0		0.08 M Bis-Tris
	pH 7					pH 6.5		pH 6			pH 6.0
B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12
19 % PEG 1000	18 % PEG 2000 MME	3.5 % PEG 3350	14 % PEG 350 MME	35 % PEG 550 MME	28 % PEG 400	25 % PEG 350 MME	36 % MPD	11 % PEG 8000	26 % PEG 400	32 % PEG 400	3 % PEG 4000
0. 1 M NaCl	0.01 M $NiSO_4$	0.02 M $MgCl_2$	0.02 M NaCl	0.025 M $MgCl_2$ 0.02 M MOPS	0.03 M $MgCl_2$	0.04 M NaCl 0.04 M Tris	0.04 M MgAc 0.1 M MES	0.05 M ZnAc	0.05 M MgAc	0.05 M MgAc	0.066 M NaCl
0.15 M $(NH_4)_2SO_4$	0.1 M NaCitrate pH 6.0	0.02 M MES	0.05 M ME S pH 5.5	pH 7.0	0.1 M MES	pH 8	pH 6	0.05 M ADA	0.1 M MES	0.1 M Glycine pH 9.5	0.02 M Tris
0.01 M MES		pH 6.0			pH 6.5			pH 6.3	pH 6.5		pH 7.5

(continued)

Table 1
(continued)

A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12
pH 6.5											
C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
30 % PEG 2000 MME	14 % PEG5000 MME	2.0 M $(NH_4)_2SO_4$	22 % PEG250 DME	13 % PEG 8000	17 % PEG 3350	18 % PEG 4000	18 % PEG 200	22 % PEG 4000	22 % PEG 8000	23 % PEG 3350	32 % PEG 400
0.075 M $MgCl_2$	0.08 M MgAc	0.05 M ZnAc	0.087 M $(NH_4)_2SO_4$	0.1 $MgCl_2$	0.1 M Mg formate	0.1 M KCl	0.1 M KCl	0.1 M MgAc	0.1 M CaAc	0.1 M $(NH_4)_2SO_4$	0.1 M KCl
0.1 M Na Cacodylate	0.1 M Na citrate pH 6.0	0.1 M MES	0.5 M Tris	0.1 M Tris	0.1 M MOPS	0.1 M Bis-Tris	0.1 M KPO_4	0.1 M MES	0.1 M MES	0.1 M HEPES	0.1 M MES
pH 6.5		pH 6	pH 7.0	pH 7.5	pH 7	pH 6	pH 7.5	pH 6.0	pH 6	pH 8.5	pH 6
D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12
36 % PEG 300	45 % PEG 550 MME	35 % PEG 400	13 % PEG 4000	14 % PEG 2000 MME	19 % PEG 6000	19 % PEG 6000	19 % PEG 3350	20 % PEG 350 MME	20 % PEG 3350	33 % PEG 1000	24 % PEG 400
0.1 M NaCl	0.1 M NaCl	0.15 M $CaCl_2$ 0.1 M Glycine pH 9.0	0.2 M $(NH_4)_2SO_4$	0.2 M choline Cl	0.2 M NaCl	0.05 M NaCl	0.2 M Mg formate 0.05 M Tris	0.2 M $CaCl_2$	0.2 M $(NH4)_2NO_3$	0.5 M $MgCl_2$	0.2 M CaAc
0.1 M MES	0.1 M Bicine		0.5 M ADA	0.1 M Tris	0.05 M MOPS	0.05 M MOPS	pH 8	0.1 M MES	0.05 M HEPES	0.02 M LiCl	0.1 M HEPES
pH 6.5	pH 9		pH 6.5	pH 7.5	pH 7.0	pH 7.0		pH 5	pH 7.0	0.02 M Glycine	pH 7
										pH 10	
E1	E2	E3	E4	E5	E7	E7	E8	E9	E10	E11	E12

28 % PEG 400	29 % PEG 400	29 % PEG 400	31 % pentaery thritol ethoxylate	35 % PEG 3350	38 % PEG 400 0.2 calcium acetate—none 0.1 Sodium Acetate 5.0	38 % PEG 400 0.2 sodium chloride—none 0.1 MOPS 7.5	2.0 M Ammonium Sulfate 0.2 sodium chloride—none 0.1 Sodium Cacodylate 6.5	12 % PEG 4000 0.225 ammonium sulfate—none 0.05 Sodium Acetate 4	33 % PEG 400	22 % PEG 3000 0.25 M Mg formate	40 % PEG 1000 0.25 M $MgCl_2$
0.2 M NaAc	0.2 M NaCl	0.2 M NaCl	15/04 0.2 ammonium formate—none 0.1 Tris 7	0.2 M $(NH_4)_2SO_4$—none 0.1 Tris 8.5					0.23 M NaCl	0.1 Na Cacodylate	0.1 M Tris
0.1 M MES	0.05 M CaAc	0.1 M HEPES							0.05 M NaAc	pH 6.5	pH 8.5
pH 6.5	pH 5.0	pH 7							pH 4.5		
F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12
25 % PEG 400	33 % PEG 550 MME	34 % PEG 400	14 % PEG 4000 0.32 M LiCl	12 % PEG 4000 0.34 M $(NH_4)_2SO_4$	11 % PEG 600	22 % PEG 400	10 % PEG 3350	32 % PEG 400	12 % PEG 400	15 % PEG 4000	16 % PEG 4000
0.3 M $LiSO_4$	0.3 M NH_3 formate	0.3 M BaCl2	0.1 M Na Citrate	0.1 M Na citrate pH 5.5	0.35 M $LiSO_4$	0.37 M KNO_3	0.4 M $(NH_4)_2SO_4$	0.05 M NaCl	0.4 M KCl	0.4 M ammonium thiocyanate	0.4 M Na thiocyanate
0.1 M MES	0.05 M Tris	0.1 M MES	pH 5.5		0.1 M NaAc	0.1 M MES	0.1 M MES	0.04 M $MgCl_2$	0.05 M HEPES pH 7.50	0.1 M NaAc	0.1 M NaAc
pH 6.5	pH 9.0	pH 6			pH 4.5	pH 6.5	pH 6.5	0.1 M HEPES		pH 4.5	pH 4
								pH 7.5			
G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12

(continued)

Table 1
(continued)

A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12
20 % PEG 400	21 % PEG 350 MME	11 % PEG 4000	9 % PEG 8000	11 % PEG 20000	13 % PEG 400	14 % PEG 6000	17 % PEG 350 MME	22 % PEG 300	24 % PEG 400	24 % PEG 1500	28 % PEG 600
0.5 M KCL	0.5 M $MgCl_2$	0.8 M K formate	0.1 M MOPS	0.1 M MES	0.1 M MES	0.1 M ADA	0.05 M Tris	0.07 M Na citrate pH 4.5	0.05 M ADA	0.1 Na Cacodylate pH 6.5	0.1 M HEPES
0.05 M HEPES	0.05 M Tris	0.1 M NaAc	pH 7	pH 6.0	pH 6.5	pH 5.5	pH 7.5		pH 6.5		pH 7.5
pH 6.5	pH 7.5	pH 5.0									
H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12
28 % PEG 400	30 % PEG 400	31 % PEG 600	32 % PEG 550 MME	33 % PEG 400	34 % PEG 3350	44 % PEG 200	65 % MPD	2.75 M NH_4Cl	2.8 M NH_4Cl	3.0 M $(NH_4)_2SO_4$ 0.1 M MES	3.25 M 1,6-Hexanediol
0.05 M Tris	0.1 M Bicine	0.1 M ADA	0.10 M Tris	0.1 M HEPES	0.18 M Na citrate pH 4.0	0.1 M Tris	0.1 M Tris	0.0 25 M Bis-Tris pH 7	0.075 M HEPES	pH 5.5	0. 01 M HEPES
pH 8.5	pH 9	pH 7.0	pH 8.5	pH 7.5		pH 8.5	pH 8		pH 7.5		pH 7.5

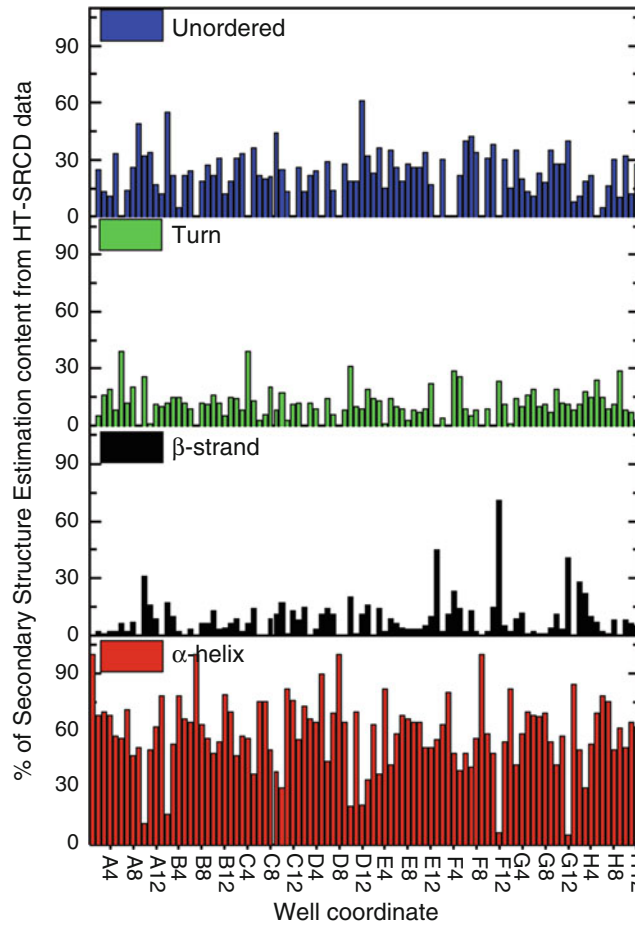

Fig. 3 Bar chart of the percentage of Protein Secondary Structure Estimation (PSSE) for unordered (*blue*), turn (*green*), β-strand (*black*), and α-helix (*red*) conformations from 96 SRCD data of myoglobin in MemGold2™ crystallization solvent conditions (*see* Fig. 2) using CONTIN/LL [5]

In some of the wells (A6, A10, B2, C4, D2, D10, E12, F5, F11, F12, H2, and H3 of Fig. 2), the high salt content interfered with the measurement leading to an underestimate of the alpha helical content. Spectral CD changes in the far-UV region associated with protein folding were observed in several wells (Fig. 3). The protein secondary structure content estimated from CD data indicated the appearance of significant β-sheet content for myoglobin under several formulation conditions (Fig. 3).

The Peltier 6-cell turret holder, introduced by On-Line Instrument Systems [www.olisweb.com] since 2005, available at B23 since 2011 and from 2013 for Chirascan and Jasco CD instruments [www.jasco.uk], can be used to carry out experiments as a function of many variables such as temperature, pH, and ligand interaction. The 6-cell Peltier turret on B23 (Fig. 4) allows the

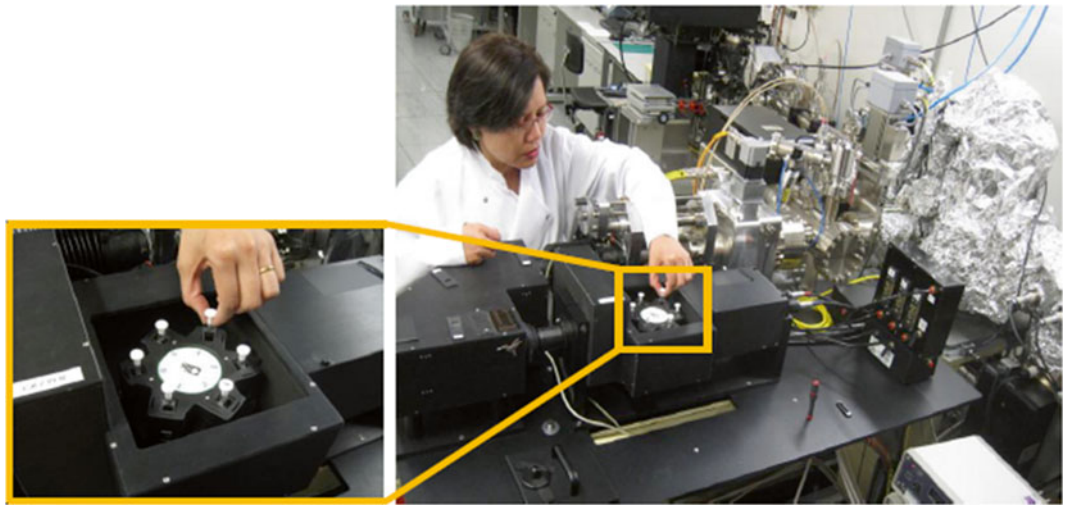

Fig. 4 6-Cell Turret of Diamond B23 module B beamline. The turret is used for the SRCD UV-protein denaturation assay experiment in the far-UV region and/or variable temperature measurements in the 5–95 °C temperature range

measurements of six rectangular cuvette cells from 0.1 to 10 mm path length [Hellma (www.hellma-analytics.com) and Starna (www.starnacells.com)] as well as demountable from 0.01 to 0.1 mm path length [Hellma]. The measurements can be conducted on the same sample as a function of solvent, buffer, pH, concentration, ligand binding interaction, and temperature. One of the important applications of B23 beamline, the protein UV-denaturation assay [5–7], makes use of the 6-cell turret. In Fig. 5a are illustrated the conformational behaviors of a monoclonal antibody (Mab1) under six different formulation conditions when irradiated in the far-UV region under the same parameters of photon flux, irradiation time number of repeated consecutive scans, and antibody concentration. Each sample was irradiated by scanning 30 consecutive repeated spectra in the 180–260 nm region corresponding to an irradiation time of approximately 90 min for that wavelength region. The collapse of the positive CD band at about 200 nm associated with the π–π^* transition of the β-sheet conformation is a direct indication of the loss of secondary structure as the antibody unfolds and can be related to a reduction in protein stability. The complete experiment was carried out overnight over 9 h as a single multiscript experiment that controlled the operation of the B23 module B beamline. The rate of UV-irradiation assay revealed that Mab1 in buffer formulation EC4 was the most stable while Mab1 in EC6 the least stable (Fig. 5b). In terms of relative stability, the six formulations can be ranked qualitatively as follows: EC4 > EC5 > EC1 = EC2 ≫ EC6.

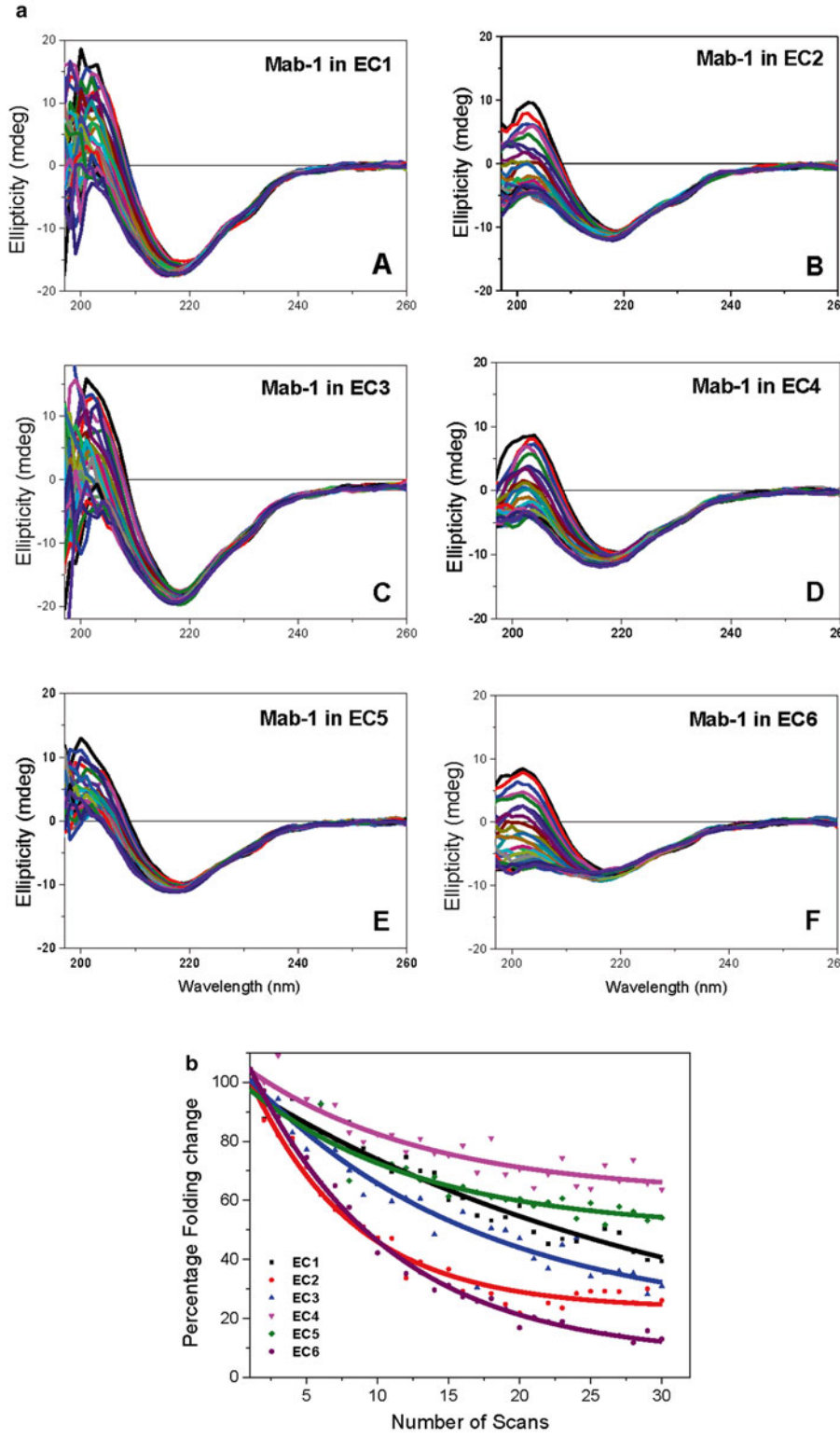

Fig. 5 (**a**) SRCD UV-denaturation assay in the far-UV region of a monoclonal antibody (Mab1) in six different formulations (EC1 to EC6). In this example the assay consisted of 30 repeated consecutive scans measured with Diamond B23 beamline. (**b**) Rates of UV-denaturation for Mab1 in six different formulations. From the rate of UV-denaturation, Mab1 in EC6 is the least stable while Mab1 in EC4 is the more stable one

4.2 Quality Control of Protein Folding and Determination of Ligand Binding by SRCD

The study of the function/activity of recombinant proteins in solution includes the analysis of folding that is the secondary, tertiary, and quaternary structure, determination of ligand binding all of which can be measured by SRCD. Critically, if several different batches of a protein are studied, it is important to establish that protein folding is the same for each batch. Subtle changes in the purification procedure can, for instance, produce protein batches of different quality and one effect that is commonly observed is a change in protein folding. CD spectroscopy is the ideal technique to characterize and estimate directly the protein folding in terms of secondary structure content and in particular the conformational changes as a function of environmental perturbations such as solvent composition, ionic strength, concentration, pH, temperature, chemical and detergent agents, and ligand interaction.

Ligand binding interaction can be determined both qualitatively and quantitatively with the method developed by Siligardi used to calculate dissociation constant (Kd) following [15–17]. In this method, the data of the CD titrations are reported as difference CD spectra calculated by subtracting from the spectra of the protein–ligand mixtures the equivalent spectra for each addition of the chiral ligand. In this way the spectral changes observed in the difference CD spectra are unambiguously indicative of ligand binding. The dissociation constant Kd is determined from the CD data at fixed wavelength as a function of ligand concentration using a nonlinear regression analysis [15]. Many other techniques, such as fluorescence, SPR, ITC, and AUC, are used to determine protein ligand binding interactions; however, only CD spectroscopy determine directly whether the interaction has induced changes in protein secondary and/or tertiary structures.

Reproducible gel filtration and mass spectrometry data for different batches of a recombinant protein do not necessarily mean that folding of the proteins is identical. Therefore, the implementing a quality control (QC) step for folding is therefore a must. The following examples illustrate some of the common problems that are often encountered in studying recombinant proteins in solution that have been revealed by CD spectroscopy.

4.2.1 Example 1: Yeast Heat Shock Protein 90 (hsp90)

Heat shock protein 90 (hsp90) is a ubiquitous and abundant molecular chaperone that mediates protein folding and activation of many signal transduction and cell regulatory proteins. Figure 6 shows the CD spectra of nine recombinant hsp90 batches produced over 3 years (from January 1998 till February 2001) during collaborative research with Dr. Christopher Prodromou and Professor Laurence Pearl [18]. The CD spectra recorded in the near-UV region were characteristic of the local tertiary structure of the side chains of aromatic amino acid residues (Trp, Tyr, and Phe) [19] and dihedral angle of disulfide bonds [5]. Out of the nine batches, two batches (09/02/2000 and 22/02/2000 of Fig. 6) showed significant CD spectral changes in the near-UV region at

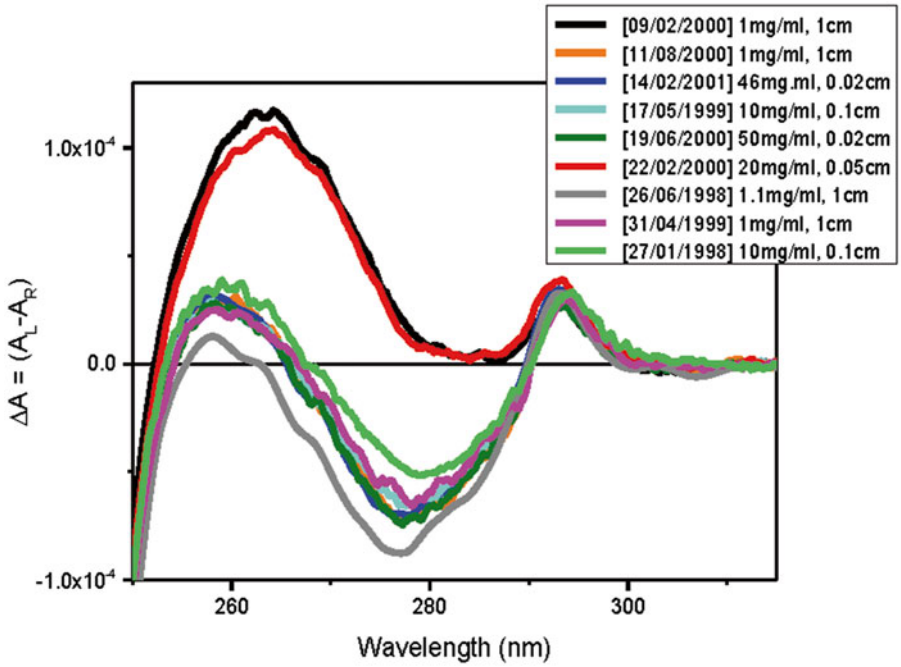

Fig. 6 CD spectra in the near-UV region of nine batches of recombinant Hsp90 protein measured within 4 years period from 1998 to 2001. The spectra were measured with Jasco J720 CD spectropolarimer

around 280 nm associated with differences in local tertiary structure of Tyr side chain residues compared to the other similar six batches. However, the protein binding properties of hsp90 were not affected as the two batches representing the two types of local tertiary structure (batches 22/02/2000 and 19/06/2000 of Fig. 6) still bound to hP50 protein with the same dissociation constant Kd of 250 nM (Fig. 7). In addition, there was difference in the ATPase activity of the two batches of hsp90. One interesting observation was that for batch 22/02/2000 (Fig. 6) the complex remained soluble throughout the titration with hP50 while for the other hsp90 batch the complex started to precipitate at higher hP50 molar ratio than 1:1 [15].

4.2.2 Example 2: Link Domain of TSG-6 Glycoprotein

The Link module of human TSG-6 glycoprotein is involved in the formation of the extracellular matrix and cell migration by interacting with hyaluronan 10 (HA_{10}) [20]. This interaction was studied using two different batches of the Link module protein (batch 1 and 2) by CD spectroscopy. The CD spectral profiles of the two protein batches were not superimposable in the near-UV region (Fig. 8) indicating significant differences in the local structure of the side chains of aromatic residues. The Link module glycoprotein contains two tryptophan (Trp), eight tyrosine (Tyr), and three phenylalanine residues (Phe) (1tsg.pbd). The positive CD bands of the two batches

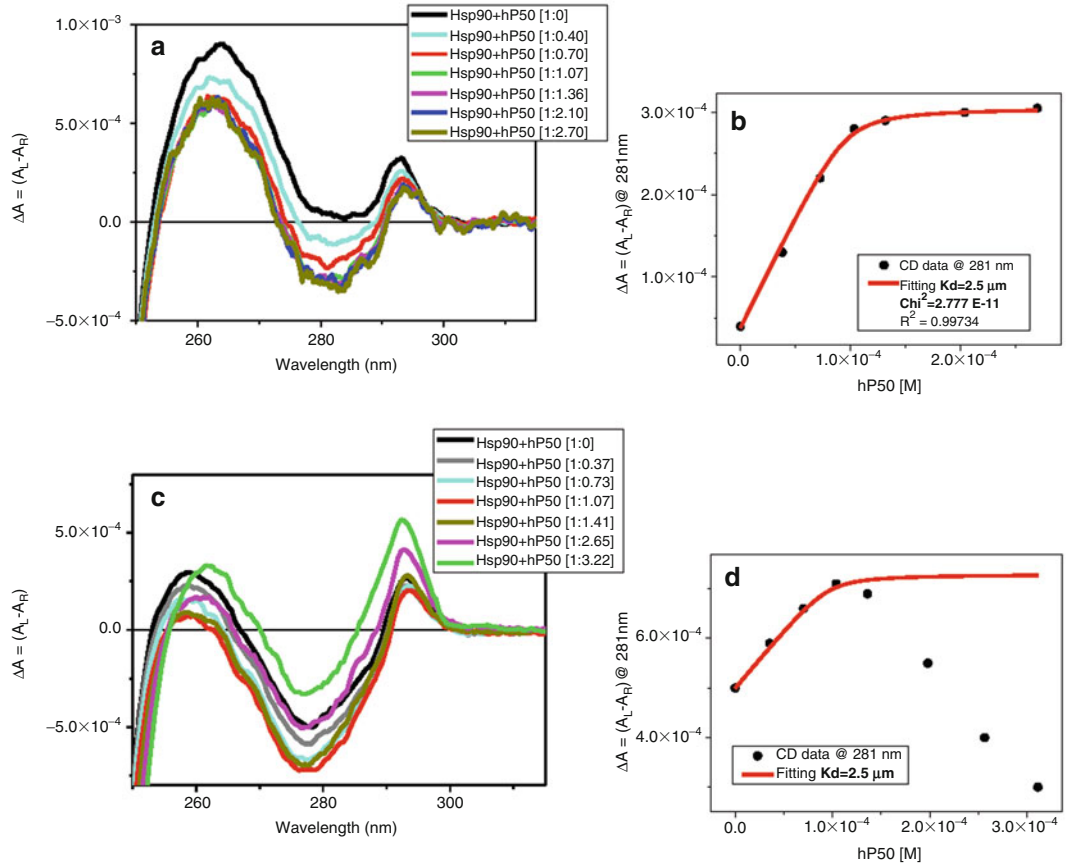

Fig. 7 (**a**) Difference CD spectra of the titration of hP50 into Hsp90 (batch 22/02/2000). The difference spectra were calculated subtracting from the each spectrum of (Hsp90 + hP50) mixtures at [1:*n*] molar ratio the equivalent spectrum of hP50 at [*n*] molar ratio. (**b**) Plot of difference CD intensity reported in $\Delta A = (A_L - A_R)$ measured at 281 nm versus the ligand hP50 concentration. The best fitting of the experimental data of (**a**) calculated using the nonlinear regression method of Siligardi et al. [15] was achieved for a Kd = 0.250 μM and $R^2 = 0.997$. (**c**) Difference CD spectra of the titration of hP50 into Hsp90 (batch 19/06/2000). (**d**) Plot of difference CD intensity reported in $\Delta A = (A_L - A_R)$ measured at 281 nm versus the ligand hP50 concentration. The best fitting of the experimental data of figure (**c**) calculated using the nonlinear regression method of figure (**b**) was achieved for a Kd = 0.250 μM. Due to protein association [15], only the first four experimental data were fitted and therefore R^2 could not be calculated

at around 298 nm, assigned to the Trp residues, indicated a similar local tertiary structure (Fig. 8a), whereas a major CD spectral difference below 290 nm, assigned to the Tyr residues, indicated a significantly different local tertiary structure (Fig. 8a). The affect on hyaluronan binding was investigated by CD spectroscopy.

For the Link protein batch 1, the addition of $(HA)_{10}$ at molar ratio of [1:1] induced an increase in the intensity of the 285 nm positive CD band assigned to a Tyr residue that reached saturation at 2:1 molar ratio (Fig. 8b). The fact that the rest of the CD spectrum remained unchanged indicated that the other aromatic residues in the protein the two Trp residues and the majority of Tyr

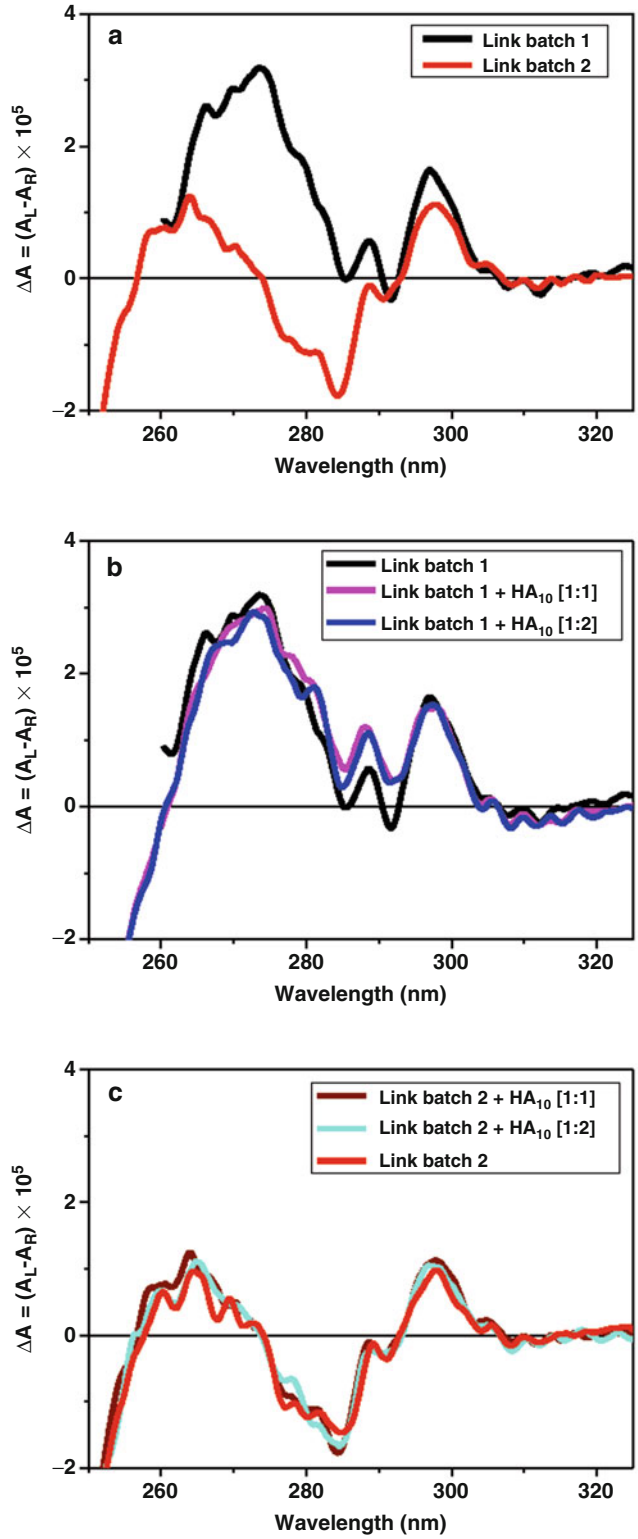

Fig. 8 (**a**) Near-UV CD spectra of batch 1 (*black*) and 2 (*red*) of TSG-6 Link Module protein. (**b**) CD spectra of bath 1 of TSG-6 Link module (*black*) with 1 (*pink*) and 2 (*blue*) molar ratios of hyaluronan 10 (HA_{10}). (**c**) CD spectra of bath 2 of TSG-6 Link module (*red*) with 1 (*brown*) and 2 (*cyan*) molar ratios of hyaluronan 10 (HA_{10}). The CD spectra were measured with Jasco J720 spectropolarimer

residues were not involved in the ligand binding interactions (Fig. 8b). This was consistent with the NMR data of the complex formation [21]. For the Link protein batch 2, the hyaluronan binding interaction study showed no detectable CD changes in the near-UV region (Fig. 8c). This does not rule out some interaction between the Link protein batch 2 and hyaluronan as in the far-UV region, the CD spectra of Link batch 2 with and without hyaluronan were not superimposable indicating ligand binding interaction (data not shown). Thus the CD measurements in the far- and near-UV regions confirmed the hyaluronan binding property of the Link module protein but it also revealed folding differences between the two samples of the protein.

4.2.3 Example 3: K+ Transporter TrkA Peripheral Membrane Protein

TrkA is a peripheral membrane protein of the Trk system that requires ATP for the transport of K^+ in prokaryotic and eukaryotic cells [22]. Five samples (p001, p002, p003, p005, and p006) of purified *H. influenzae* TrkA peripheral membrane protein in 500 mM NaCl, 50 mM Tris pH 7.5, 5 % glycerol, and 0.1 mM TCEP solutions were investigated using CD spectroscopy. The protein does contain, nine tyrosine residues and but no tryptophans or disulfide bonds [23]. In the near-UV region (250–315 nm) the CD spectra of the five equimolar solutions of TrkA were not identical (Fig. 9) indicating local conformational differences in the

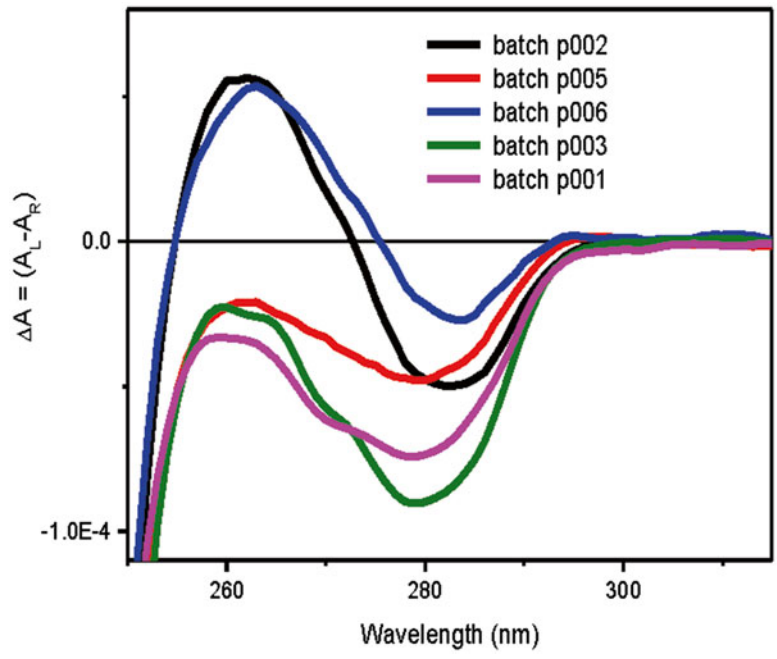

Fig. 9 Near-UV CD spectra of five batches (batch p001 (*pink*), p002 (*black*), p003 (*green*), p005 (*red*), and p006 (*blue*)) of 20 μM TrkA peripheral membrane protein in 500 mM NaCl, 0.5 % glycerol, 0.1 mM TCEPT, and 50 mM Tris measured in 1 cm pathlength cell using Chirascan Plus CD spectropolarimeter (Applied Photophysics Ltd)

aromatic side chains of the Tyr residues. To investigate the effect of these conformational differences on the function of TrkA protein, a CD titration with ATP ligand was conducted for each protein batch. ATP was found to bind to TrkA and two main behaviors were observed. In Fig. 10, the data for the difference CD spectra, calculated by subtracting from each mixture of [TrkA+ ATP] at different (1:*n*) molar ratios the equivalent spectra for ATP (*n*),, were fitted using a nonlinear regression method [15] (inserts of Fig. 10). The results for Sample p003 indicated that the TrkA has a single ATP binding site whereas the data for samples p001 and p006 (Fig. 10) suggested that there are two ATP binding sites. The results for the other batches of TrkA fitted with either a one-site or two-site model (Fig. 11). From the local tertiary structure, the batches of TrkA protein could be clustered into two groups: p002 with p006 and p001 with p003 and p005 as having qualitatively similar CD spectra (Fig. 9) while according to ATP binding stoichiometry, the groups could be rearranged differently: p002 with p003 and p001 with p005 and p006 (Fig. 11). Further studies will be required to clarify the functional significance of these observations but these behaviors would not have been fully revealed without the analysis of different batches of protein by CD spectroscopy.

5 Summary

HTCD offers a rapid way of assessing protein folding in solution and the effect of buffer conditions on secondary structural features. This in turn may inform how a protein sample behaves in crystallization trials. HTCD also allows the screening of the binding properties of the proteins in solutions under different conditions including in crystallization buffers. Another important application of HTCD is monitoring variations in protein folding between different batches of the same protein. In fact, the quality control of conformation and binding interactions of recombinant proteins by CD spectroscopy ought to play a more important role in structural biology.

Acknowledgements

We would like to thank Dr. Christopher Prodromou and Prof Laurence Pearl for providing the batches of Hsp90 proteins, Prof Antony Day for TSG-6 Link Module protein, and Dr. Petra Lukacik and Dr. Martin Walsh for TrkA protein. We would like to thank Dr. Tamas Javorfi for his assistance in commissioning the vertical chamber for Diamond B23 module A beamline.

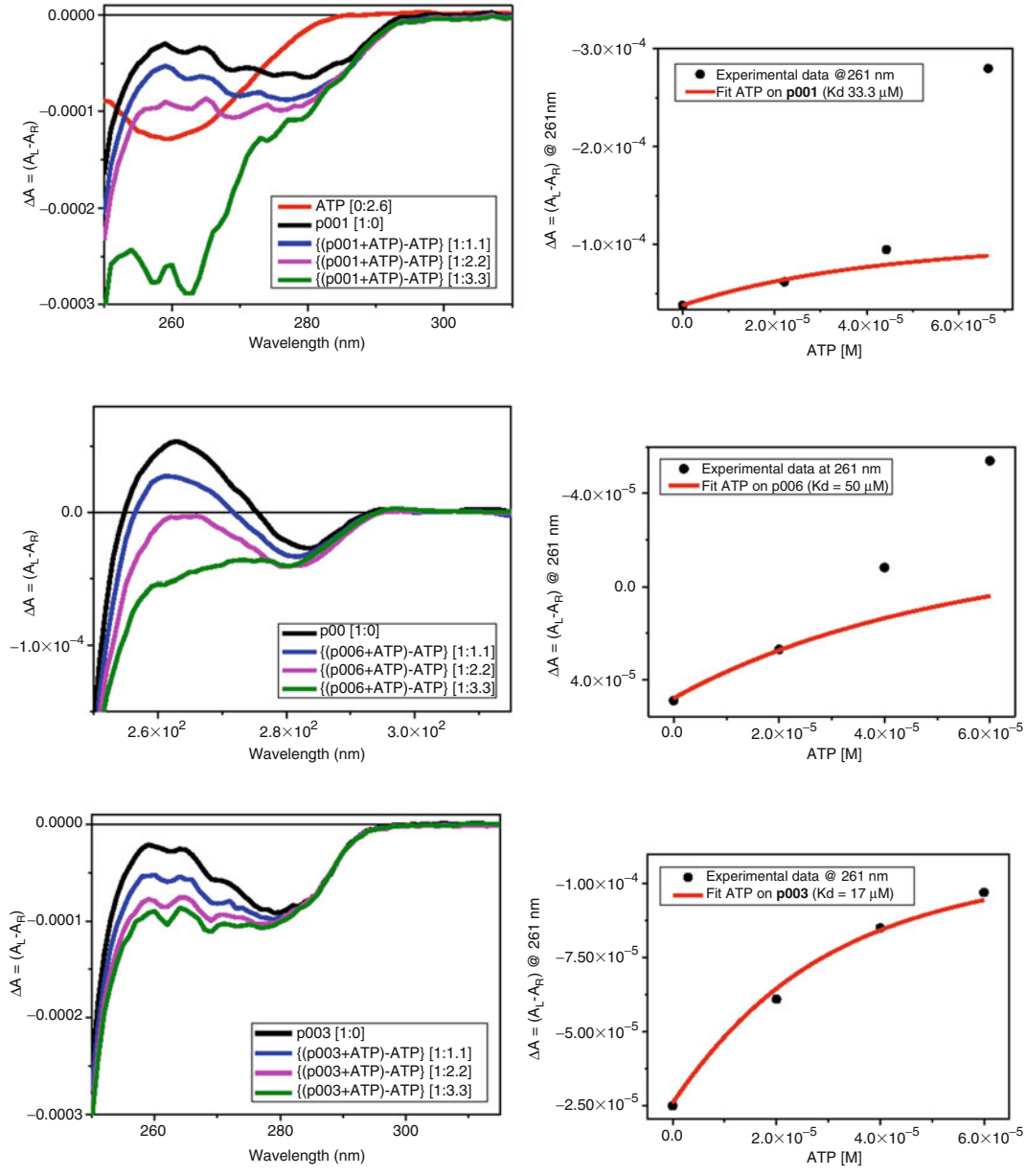

Fig. 10 (*Top left*) Difference CD spectra of the titration of ATP into TrkA (batch p001) in a 1 cm pathlength cell. The difference spectra were calculated subtracting from the each spectrum of (TrkA+ ATP) mixtures at [1:*n*] molar ratio the equivalent spectrum of ATP at [*n*] molar ratio. (*Top right*) Plot of difference CD intensity reported in $\Delta A = (A_L - A_R)$ measured at 261 nm versus the ligand ATP concentration. The best fitting of the experimental data calculated using the nonlinear regression method of Siligardi et al. [15] was achieved for the first two data supportive of ATP stoichiometry of 2. (*Middle left*) Difference CD spectra of the titration of ATP into TrkA (batch p006). (*Middle right*) Plot of difference CD intensity reported in $\Delta A = (A_L - A_R)$ measured at 261 nm versus ATP concentration. The best fitting of the experimental data was achieved like for batch p001 for the first two data supportive of ATP stoichiometry of 2. (*Bottom left*) Difference CD spectra of the titration of ATP into TrkA (batch p003). (*Bottom right*) Plot of difference CD intensity reported in $\Delta A = (A_L - A_R)$ measured at 261 nm versus ATP concentration. The best fitting of all experimental data was achieved indicating an ATP stoichiometry of 1

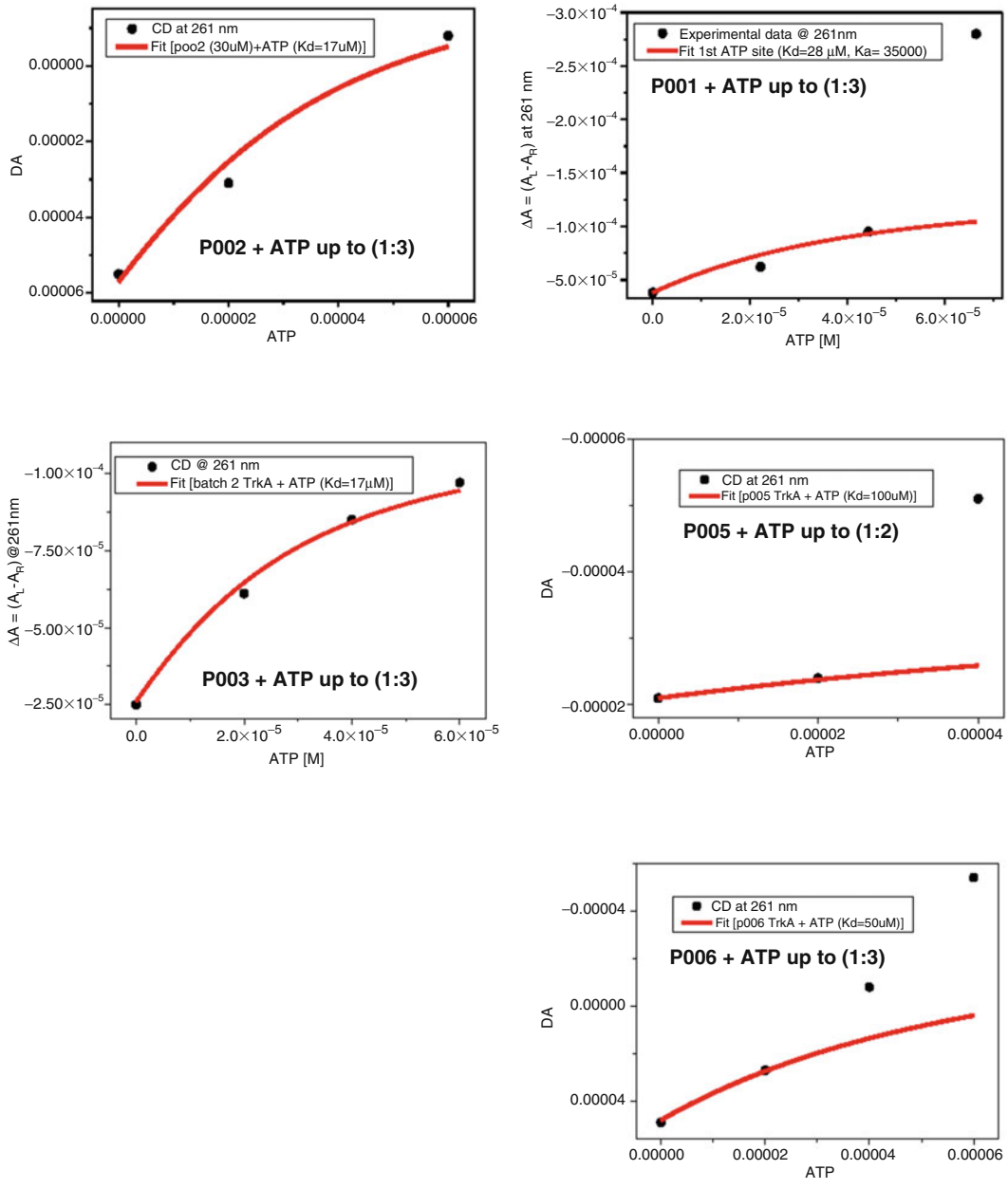

Fig. 11 Fitting of the experimental data of the plot CD data in ΔA versus ATP concentration for all five TrkA batches. The fittings of the experimental data measured with Chirascan Plus were calculated using the non-linear regression method of Siligardi et al. [15]

References

1. Carvalho AL, Trincão J, Romão MJ (2009) X-ray crystallography in drug discovery. Methods Mol Biol 572:31–56
2. Pellecchia M, Bertini I, Cowburn D et al (2008) Perspectives on NMR in drug discovery: a technique comes of age. Nat Rev Drug Discov 7:738–745
3. Pennisi E (2012) ENCODE project writes eulogy for junk DNA. Science 337: 1159–1160
4. Fasman GD (1996) Circular dichroism and the conformational analysis of biomolecules. Plenum, New York
5. Hussain R, Jávorfi T, and Siligardi G (2012) Spectroscopic Analysis: Synchrotron Radiation Circular Dichroism. In: Carreira E.M. and Yamamoto H. (eds.) Comprehensive Chirality, Volume 8, pp. 438–448. Amsterdam: Elsevier
6. Hussain R, Jávorfi T, Siligardi G (2012) Circular dichroism beamline B23 at the Diamond light source. J Synchrotron Radiat 19:132–135
7. Calzolai L, Laera S, Ceccone G et al (2013) Gold nanoparticles' blocking effect on UV-induced damage to human serum albumin. J Nanoparticle Res 15:1412–1416
8. ICH, Topic Q1B, EMEA (1998) CPMP/ICH/279/95
9. Siligardi G, Campbell MM, Gibbons WA, Drake AF (1991) Conformational analysis of the melanin concentrating hormone (MCH) by CD spectroscopy: disulphide bridge and aromatic tyrosyl contributions. Eur J Biochem 206:23–29
10. Sreerama N, Woody RW (2000) Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON and CDSSTR methods with an expanded reference set. Anal Biochem 287:252–260
11. Fiedler S, Cole L, Keller S (2013) Automated Circular Dichroism spectroscopy for medium throughput analysis of protein conformation. Anal Chem 85:1868–1872
12. Kabsch W, Sander C (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 22:2577–2637
13. Provencher SW, Glockner J (1981) Estimation of globular protein secondary structure from circular dichroism. Biochemistry 20:33–37
14. Van Stokkum LHM, Spoelder HJW, Bloemendal M et al (1990) Estimation of protein secondary structure and error analysis from CD spectra. Anal Biochem 191:110–118
15. Siligardi G, Panaretou B, Meyer P et al (2002) Regulation of Hsp90 ATPase activity by the co-chaperone Cdc37p/p50cdc37. J Biol Chem 277:20151–20159
16. Siligardi G, Hussain R (2010) Lindon J, Tranter G, Koppenaal D (eds) In: Applications of circular dichroism encyclopedia of spectroscopy and spectrometry, 2nd ed. vol 1. Elsevier, Oxford. pp 9–14
17. Martin SR, Schilstra MJ, Siligardi G (2011) Chapter 7: circular dichroism. In: Podjarmy A, Dejaegere A, Kieffer B (eds) Biophysical approaches determining ligand binding to biomolecular targets, detection, measurement and modelling. RSC Publishing, Cambridge, pp 226–246
18. Prodromou C, Siligardi G, O'Brien R et al (1999) Regulation of Hsp90 ATPase activity by tetratricopeptide repeat (TPR)-domain co-chaperones. EMBO J 18:754–762
19. Strickland EH (1974) Aromatic contributions to circular dichroism spectra of proteins. CRC Crit Rev Biochem 2:113–175
20. Kohda D, Morton CJ, Parkar AA et al (1996) Solution structure of the link module: a hyaluronan-binding domain involved in extracellular matrix stability and cell migration. Cell 86:767–775
21. Blundell CD, Mahoney DJ, Andrew A et al (2003) The link module from ovulation- and inflammation-associated protein TSG-6 changes conformation on hyaluronan binding. J Biol Chem 278:49261–49270
22. Stewart LM, Bakker EP, Booth IR (1985) Energy coupling to K+ uptake via the Trk system in Escherichia coli: the role of ATP. J Gen Microbiol 131:77–85
23. Bertrand T, Kothe M, Liu J et al (2012) The crystal structures of TrkA and TrkB suggest key regions for achieving selective inhibition. J Mol Biol 423:439–453