THE ANALYSIS OF THE FAR-ULTRAVIOLET CIRCULAR DICHROISM SPECTRA OF PROTEINS

D.G. DALGLEISH

Physical Biochemistry Laboratory, Nuffield Department of Clinical Biochemistry Annexe, University of Oxford, Old Radcliffe Observatory, 43 Woodstock Road, Oxford, OX2 6HF, England

Received 10 April 1972

1. Introduction

It has generally been assumed that three components are sufficient to analyse the circular dichroism (CD) spectra of proteins, and on this basis curve-fitting methods have been developed for the determination of α-helix and β-sheet contents [1–3], none of which are completely satisfactory. I have used a matrix rank analysis [4] of the far-ultraviolet CD spectra of several proteins to estimate the number of contributing components. This analysis showed that four parameters, rather than three, must be used in the majority of cases. However, some of the CD spectra studied cannot be fitted even by these four parameters. It is clear, therefore, that the determination of regular structure contents of proteins from analysis of their CD spectra must await further determination of the nature and sources of the components involved.

2. Methods

The most frequently used method of three-parameter fitting is based on the assumption that the α-helical, β-sheet and unordered conformations of poly-L-lysine are suitable representations of the α-helical, β-sheet and non-regular protions of protein tertiary structure. The CD spectra of the three reference conformations are then used as the components in a three-parameter fit to the observed CD spectra of proteins [1]. Alternatively, poly-L-serine in 8 M lithium chloride can be used to represent non-regular structure, and some improvement is found in the analysis [3]. Analysis of the CD spectra of proteins whose tertiary structures are known from X-ray crystallography can yield estimates of the CD spectra of the α-helix, β-sheet and non-regular conformations, and these can then be used to analyse the CD spectra of other proteins [2].

However, it has not yet been clearly established that three parameters are necessarily sufficient for the general fitting of the CD spectra of proteins, and it has been pointed out that contributions from other structures should be expected. It was therefore considered worthwhile to examine the three-parameter assumption, and to determine how many components are needed to fit the CD spectra of most proteins. The technique used was matrix rank analysis [4, 5].

The CD spectra used in the analysis were those of 27 solutions of histone fractions from chicken erythrocyte, at different ionic strengths and pH values (see Appendix), together with those of myoglobin, lysozyme, and ribonuclease given in [2]. In addition, the CD spectra of β-lactamases [6] and bovine [7] and human [8] erythrocyte cupro–zinc protein were included. These CD spectra were arranged in a matrix format, using values of ellipticity at 2.5 nm intervals from 230 to 200 nm, and the number of contributing components was estimated by the reduction of the matrix as described in [5].

3. Results and discussion

The results may be summarized thus:

a) For the histone spectra alone, three components were

North-Holland Publishing Company – Amsterdam

needed to fit the observed CD spectra between 230 and 205 nm: inclusion of points at 202.5 and 200 nm gave evidence that another contribution was effective in this region;
b) when the spectra of myoglobin, lysozyme and ribonuclease were included with the histone spectra, four components were needed for the analysis between 230 and 205 nm;
c) the CD spectra of the α-helix, β-sheet and 'remainder' conformations as calculated by Saxena and Wetlaufer [2] did not satisfactorily fit the histone spectra;
d) four components were needed to fit a matrix composed of the CD spectra of three histone factions, myoglobin, lysozyme, ribonuclease, and two β-lactamases;
e) erythrocyte cupro–zinc protein CD spectra were not satisfactorily fitted by the components of (b) or (d): this misfitting was not caused by additional copper transitions being present since the apoprotein CD spectra also gave bad fits.

These results show that general three-parameter fitting of the CD spectra of proteins is not possible, and that at least four components are necessary. Even so, the erythrocyte cupro–zinc proteins and their apoproteins show that it is not possible to fit all protein CD spectra using even four components. If the erythrocyte proteins are excluded from consideration on account of their uncommon amino acid composition, which may give rise to a hitherto undefined regular structure, then the problem of the sources of the other four components remains.

Two of the four are likely to arise from the α-helix and antiparallel-β structures. The CD spectrum of the α-helix is fairly well defined, and similar results are found for its shape from poly-L-lysine or from analysis of the CD spectra of three proteins [1, 2]. The CD spectrum of β-structure is not so well defined, since different results are obtained from poly-L-lysine and the analysis of three spectra [9, 2]. Even if these two structures give rise to two out of the four spectral components found in the calculations, it is at present not possible to define the sources of the other two. They may represent two components into which the general 'remainder' conformation may be divided, or they may arise from a general 'remainder' conformation plus contributions from other regular structures or from aromatic side-chains.

It is clear, therefore, that the results from three-parameter fitting of the far-ultraviolet CD spectra of proteins can only be taken as rather uncertain estimates of the amounts of regular structures present in the molecules. Extension of this work is at present in progress to attempt to define more clearly the structures which give rise to four components.

Acknowledgements

I would like to thank the Medical Research Council for supporting grants (to Dr. A.R. Peacocke). Computing facilities were made available by the Department of Biomathematics and the University Computing Laboratory, Oxford. My thanks are also due to Professor W. Bannister, The Royal University of Malta, for the provision of human and bovine erythrocyte cupro–zinc proteins.

References

[1] N. Greenfield and G.D. Fasman, Biochemistry 8 (1969) 4108.
[2] V. Saxena and D.B. Wetlaufer, Proc. Natl. Acad. Sci. U.S. 68 (1971) 969.
[3] H. Rosenkranz and W. Scholtan, Hoppe-Seyler's Z. Physiol. Chem. 352 (1971) 896.
[4] R.M. Wallace, J. Phys. Chem. 64 (1960) 899.
[5] D.W. McMullen, S.R. Jaskunas and I. Tinoco, Biopolymers 5 (1967) 585.
[6] D.G. Dalgleish and A.R. Peacocke, Biochem. J. 125 (1971) 155.
[7] E. Wood, D. Dalgleish and W. Bannister, European J. Biochem. 18 (1971) 187.
[8] E. Wood, D. Dalgleish and W. Bannister, manuscript in preparation.
[9] L.K. Li and A. Spector, J. Amer. Chem. Soc. 91 (1969) 220.

APPENDIX

D.G. DALGLEISH, D.C. CURSONS and A.R. PEACOCKE

Values of the ellipticity of histones from 230 to 200 nm

The values of the ellipticity of whole histone, and fractions 2a1, 2a2 and 2b from chicken erythrocyte are given in table 1. The measurements were made on solutions of different ionic strengths, at pH 2.0, 4.5 and 7.5. The histones were dissolved to a concentration of 1 mg/ml in NaCl solutions of the given ionic strengths, and the pH was adjusted by direct titration: The ionic strength values given in the table are those of the solutions in which the histones were originally dissolved.

Measurements were made with a Jouan Dichrographe CD 185 instrument at 20°, using cuvettes of path lengths between 10 and 0.1 mm. The values of ellipticity given in table 1 have been changed in sign, since all are minus in the original measuments.

We wish to thank the Science Research Council for a grant to purchase the Dichrographe, and the Medical Research Council for supporting grants to D.G.D. and A.R.P.

Table 1

Values of the ellipticity ($-[\theta]$ in deg cm^2/decimole) for histones.

Fraction	pH	μ	Wavelength (nm) 230	225	220	215	210	205	200
whole	2.0	0	1750	2350	2500	2650	4400	9000	11200
whole	4.5	0	3700	5200	5700	5300	6750	8300	4100
whole	7.5	0	4150	6050	6450	6300	7950	9100	5450
whole	2.0	0.1	2800	4050	4750	4850	5650	8900	8500
whole	4.5	0.1	3500	5200	6000	6200	7050	8750	7000
whole	7.5	0.1	3800	5700	7000	7050	7650	8600	4600
whole	2.0	0.5	4500	6600	7200	7100	8250	9600	4500
whole	4.5	0.5	4850	7200	8000	8100	8900	9550	1800
2a1	4.5	0.05	2850	4200	4700	4500	5900	7300	6850
2a1	7.5	0.05	3200	5000	6200	6700	6950	8350	5800
2a1	4.5	0.1	3300	4900	5600	5350	6300	7850	4500
2a1	2.0	0.5	3550	5800	7650	8500	8200	6100	1600
2a1	4.5	0.5	3300	5600	7150	7900	7350	6800	500
2a1	7.5	0.5	3200	4900	6300	6800	6350	3700	−2300
2a2	2.0	0.5	3500	5500	6300	6200	6300	6900	1100
2a2	4.5	0.5	3700	5900	6600	6600	7200	8000	2500
2a2	7.5	0.5	4000	6250	7000	7000	7800	8700	3800
2a2	2.0	0.05	2500	3700	4200	4300	5000	8100	7800
2a2	4.5	0.05	2700	3800	4300	4400	5300	8400	6800
2a2	7.5	0.05	3800	5700	6200	6200	6800	8200	1800
2b	2.0	0.05	2000	2700	3100	3600	5100	9300	13000
2b	4.5	0.05	2800	4300	4900	4800	6400	8400	9600
2b	7.5	0.05	4400	6800	7900	7600	8400	8200	3000
2b	4.5	0.1	3500	5000	5600	5700	6400	9000	7500
2b	2.0	0.5	4100	7100	8000	7800	8700	10000	4900
2b	7.5	0.5	4700	7200	8400	8200	8700	8000	1100
2b	4.5	0.5	5200	8000	9600	9400	9700	9300	3000