PROTEINS: Structure, Function, and Genetics 13:57–69 (1992)

Deconvolution of the Circular Dichroism Spectra of Proteins: The Circular Dichroism Spectra of the Antiparallel β-Sheet in Proteins

A. Perczel, K. Park, and G.D. Fasman
Department of Biochemistry, Brandeis University, Waltham, Massachusetts, 02254-9110

ABSTRACT **A recently developed algorithm, called Convex Constraint Analysis (CCA), was successfully applied to determine the circular dichroism (CD) spectra of the *pure β-pleated sheet* in globular proteins. On the basis of X-ray diffraction determined secondary structures, the original data set used (Perczel, A., Hollosi, M., Tusnady, G. Fasman, G.D. Convex constraint analysis: A natural deconvolution of circular dichroism curves of proteins, Prot. Eng., *4*:669–679, 1991), was improved by the addition of proteins with high β-pleated sheet content. The analysis yielded CD curves of the pure components of the main secondary structural elements (α-helix, antiparallel β-pleated sheet, β-turns, and unordered conformation), as well as a curve attributed to the "aromatic contribution" in the wavelength range of 195–240 nm. Upon deconvolution the curves obtained were assigned to various secondary structures. The calculated weights (percentages determining the contributions of each pure component curve in the measured CD spectra of a given protein) were correlated with the X-ray diffraction determined percentages in an assignment procedure and were evaluated. The Pearson product correlation coefficients *(R)* are significant for all five components. The new pure component curves, which were obtained through deconvolution of the protein CD spectra alone, are promising candidates for determining the percentages of the secondary structural components in globular proteins without the necessity of adopting an X-ray database. The CD spectrum of the CheY protein was interesting because it has the characteristic shape associated with the α-helical structure, but upon analysis yielded a considerable amount of β-sheet in agreement with the X-ray structure.**

Key words: protein conformation, aromatic contribution, disulfide contribution, CD spectra of the main secondary structural elements in proteins

INTRODUCTION

The conformation of many proteins share highly similar structural motifs; the secondary structural elements, α-helices, β-pleated sheets, and β-turns. In globular proteins these structural units may exhibit a wide variety of super secondary structural forms. They are built more or less from the same conformational "monomers" $(\phi_i, \psi_i)_{1<i<n}$. As an example of the large variation of one structural unit, the β-pleated sheet is present in several forms: parallel-β-sheets[1], antiparallel-β-sheets[1,2], twisted and extended-β-sheets[3], mixed β-sheets forming a saddle-shaped structure, and β-barrels[4–7]. These superstructures can be regarded as "conformational polymers," being built from similar conformational subunits $(\phi_i \simeq 150°, \psi_i \simeq 150°)_{1<i<n}$.

Circular dichroism (CD) spectroscopy is a powerful and unique technique[8] for studying these secondary structural elements, which upon association provide the overall conformation of peptides and proteins in solution. The method, used even at low concentrations (0.1 mg/mL), can yield important information on the back-bone conformation of oligopeptides and proteins. The various *assignments* of the secondary structural elements, which fold into the tertiary structure of proteins, *are mostly based on empirical correlations* with theoretically calculated and/or measured pure component spectra[9–11]. These methods have been extensively evaluated[8–14] since the plausible identification of the pure component curves is of basic importance for further assignment of conformers.

Theoretical calculations[15–33] of extended systems are inaccurate due to the method used for calculating the electric and magnetic dipole transition moments for the isolated chromophores. Geometric distortions (deviation of back-bone torsion angles from the ideal values) could be a cause[34,35] of further com-

Received May 7, 1991; revision accepted July 11, 1991.
Address reprint requests to Gerald D. Fasman, Graduate Department of Biochemistry, Brandeis University, Waltham, MA 02254-9110.

plications in calculating the shape of the pure CD component curves of the α-helix, β-sheet, and β-turns.

The information derived from the CD spectra of synthetic and model peptides yields excellent data[36,37], however the reliability of such information decreases with the increase in the number of amino acids involved in the model, due to their conformational flexibility. On the other hand, the direct use of the CD spectra of some proteins to obtain the pure component curves would not be feasible because of the low probability of finding proteins containing high percentages of one secondary structural element. To study the chiroptical properties of the pure α-helical conformation in an aqueous environment, the CD spectra of myoglobin was successfully introduced with a slight extrapolation, due to its high content of α-helix (≥ 80%)[38]. In contrast, the proteins containing a high percentage of the β-form are rare and usually insoluble in water[39]. Applying any of the existing methods to deconvolute the CD curves of proteins with a low content of β-sheet produces results of questionable reliability. Therefore *the direct use* of the CD spectra of proteins, which have a high β-sheet content as reference spectra, to obtain the pure CD spectrum of the β-pleated sheet, was thought to be implausible.

Sequences of oligopeptides and proteins can adopt three fundamentally different types of β-pleated sheet[7] conformations; antiparallel, parallel, and mixed sheets. Furthermore the twisting of these different β-form units can perturb the shape of the CD curves[4]. Whereas the CD spectra of the α-helix can vary with the length of this secondary structural element[37–40], the intensities of the CD spectra of β-pleated sheet like conformations, in general, show a remarkably high tendency of band intensity fluctuation. This phenomenon was also observed for twisted β-sheets[34]. Therefore the identification of the rotational strengths and their intensities of the pure component curve(s) for the β-conformation is of basic importance for the further assignment of these conformers in proteins.

For a pure component curve of the β-pleated sheet conformation, the CD spectra of homopolypeptides [poly(Lys), poly(Ser), and poly(Cys)] were directly used in early studies[41–43]. The overall shape of the observed CD curves were found to be similar to one another (negative band at 215±3 nm and a positive band near 197±3 nm). However due to the significant variation of the amplitudes, those CD curves cannot be regarded as definitive models for the β-sheet conformation. More recently oligopeptides were introduced as candidates yielding predominantly the β-pleated sheet conformation. Hydrophobic amino acids, Ala, Val, Leu, and Ile, were used[44,45] for the synthesis of Boc-(amino acid)$_n$-OMe, where the value of n was found to have a critical influence on the conformational behavior of the whole molecule; for shorter systems ($n < 6$) the unordered conformation was usually assigned based on the CD measurement, while for peptides containing six, seven or more amino acids ($n \geq 6$) a β-conformation was favored[46]. Here again the peak positions remained approximately at the same wavelengths, but the amplitudes of the two observed bands varied ten-fold. Another approach, in addition to that of theoretical calculations and information derived from model systems, is the deconvolution of the CD curves of globular proteins[9,10,47,48].

The CD curve of the CheY protein[49] is reported herein (Fig. 1) as an excellent example to underline the necessity of deconvoluting the measured spectra; the overall appearance can be very deceptive, leading to erroneous estimations of the secondary structures. One could assume from the overall shape of the CD curve that the conformation of this protein is mostly α-helical (see Fig. 1), if the low intensity of the bands is ignored. However, the analysis of the X-ray diffraction data of the crystallized protein reports only 41% helix[49]. In addition, 25% of the crystal conformation was assigned to the β-pleated sheet, which appears to be "hidden" in the measured CD curve. This contradiction will be discussed in the conclusion part of this paper, with the aid of the newly developed pure component curve set.

The results herein were obtained by using the deconvolution method called CCA[48]. This algorithm deduces the spectral contribution of the secondary structures *directly from the experimental CD curves* without using any other "background" information in the deconvolution procedure. As demonstrated earlier[48], the method was found to be sensitive only to the deconvoluted data set; therefore poorly represented conformers (low percentages) are expected to be less accurately deconvoluted. Representing within a triangle constraint surface, the location of the vertices of the triangle for the conformer (Fig. 2 represents the conformational distribution of the proteins in the data set on the triangle) would be poorly defined, resulting in less accuracy. Examination of the X-ray diffraction data of the proteins used (Table I) in the earlier analysis[48] revealed that the percentages of the β-pleated sheet were very low. Therefore highly β-containing water soluble proteins were added in the analysis herein in order to improve the most probable CD curve of the pure β-pleated sheet.

Abbreviations

CD	circular dichroism
CCA	Convex Constraint Analysis
R	Pearson Product Correlation Coefficient

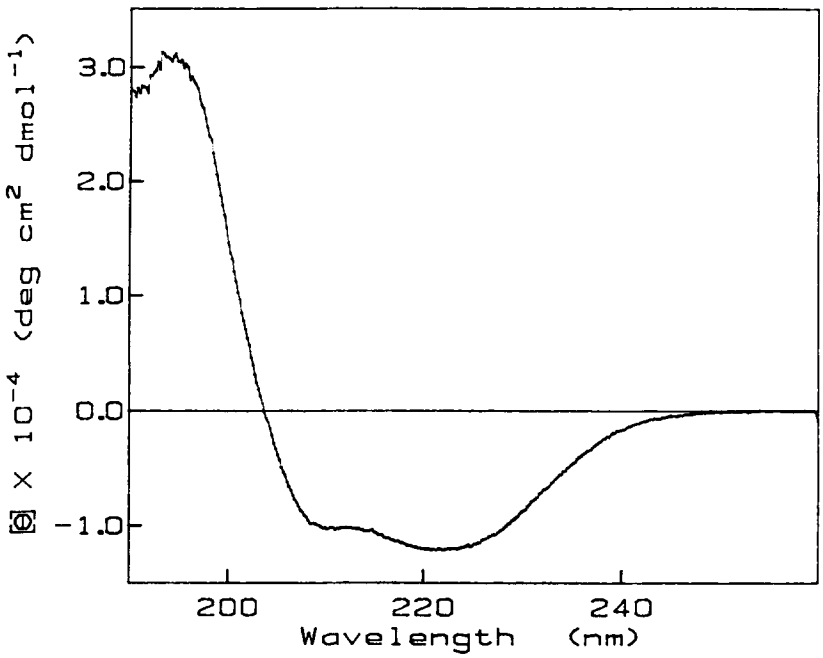

Fig. 1. The circular dichroism spectrum of the CheY protein. Conc. = 1.04 mg/mL. Path length = 0.01 cm. Temp. = 22°C.

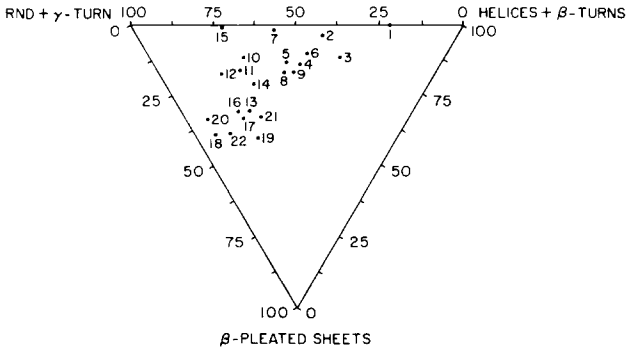

Fig. 2. The secondary structure "distribution" of the 22 proteins involved in the analysis. The three vertices of the triangle (simplex) are (RND + γ-turns), (α-helices + β-turns), and β-pleated sheets. The 3D structures of the proteins are mostly mixtures of these secondary structural elements. (Note that only one protein, the 1MBD is close to a vertex.) The simplex is a basic geometric element in a Euclidean space, being a line segment in one dimension, a triangle in two dimensions, a tetrahedron in three dimensions, and so on. The abbreviations are from Brookhaven Protein Data Bank, see Tables I and II. RND = random or unordered.

EXPERIMENTAL

Method of Analysis

Assuming the additivity of the chiral contributions of the different secondary structural elements such as α-helix, β-pleated sheet, etc., the circular dichroism (CD) spectra of a protein is therefore a function [f(λ)] which may be decomposed by a suitable method. If P is the number of the pure components and w_i is the weight (%) of the i-th pure component curve [$g_i(\lambda)$], the measured CD curve [$f^m(\lambda)$] can always be fitted by a calculated curve [$f^c(\lambda)$] which has the following form:

$$f^c(\lambda) = \sum_{i=1}^{P} w_{ij} \cdot g_i(\lambda) \,. \qquad (1)$$

A method solving such an equation has been developed earlier[44] and called the CCA method. This algorithm differs from the all alternative approaches[10,28] in that it uses only "natural constraints" (a, b, and c)[48] and yields adequate pure component curves and weights, respectively. The sum of the weights of the pure components for a protein must be equal to 1 (constraint a) with all weights being positive (constraint b). Proteins can

TABLE I. Secondary Structural Percentages of Proteins in the Standard Data Set as Determined by the Kabsch and Sander Method[56]*

Protein No.	X-ray Name†	% β-pleated sheet		% turns		% helix		% RND**
		PAR‡	ANT-P§	γ-	β-	α-	π	
1	1MBD	0.0	0.0	12.4	2.0	63.4	2.0	20.2
2	1CPV	0.0	3.7	17.6	2.8	38.0	0.9	37.0
3	3ADK	11.3	0.0	9.3	7.2	45.4	3.1	23.7
4	1INS	0.0	11.8	15.7	4.9	29.4	2.0	36.2
5	4LDH	7.3	5.8	12.2	8.8	27.1	1.8	37.0
6	1LZT	1.6	8.5	18.6	7.0	27.1	3.1	34.1
7	4CYT	0.0	1.9	9.7	5.8	32.0	1.0	49.6
8	5CPA	9.4	5.5	9.1	6.5	30.9	1.0	37.6
9	3TLN	3.8	10.8	8.5	6.6	33.5	2.2	34.6
10	2SBT	9.5	2.2	9.1	10.2	18.5	0.7	49.8
11	2PAP	1.4	14.6	5.2	6.1	19.2	0.9	52.6
12	5PTI	0.0	17.2	5.2	12.1	12.1	1.7	51.7
13	1RNS	1.6	29.0	6.5	9.7	13.7	0.8	38.7
14	2SNS	2.8	18.4	9.2	11.3	16.3	1.4	40.6
15	1RN3	1.6	2.9	8.9	4.0	12.9	1.6	68.1
16	4CHA	0.8	30.2	10.3	7.5	6.7	0.2	44.3
17	3EST	0.0	32.9	12.5	8.3	4.2	0.4	41.7
18	2CNA	0.0	39.2	3.8	10.1	2.1	0.0	44.8

*The classification of backbone units is based on the number of residues in the H-bonded unit[37].
†Abbreviations are from Brookhaven Protein Data Bank. 1. MBD, Myoglobin; 2. 1CPV, Calcium-binding Parvalbumin; 3, 3ADK, Adenylate Kinase; 4. 1INS, Insulin; 5. 4LDH, Lactate Dehydrogenase; 6. 1LZT, Lysozyme; 7. 4CYT, Cytochrome C; 8. 5CPA, Carboxypeptidase; 9. 3TLN, Thermolysin; 10. 2SBT, Subtilisin Novo; 11. 2PAP, Papain; 12. 5PTI, Trypsin Inhibitor; 13. 1RNS, Ribonuclease S; 14. 2SNS, Staphylococcal Nuclease; 15. 1RN3, Ribonuclease A; 16. 4CHA, Alpha-Chymotrypsin; 17. 3EST, Elastase; 18. 2CNA, Concanavalin A; 19. 2APP, Acid Proteinase; 20. 3FAB, Lambda Immunoglobulin FAB; 21. 1PSG, Pepsinogen; 22. 1GCR, Gamma-II Crystallin.
‡PAR, parallel β-pleated sheet.
§ANT-P, antiparallel β-pleated sheet.
**RND, unordered secondary structure.

TABLE II. Secondary Structural Percentages of Proteins From Beta Data Set as Determined by the Kabsch and Sander Method[56]*

Protein No.	X-ray Name†	% β-pleated sheet		% -turns		% helix		% RND**
		PAR‡	ANT-P§	γ-	β-	α-	π	
18	2CNA	0.0	39.2	10.1	3.8	2.1	0.0	44.8
19	2APP	5.9	34.7	12.1	11.1	6.5	0.6	29.1
20	3FAB	1.4	32.2	14.0	5.4	0.9	0.2	45.9
21	1PSG	5.8	27.1	11.2	12.3	8.8	1.6	33.2
22	1GCR	0.0	38.5	4.6	5.7	4.0	1.1	46.1

*The classification of backbone units is based on the number of residues in the H-bonded unit[37].
†Abbreviations are from Brookhaven Protein Data Bank. (see Table I).
‡PAR, parallel β-pleated sheet.
§ANT-P antiparallel β-pleated sheet.
**RND, unordered secondary structure.

be represented on a simplex, where the volume of the simplex (Fig. 2) should be minimal (constraint c) (for more details, see Ref. 48). An application is reported herein of the previously discussed method for N proteins, where the fitting error of the measured and calculated CD spectra was minimized;

$$\left\{\sum_{j=1}^{N} f_j^m(\lambda) - \sum_{j=1}^{N} f_j^c(\lambda)\right\}^2 = \left\{\sum_{j=1}^{N} f_j^m(\lambda) - \sum_{j=1}^{N}\sum_{i=1}^{P} w_{ij} \cdot g_i(\lambda)\right\}^2 \rightarrow \text{minimized.}$$

Data Bases

The analyses are based on a combined data set (standard + beta) where the CD curves of the proteins assembled in the standard data set (Table I) were reported by Yang et al.[37] The CD spectra of eighteen proteins were measured by Chen et al.[37] in the wavelength range 190–240 nm. The standard set alone was found to have insufficient data points near the "β-sheet peak" of the triangle representing the conformational distribution of the proteins (Fig. 2). The proteins included in the beta set (Table II) of the data base were therefore selected and, after their CD spectra were taken, added to the standard data set. By the addition of the beta set proteins to the standard data set, the beta sheet vertex became more stabilized (see Fig. 2). The overall data base contains the CD curves of 23 proteins in the 190–240 nm wavelength range.

Materials

Mouse immunoglobulin G (IgG) was kindly provided by Dr. Nisonoff, Department of Biology, Brandeis University. Penicillopepsin was supplied by Dr. T. Hofmann, Department of Biochemistry, University of Toronto, Canada; γ-II crystalline was obtained from Dr. D. Moss, Birkbeck College, University of London, London, England. CheY protein from *S. typhimurium* was a gift of Dr. A. Stock, Department of Biochemistry, Rosentiel Basic Medical Sciences Research Center, Brandeis University. Concanavalin A and porcine stomach pepsinogen were purchased from Sigma Chemical Co., St. Louis, MO (lot number for the proteins were 107F-8920 and 24F-8090, respectively). Other chemicals, unless otherwise stated, were reagent grade from Fisher Scientific Co. (Pittsburgh, PA).

Circular Dichroism (CD) Spectroscopy

All the CD spectral measurements were obtained using a Jovin-Yvon Mark V autodichrograph with sensitivity set at 2×10^{-6} to 5×10^{-6} as previously described[50]. Circular quartz cells with the path length of either 0.02 or 0.05 cm were used. The response time setting of the spectrometer was 2 sec with the data acquisition time set at 5 sec, which ensured good noise reduction as well as full-magnitude signal conversion from analog to digital. Each measurement was the average result of five repeated scans in steps of 0.2 nm at ambient temperature.

Protein concentration for IgG, in 0.14 M NaCl, 0.05 M sodium phosphate at pH 7.2, was 0.51 mg/mL, determined by UV absorbance at 280 nm ($\epsilon^{280}_{1mg/mL}$ = 1.2, 1 cm cell) after clarification by centrifugation in an Eppendorf microcentrifuge for 15 min (personal communication with Dr. Nisonoff). Similarly, the concentration of the γ-II crystalline, in 2.5 mM sodium phosphate at pH 7.4, was 0.34 mg/mL

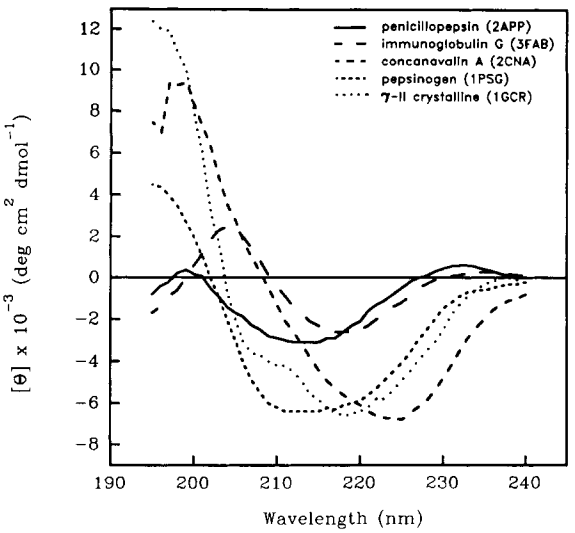

Fig. 3. The circular dichroism curves of the five proteins of high β-pleated sheet content. See experimental section for conditions.

($\epsilon^{280}_{1mg/mL}$ = 2.1, 1 cm cell)[51]. The concentrations were as follows: Concanavalin A[52] in 0.2 M NaCl, 0.02 M sodium acetate at pH 5.2 was 1.0 mg/mL; pepsinogen[53] in 2.5 mM sodium phosphate at pH 7.4, was 1.0 mg/mL and penicillopepsin[54] in 0.2 M NaCl, 0.02 M sodium acetate at pH 5.2[54] was 1.12 mg/mL, determined gravimetrically. The concentration of CheY protein (1.04 mg/mL) was determined using $\epsilon^{280}_{1mg/mL}$ = 0.5, 1 cm cell (personal communication from Dr. A. Stock).

RESULTS AND DISCUSSION

The secondary structural elements of proteins were chosen on the basis of X-ray diffraction crystallographic data. However, even on the basis of X-ray coordinates, the precise assignment of the secondary structural units can be ambiguous[48,55]. At least four different secondary structures are expected to occur in globular proteins: the α-helix, β-pleated sheet, and β- and γ-turns. The most accurate assignments yielded six or seven distinguishable units, using the typical back-bone torsion angles or the H-bond classifications. The Kabsch and Sander[56] method utilizes two types of helix (α- and π-) and two types of β-sheet (parallel and antiparallel), as well as the major turn types (β- and γ-). However in the 23 proteins reported herein the amount of π-helix and parallel β-pleated sheet were of low content (see Table I). Consequently the sum of the three major secondary structural elements (α-helix, antiparallel β-pleated sheet, and β-turns) with the percentage of the RND (unordered secondary structure) were generally 90% or larger. The CD spectra of proteins of high β-pleated sheet content were measured (Fig. 3) and a large variation in CD spectra was observed.

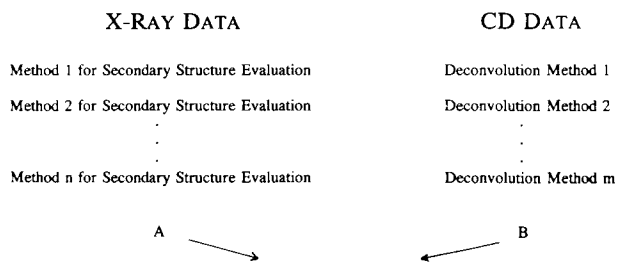

Fig. 4. The calculation of the Pearson Product Correlation Coefficient *(R)* is the final step whereby errors of the overall procedure are accumulated and expressed by a decrease in correlation coefficient. Accumulation of the errors can be remarkably high during the individual steps of the various procedures, such as measurements (X-ray diffraction or CD), and calculations of the secondary structural elements using a given method. All inaccuracies are summed and expressed by decreasing the theoretical values ($-1<R<1$) of the correlational coefficient.

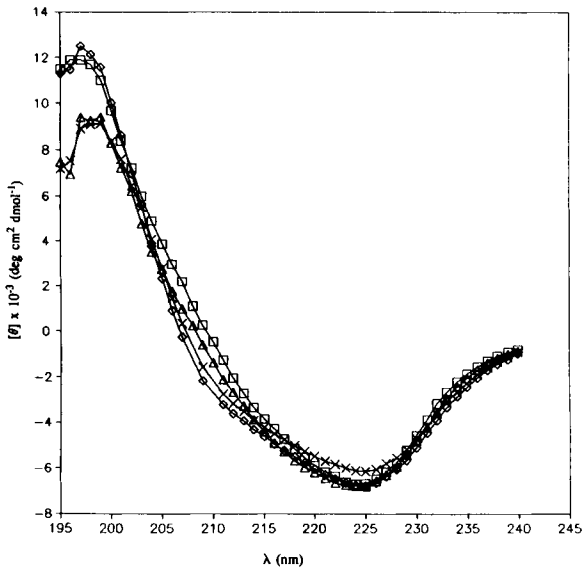

Fig. 5. The circular dichroism curves of CONCANAVALIN A measured in two different laboratories. The deviation between the two measured CD spectra is 1.27×10^{-3} (deg cm dmol^{-1}). Notice that each of the computed curves fit better to the measured, than the measured fit to each other. (-□-□-□-), reported by Yang et al.[37]; (-◇-◇-◇-), measured herein; (-x-x-x-), the fitted curve for experimental curve (-□-□-□-); (-△-△-△-), the fitted curve for experimental curve (-x-x-x-).

By calculating the dependence of the standard deviation (σ[P]) on the numbers of the pure component spectra (P) as described earlier[48], P was chosen as five in the analysis, in accordance with the results of the X-ray data analysis. Three parameters were taken into consideration as the final criteria for the above assignment of pure components: 1) The final accuracy, 2) the shape of the pure component curves, and 3) the correlation of the weights of the pure component curves with weights obtained from X-ray crystallography should be maximized. The average deviation of the measurement, $[\theta]_M^{meas.}$ from the calculated $[\theta]_M^{calc.}$, curve sets is expected to be in the range of $\pm 2 \times 10^3$ deg cm^2 dmol^{-1}. The correlational coefficient between the calculated weights (w_i) and the X-ray percentages are expected to be reasonably good; however one cannot expect values near 1. (The expected range of the correlation coefficient is discussed in the Error Analysis section.) The shape of the calculated pure component spectra must be similar to the theoretical and/or experimentally found pure component CD spectra (for details, see Ref. 8).

Error Analysis

Before the results of the analysis are discussed, an estimate of the errors must be made. A schematic representation of the sequential steps leading to the two types of secondary structural analysis is shown in Figure 4, illustrating two-separate procedures: 1)

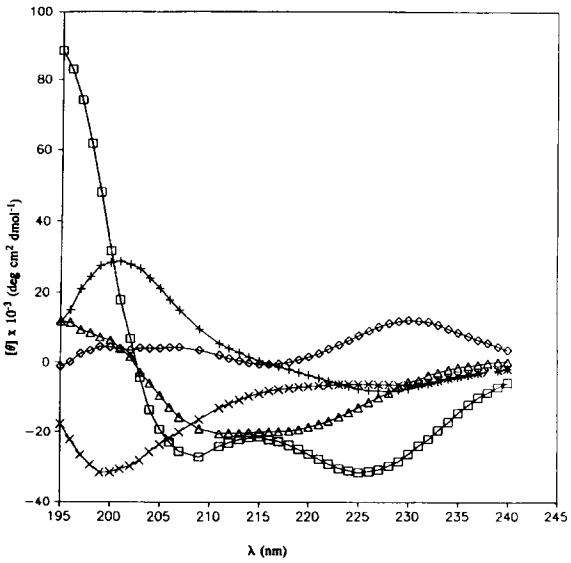

Fig. 6. The five pure component spectra from the deconvoluted 23 proteins at 46 wavelength values (standard + beta data set). The average for the final deviation for each data point was found to be ± 1.6 × 10^{-3} deg^2 cm^{-1} dmol.

$$(Dev)^2 = \sum_{i=1}^{23} \sum_{i=j}^{46} [f_{ij}^{comp.}(\lambda) - f_{ij}^{meas.}(\lambda)]^{2^a},$$

where $f_{ij}^{comp.}$ are the computed and $f_{ij}^{meas.}$ are the measured values of the CD curves. Correlation was established between: curve (1), -□-□-□-, and (α-helix + π-helix), curve (2), -△-△-△-, and (β-turns + parallel β-sheet), curve (3), -◇-◇-◇-, and [nonsecondary structure originated chiral contributions (aromatic and/or S-S)], curve (4), -x-x-x-, and (RND + γ-turn), and curve (5) (-+-+-+-) (antiparallel β-sheet). For correlation coefficients see Table III and Table IV. [a] deviation as defined in equation.

determination on the basis of X-ray diffraction and 2) determination on the basis of the CD measurement followed by a deconvolution method. The final step is the calculation of the correlation coefficient, whereby the accumulation of all errors are evaluated. The algorithm operates only on the measured CD data set. As CD is not a primary method, additional information is required for the assignment of the secondary structure. The inaccuracy of the measurements result in errors[48,55] which are added to the errors arising from the different definitions of the secondary structural elements in different methods. These errors will be accumulated and expressed by decreasing the final correlation coefficient. Therefore the theoretical values of the correlation coefficients ($-1<R<1$) are not attainable. One way to decrease the error of the overall procedure is to decrease the error of each step. The CCA reported earlier was scaled and found to produce sufficiently high correlation coefficients ($R>0.9$)[48] if the deconvolution was isolated from the previously coupled steps. This was accomplished by the deconvolution of the simulated CD curves of proteins. (These spectra were created as the linear combination of the pure component curves with different weights. The CCA resulted in correlation coefficients near 1 derived from the calculated weights and weights used as a linear combinational coefficient during the creation of the "artificial proteins" CD spectra.) The inaccuracies introduced by the CD measurements were minimized by using data obtained from only two instruments and using only well-characterized proteins. However, an "internal error-marker" was built into the analysis by involving the CD of Concanavalin A twice in the data set. One can estimate the accuracy of the CD measurements by checking the differences between the two curves (Fig. 5). (Note that the CCA treated the two curves as separate entities, which resulted in slightly different weights, underlining the higher accuracy of the deconvolution than the CD measurement.) The difference of the weights for the two separately analyzed Concanavalin A data sets is less than 5% and can be regarded as an internal signal for the estimation of the accuracy of the overall CD procedure (right side of Fig. 4).

Assignment and Discussion of the Pure Component CD Curves

Once the number of the pure components is a priori determined, the algorithm, using the a, b, and c constraints, operates only on the measured CD data set. The assignment of the conformational weights and related pure CD conformational curves is performed next, separately from the algorithm. For the

TABLE III. Secondary Structural Percentages of Proteins From (Standard + Beta) Data Set as Determined by the Kabsch and Sander[56] Method Using X-Ray Diffraction Data and by Convex Constraint Analysis Using Only the CD Curves of the Proteins

	X-ray Method[56] Secondary structural elements %						CD Method[48] Pure component curves[‡‡] %				
Name*	α-+ π-helix	β-turns + PAR[†]	RND[‡] + γ-turn	ATN-P[§]	S-S**	AROM[††]	1	2	4	5	3
1MBD	65.4	12.4	22.2	0.0	0	7	47.4	52.6	0.0	0.0	0.0
1CPV	38.9	17.6	39.8	3.7	0	9	22.6	32.1	20.6	18.8	5.9
3ADK	48.5	20.6	30.9	0.0	0	5	26.9	32.4	15.7	10.9	14.1
1INS	31.4	15.7	41.1	11.8	6	20	28.1	1.3	11.7	19.6	39.3
4LDH	28.9	19.5	45.8	5.8	0	6	24.3	22.8	23.8	19.2	9.9
1LZT	30.2	20.2	41.1	8.5	3	9	22.5	21.6	35.9	7.5	12.5
4CYT	33.0	9.7	55.4	1.9	0	10	12.9	22.2	23.2	30.2	11.5
5CPA	31.9	18.5	44.1	5.5	1	5	9.4	43.4	11.7	32.4	3.1
3TLN	35.7	12.3	41.2	10.8	0	13	28.4	27.1	16.5	2.2	25.8
2SBT	19.2	18.6	60.0	2.2	0	6	15.6	11.1	36.3	18.7	18.3
2PAD	20.1	6.6	58.7	14.6	1	13	0.3	13.0	51.7	31.7	3.3
5PTI	13.8	5.2	63.8	17.2	5	14	5.4	11.7	63.1	2.1	17.7
1RNS	14.5	8.1	48.4	29.0	3	7	7.7	28.0	23.9	12.3	28.1
2SNS	17.7	12.0	51.9	18.4	0	8	13.2	17.4	37.2	18.2	14.0
1RN3	14.5	10.5	72.1	2.9	0	7	7.4	26.9	25.7	11.1	28.9
4CHA	6.9	11.1	51.8	30.2	2	8	8.1	1.2	55.9	17.3	17.5
3EST	4.6	12.5	50.0	32.9	2	8	0.3	2.7	57.2	13.6	26.2
2CNA	2.1	3.8	54.9	39.2	0	9	11.8	1.4	22.3	42.9	21.6
2APP	7.1	17.0	41.2	34.7	0	11	5.5	5.9	32.9	19.3	36.4
3FAB	1.1	6.8	59.9	32.2	1	8	0.2	0.0	32.3	33.7	33.8
2CNA	2.1	3.8	54.9	39.2	0	9	7.1	3.2	25.3	45.6	18.8
1PSG	10.4	18.1	44.4	27.1	1	10	7.4	15.0	27.4	19.0	31.2
1GCR	5.1	5.7	50.7	38.5	0	16	15.6	14.6	18.6	18.3	32.9

*Abbreviations from the Brookhaven Protein Data Bank.
[†]PAR, parallel β-sheet.
[‡]RND, unordered secondary structure.
[§]ANT-P, antiparallel β-sheet.
**S-S, disulfides.
[††]AROM, aromatic amino acids.
[‡‡]CCA Component 1 correlated with α-helix and π-helix; CCA Component 2 correlated with β-turns and parallel β-sheet; CCA Component 4 correlated with unordered conformations and γ-turn; CCA Component 5 correlated with antiparallel β-sheet; CCA Component 3 correlated with S-S (disulfide bonds (%)) and AROM [aromatic amino acids (%)].

TABLE IV. Overall Pearson Product Correlation Coefficient *(R)* Between the Total Set of Secondary Structures, Determined by X-Ray Analysis[56] and by Convex Constraint Analysis[48]

		R
Complete Set	=	0.57
For pure comp 1	=	0.84 correlated with α-helix and π-helix
For pure comp 2	=	0.37 correlated with β-turns and parallel β-sheet
For pure comp 4	=	0.56 correlated with unordered conformations and γ-turns
For pure comp 5	=	0.41 correlated with antiparallel β-sheet
For pure comp 3	=	0.48 correlated with S-S [disulfide bonds (%)] and AROM [aromatic amino acids (%)]

different structural assignments one must use other information on the analyzed data set (X-ray determined secondary structural percentages, NMR determined secondary percentages, molecular dynamics calculated structures, etc.).

The shape of the calculated pure component curve 1 (Fig. 6) shows a great resemblance to those calculated for the α-helix back-bone geometry[8,50] and measured for highly α-helix containing proteins or polypeptides[42,57]. The calculated exiton couplet [190–210 nm] with the n→π* transition about 222 nm are in good agreement with literature data. Utilizing X-ray diffraction data, Chen et al.[38] performed a correlation between the chain lengths

TABLE V. The S-S Bond and Aromatic Residue Percentages of the Proteins From (Standard + Beta) Data Set as Determined by the Kabsch and Sander Method[56] and the Calculated Weights for Pure CD Component Curve (3), Figure 4

Name*	No. residues	S-S		Aromatic					Weights of component 3 (%)
		No.	%	F	W	Y	Total	%	
1MBD	153	0	0	6	2	3	11	7	0.0
1CPV	108	0	0	10	0	0	10	9	5.9
3ADK	194	0	0	4	0	6	10	5	14.1
1INS	102	6	6	9	0	11	20	20	39.3
4LDH	329	0	0	7	7	7	21	6	9.9
1LZT	129	4	3	3	6	3	12	9	12.5
4CYT	103	0	0	3	2	5	10	10	11.5
5CPA	307	1	1	16	7	19	42	5	3.1
3TLN	316	0	0	10	3	28	41	13	25.8
2SBT	275	0	0	3	3	10	16	6	18.3
2PAD	213	3	1	4	5	19	28	13	3.3
5PIT	58	3	5	4	0	4	8	14	17.7
1RNS	124	4	3	3	0	6	9	7	28.1
2SNS	141	0	0	3	1	7	11	8	14.0
1RN3	124	0	0	3	0	6	9	7	28.9
4CHA	477	10	2	12	16	8	36	8	17.5
3EST	240	4	2	3	6	11	20	8	26.2
2CNA	237	0	0	11	4	7	22	9	21.6
2APP	332	1	0	20	3	14	37	11	36.4
3FAB	428	5	1	11	7	17	35	8	33.8
2CNA	237	0	0	11	4	7	22	9	18.8
1PSG	365	3	1	16	4	17	37	10	31.2
1GCR	174	0	0	9	4	15	28	16	32.9

*Abbreviations from the Brookhaven Protein Data Bank. See Tables I and II for definitions.

(number of amino acid residues in the α-helix) and the intensities of the bands. The amplitude of the calculated curve (1) herein (Fig. 6) can be obtained from their calculation if the helical units are estimated to be in the range of 10–20 amino acids, which agrees with the average helix lengths of the helices in the proteins used in the data set, based on X-ray diffraction analysis. All the characteristics of this pure component curve agree well with literature data, as well as the high correlation ($R = 0.85$) between the calculated[56] and measured percentages.

The spectral curve (4) in Figure 6 shows a negative minimum near 200 nm, as was observed in the spectra of unordered polypeptides or proteins[43]. The correlation between X-ray based RND percentages, following Kabsch and Sander method[56], and the weight of this pure component curve was found to be significant. However, by adding the γ-turn percentages (I + 2, I units, following Kabsch and Sander definition) to the RND percentage, the significance increased (0.59, Tables III and IV). Therefore in this analysis the γ-turn conformation is treated as part of the unordered structure (RND). The unsuccessful determination of the pure CD spectrum of the γ-turn is mainly caused by the low contend (<10%) of the γ-form in the analyzed data set.

The spectral curve labeled (5) in Figure 6 was correlated with the antiparallel β-pleated sheet conformation. The maxima near 200 nm, with a negative band in the CD spectra near 225 nm, agrees with the general finding that antiparallel β-sheets, which occur more frequently than parallel β-sheets, produce a spectrum similar to this. According to theoretical considerations[58] the parallel β-form is expected to give a slightly different spectra. This is in agreement with the Hennessey and Johnson finding that the pure CD spectra of the parallel and antiparallel β-forms are different[47]. This statement seems to be supported by the fact that the percentages of the antiparallel form correlated better (0.41, Tables III and IV) than did the total of the β-form (parallel and antiparallel) with the coefficients of curve (5). The calculated weights for the high content β-proteins (see Fig. 7) show that there are more β-pleated sheets present than helices or turns (Tables III and IV).

The calculated curve (2), Figure 6, according to comparative CD data on linear[59] and bridged[59,60] model β-turn peptides, was identified as the CD curve of "global" β-turns. As demonstrated earlier, the Kabsch and Sander method[56] usually underestimates the percentage of the β-turns by their definition (not all β-turns have a 1←4 hydrogen bond) as required by their X-ray analysis method. Moreover, the synthetic model peptides supported the idea that the type I (or III) β-turns might give different spectra than the type II. According to this fact

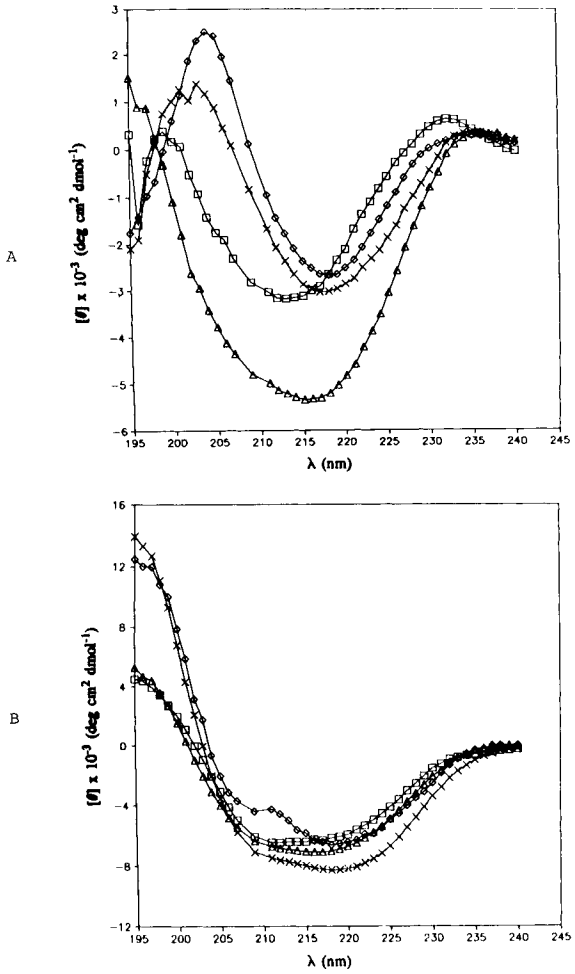

Fig. 7. The measured and the calculated pairs of circular dichroism curves [[θ] × 10^{-3} deg cm^{-1} $dmol^{-1}$] for Penicillopepsin and Pepsinogen. **A:** Penicillopepsin (2APP) and Immunoglobulin G (3FAB): PENIC, Penicillopepsin (-□-□-□-), measured; Calculated PENIC, Penicillopepsin (-△-△-△-), calculated; MIGG, Immunoglobulin G (-◇-◇-◇-) measured; Calculated MIGG, Immunoglobulin G (-x-x-x-) calculated. **B:** Pepsinogen (1PSG) and γ-II Crystalline (1GCR): PEPSI, Pepsinogen (-□-□-□-) measured; Calculated PEPSI, Pepsinogen (-△-△-△-) calculated; GIICR, γ-II Crystalline (-◇-◇-◇-) measured; Calculated GIIR, γ-II Crystalline (-x-x-x-) calculated.

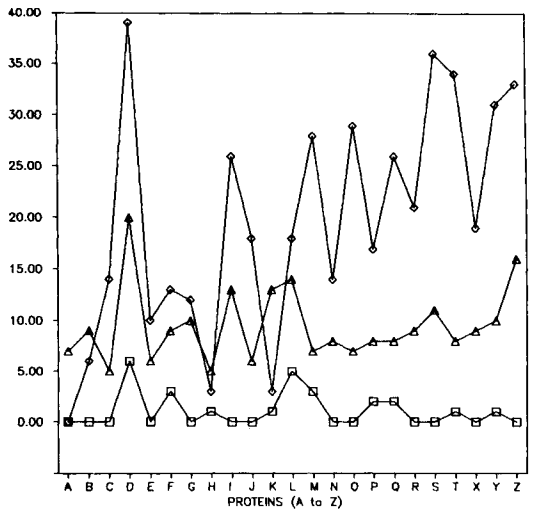

Fig. 8. The aromatic residues plus the disulfide percentages and weight calculated for the pure component curve (3), Figure VI. A, 1MBD; B, 1CPV; C, 3ADK; D, 1INS; E, 4LDH; F, 1LZT; G, 4CYT; H, 5CPA; I, 3TLN; J, 2SBT; K, 2PAD; L, 5PTI; M, 1RNS; N, 2SNS; O, 1RN3; P, 4CHA; Q, 3EST; R, 2CNA; S, 2APP; T, 3FAB; X, 2CNA; Y, 1PSG; Z, 1GCR. For abbreviations, see footnotes to Tables I and II.

the "global" representation of the β-turns (suggested by Kabsch and Sander[56]) will automatically decrease the correlation between the weights of the calculated and the X-ray determined percentages. However the correlation reached a significant level between weights of curve (2), Figure 6 (similar to a Class C CD spectra according to Woody's terminology[61]) and "global β-turns" percentages. This observation can be attributed to the finding that type I β-turns are considerably more frequent that type II in globular proteins[62]. (Usually a type I turn is expected to give a Class C CD spectra.) The correlation increased if the calculated weight for the parallel β-sheet was correlated with β-turns. However the rare occurrence of the parallel form of β-pleated-sheet (see Table I and II) makes it impractical to reach any further conclusion for the time being. Therefore, it would be implausible to conclude that the parallel β-sheet gives a class C like CD spectra[61].

The fusion of the seven typical secondary structural elements into four pure component groups (see Table III and text above) resulted in a significant correlation between weights of component curves and X-ray diffraction determined percentages by the Kabsch and Sander method[56]. Furthermore, the shape of the pure component CD curves agreed well with the commonly accepted criteria. However, the linear combinational coefficients (weights) of the third pure component curve (3), Figure 6, do not correlate well with any of the weights which were obtained by the Kabsch and Sander Method[56]. This can be accommodated by introducing other sources of CD contributions than those of secondary structural elements. The three amino acids (Phe, Trp, and Tyr), containing aromatic side chains, have a significant contribution to the CD spectra of proteins which originate from the L_a band of Tyr at 210 nm[63] or the B_b of indol[64]. Theoretical calculations demonstrated that the L_a transition can acquire rotational strength through coupling with the peptide transition. (This band can be red shifted to 225–230 nm in the case of tyrosine)[63]. Therefore the positive band near 230 nm, observed in the CD spectra of several proteins, can be attributed to the L_a band of tyrosine[65] or to the B_b band of indol in tryptophan[60]. On the other hand Hider et al.[66] demonstrated that

TABLE VI. Comparison of the Calculated Secondary Structure Percentages for the Five Highly β Proteins Using the Yang et al. Method[37] and the Convex Constraint Analysis Method[48]

	% β-Sheets			% β-Turns			% Helices			% Unordered conformation		
Protein*	KS†	Y‡	CCM§	KS†	Y‡	CCM§	KS†	Y‡	CCM§	KS†	Y‡	CCM§
2APP	41	65	33	17	15	6	7	0	6	41	20	19
3FAB	60	55	32	7	20	0	1	5	0	60	20	34
2CNA	55	00	25	4	45	3	2	35	7	55	20	46
1PSG	44	65	27	18	5	15	10	10	7	44	20	19
1GCR	50	75	19	6	0	15	5	15	15	51	10	18

*Abbreviations from the Brookhaven Protein Data Bank (see Table I and II for definition of the abbreviations).
†KS, Secondary structure of proteins using X-ray data as determined by the Kabsch and Sander method[37] (based on X-ray data).
‡Y, Secondary structure of proteins as determined by Yang et al. method[31] (based on CD curves).
§CCA, Secondary structure of proteins as determined by the Convex Constraint Analysis Method (based on CD curves)[48].

the disulfide chromophore is also capable of producing a positive band in the 220–230 nm region. Bolotina and Lugauskas[67] have previously attempted to include the contribution of aromatic amino acid residues into the CD spectra of proteins. The pure component curve (3), in Figure 6, obtained from the deconvolution, has a single maxima near 230 nm, and in accordance with the previous statements, this band can originate from the aromatic and/or disulfide chromophores. (Table V gives the detailed analysis of disulfide bonds and aromatic percentages found in proteins involved in the analysis.) The previous statement was based on the small difference between the individual correlation of aromatic percentages ($R_{AROM\rightarrow 3} = 0.47$) [$R_{AROM\rightarrow 3}$ is the Pearson product correlation coefficient between aromatic percentages and the weight calculated for the pure component curve (3)], and the summed correlation of both the S-S and aromatic weights ($R_{S\text{-}S+AROM\rightarrow 3} = 0.48$) [$R_{S\text{-}S+AROM\rightarrow 3}$ is the Pearson product correlation coefficient between the disulfide percentages and the weight calculated for the pure component curve (3)]. The similarity of the aromatic and the disulfide CD contributions in that both produce a positive band near 230 nm (see Fig. 8). (A weaker correlation was calculated between the disulfide percentages ($R_{S\text{-}S\rightarrow 3} = 0.33$) [$R_{S\text{-}S\rightarrow 3}$ is the Pearson product correlation coefficient between the disulfide percentages and the weight calculated for pure component curve (3)] and the weights of curve (3), Fig. 4.)

CONCLUSION

The CD curve of the pure antiparallel β-pleated sheet was obtained herein by deconvolution. The pure optical properties of this major secondary structural element were extracted for the first time directly from the CD spectra of globular proteins. The Pearson product correlation coefficient *(R)* between the weights of this β conformer in the investigated 23 proteins and the percentages derived from the X-ray data (using the Kabsch and Sander method[56]), was found to be *significant.*

For the proteins containing high β-sheet content, *a comparison was made between the CCA and the Yang method.* (The Yang calculation[37] is one of the best methods using simple basic functions for estimating the secondary structural percentages.) When the percentages obtained by the two fundamentally different methods were correlated with X-ray data (see Table VI) it can be seen that the use of the CCA method leads to an improvement over the Yang method[37]. Despite the accumulated errors in the CCA procedure, the accuracy is remarkably good when compared to the X-ray data (e.g., in 1PSG the difference (Δ) in secondary structures derived by the CCA method and the X-ray diffraction method is $\Delta_{\beta\text{-sheet}} = 17\%$, $\Delta_{\beta\text{-turns}} = 3\%$, $\Delta_{Helices} = 3\%$, and $\Delta_{RND} = 15\%$).

In accordance with earlier suggestions concerning the origin of the nonsecondary structural contribution (the aromatic and/or disulfide influence on the overall spectra in the range 195–240 nm), *a new pure component curve was obtained* by the use of the CCA method (curve 3, in Fig. 6), which has a single positive maxima in the wavelength range that is in accordance with the theoretical calculations. No previous method has predicted the presence of such a curve as well as the pure component curves of the major secondary structural elements.

The shapes of the CCA curves for the five pure components (α-helix, β-turns, antiparallel β-sheet, RND, and aromatic contribution) are in excellent agreement with literature expectations.

Upon applying the CCA method based on these results, the "unsolved question" reported in the introduction concerning the CheY protein can now be easily answered. The CheY protein used as an example (see Fig. 1) was successfully "reconstructed" using the pure component spectra reported herein (cf. Fig. 6). The calculated helical percentage is slightly underestimated [$\alpha\text{-helix}_{calculated}$ (%) = 31%, $\alpha\text{-helix}_{X\text{-ray}}$ (%) = 41%[49]], but the percentages found for the β-pleated sheets are in good agreement [$\beta\text{-sheet}_{calculated}$ (%) = 26%, $\beta\text{-sheets}_{X\text{-ray}}$ (%) = 25%[49]]. This example illustrates how a "hidden

conformer" can be found by the use of the CCA method.

ACKNOWLEDGMENTS

This research was supported in part by grants from NSF (DMB-8713193) and the U.S. Army Research Office (89-K-0088).

REFERENCES

1. Pauling, L., Corey, R.B., Branson, H.R. The structure of proteins: Two hydrogen-bonded helical configurations of the polypeptide chain. Proc. Natl. Acad. Sci. U.S.A. 37: 205–211, 1951.
2. Chirgadze, Yu.N. Deduction and systematic classification of spatial motifs of the antiparallel β structure in globular proteins. Acta Cryst. A43:405–417, 1987.
3. Pauling, L., Corey, R.B. Configurations of polypeptide chains favored orientations around single bonds: Two new pleated sheets. Proc. Natl. Acad. Sci. U.S.A. 37:729–740, 1951.
4. Chothia, C. Conformation of twisted β-pleated sheets in proteins. J. Mol. Biol. 75:295–302, 1973.
5. Chou, K.-C., Carlacci, L., Maggiora, G.G. Conformational and geometric properties of idealized β-barrels in proteins. J. Mol. Biol. 213:315–326, 1990.
6. Lasters, I., Wodak, S.J., Alard, P., van Cutsem, E. Structural principles of parallel β-barrels in proteins. Proc. Natl. Acad. Sci. U.S.A. 85:3338–3342, 1988.
7. Salemme, R.F. Structural properties of protein β-sheets. Prog. Biophys. Molec. Biol. 42:95–133, 1983.
8. Woody, R.W. Circular dichroism of peptides. In: "The Peptides," Vol. 7. Hruby, V.J., ed. New York: Academic Press, 1985:15–114.
9. Johnson, Jr., W.C. Protein secondary structure and circular dichroism: A practical guide. Proteins 7:205–214, 1990.
10. Provencher, S.W., Glockner, J. Estimation of globular protein secondary structure from circular dichroism. Biochemistry 20:33–37, 1981.
11. Shubin, U.V., Khazin, M.L., Efimovskaga, T.B. Prediction of secondary structure of globular proteins using CD spectra. Molek. Biol. (USSR) 24:189–201, 1990.
12. Manning, C.M., Illangesekare, M., and Woody, R.W. Circular dichroism studies of distorted α-helices, twisted β-sheets and β-turns. Biophys. Chem. 31:77–86, 1988.
13. Van Stokkum, H.M., Spoelder, H.J.W., Bloemendal, M., Van Grondelle, R., Groen, F.C.A. Estimation of protein secondary structure and error analysis from circular dichroism spectra. Anal. Biochem. 191:110–118, 1990.
14. Bobba, A., Cavatorta, P. Attimonelli, M., Riccio, P., Massotti, L., Quagliariello, E. Estimation of protein secondary structure from circular dichroism spectra: A critical examination of the CONTIN program. Prot. Seq. Data Anal. 3:7–10, 1990.
15. Kirkwood, J.G. On the theory of optical rotatory power. J. Chem. Phys 5:479–491, 1937.
16. Moffitt, W. Optical rotatory dispersion of helical polymers. J. Chem. Phys. 25:467–478, 1956.
17. Tinoco, Jr., I. Circular dichroism and rotatory dispersion curves for helixes. J. Am. Chem. Soc. 86:297–298, 1964.
18. Woody, R.W. Improved calculation of the nπ* rotational strength in polypeptides. J. Chem. Phys. 49:4797–4806, 1968.
19. Bayley, P.M., Nielsen, E.B., Schellman, J.A. The rotatory properties of molecules containing two peptide groups. J. Phys. Chem. 73:228–243, 1969.
20. Pysh, E.S. The calculated ultraviolet optical properties of polypeptide β-configurations. Proc. Natl. Acad. Sci. U.S.A. 56:825–832, 1966.
21. Rosenheck, K., Sommer, B. Theory of the far-ultraviolet spectrum of polypeptides in the β-conformation. J. Chem. Phys. 46:532–536, 1967.
22. Zubkov, V., Volkenshtein, M.W. Rotational strength of the nπ* transition in the β-form of polypeptides. Doklady Akad. Nauk. SSSR 175:492–498, 1967.
23. Volkenshtein, M.W., Subkov, V. Rotatory dispersion of the β-form of polypeptides. Biopolymers 5:465–472, 1967.
24. Woody, R.W. Optical properties of polypeptides in the β-conformation. Biopolymers 8:669–683, 1969.
25. Madison, V., Schellman, J.A. Optical activity of polypeptides and proteins. Biopolymers 11:1041–1076, 1972.
26. Weatherford, D.W., Salemme, F.R. Conformations of twisted parallel β-sheets and the origin of chirality in protein structure. Proc. Natl. Acad. Sci. U.S.A. 76:19–23, 1979.
27. Lifson, S., Sander, C. Specific recognition in the tertiary structure of β-sheets of proteins. J. Mol. Biol. 139:627–639, 1980.
28. Appelquist, J. Theoretical π-π* absorption and circular dichroic spectra of polypeptide β-structure. Biopolymers 21: 779–795, 1982.
29. Bazzi, M.D., Woody, R.W. Interaction of amphipathic properties with phospholipids: Characterization of conformation and the CD of oriented β-sheets. Biopolymers 26: 1115–1124, 1987.
30. Skolnick, J., Kolinski, A., Yaris, R. Dynamic Monte Carlo study of the folding of a six-stranded Greek key globular protein. Proc. Natl. Acad. Sci. U.S.A. 86:1229–1233, 1989.
31. Lasters, I. Estimating the twist of β-strands embedded within a regular parallel β-barrel structure. Prot. Eng. 4:133–135, 1990.
32. Chou, K.-C., Heckel, A., Némethy, G., Rumsey, S., Carlacci, L., Scheraga, H.A. Energetics of the structure and chain tilting of antiparallel β-barrels in proteins. Proteins 8:14–22, 1990.
33. Soman, K.V., Kamakrishnan, C. Identification and analysis of extended strands and β-sheets in globular proteins. Int. J. Biol. Macromol. 8:89–96, 1986.
34. Manning, M.C., Illangasekare, M., Woody, R.W. Circular dichroism studies of distorted α-helices, twisted β-sheets, and β-turns. Biophysical Chem. 31:77–86, 1988.
35. Manning, M.C., Woody, R.W. Theoretical determination of the CD of proteins containing closely packed antiparallel β-sheets. Biopolymers 26:1731–1732, 1987.
36. For review see Adler, A.J., Greenfield, N.J., Fasman, G.D. Circular dichroism and optical rotatory dispersion of proteins and polypeptides. Methods Enzymol. 27:675–735, 1973.
37. Yang, J.T., Wu, C.-S., Martinez, H.M. Calculation of protein conformation from circular dichroism. Methods Enzymol. 130:208–269, 1986.
38. Chen, Y.H., Yang, J.T., Chau, K.H. Determination of the helix and β form of proteins in aqueous solution. Biochemistry 13:3350–3359, 1974.
39. Dickerson, R., Geis, I. "The Structure and Action of Proteins." New York: Harper and Row, 1969.
40. Tinoco, Jr., I., Woody, R.W. Absorption and rotation of light by helical polymers: The effect of chain length. J. Chem. Phys. 38:1317–1325, 1963.
41. Sarkar, P.K., Doty, P. The optical rotatory properties of the β-configuration in polypeptides and proteins. Proc. Natl. Acad. Sci. U.S.A. 55:981–989, 1966.
42. Townsend, R., Kumosinski, T.F., Timasheff, S.N., Fasman, G.D., Davidson, B. The circular dichroism of the β-structure of poly-L-Lysine. Biochem. Biophys. Res. Com. 23: 163–169, 1966.
43. Stevens, L., Townend, R., Timasheff, S.N., Fasman, G.D., J. Potter. The circular dichroism of polypeptide films. Biochemistry 7:3717–3720, 1968.
44. Toniolo, C., Bonora, G.M. Structural aspects of small peptides: A circular dichroism study of monodisperse protected homo-oligomers derived from L-Alanine. Macromol. Chem. 176:2547–2558, 1975.
45. Balcerski, J.S., Pysh, E.S., Bonora, G.M., Toniolo, C. Vacuum ultraviolet circular dichroism of β-forming alkyl oligopeptides. J. Am. Chem. Soc. 98:3470–3473, 1976.
46. Toniolo, C., Bonora, G.M., Fontana, A. Three-dimensional architecture of monodisperse β-branched linear homo-oligopeptides. Int. J. Peptide Prot. Res. 6:371–380, 1974.
47. Hennessey, J.P., Johnson, Jr., W.C. Information content in the circular dichroism of proteins. Biochemistry 20:1085–1094, 1981.
48. Perczel, A., Hollósi, M., Tusnády, G., Fasman, G.D. Convex constraint analysis: A natural deconvolution of circular dichroism curves of proteins. Prot. Eng., 4:669–679, 1991.
49. Stock, A.M., Mottonen, J.M., Stock, J.B., Schutt, C.E.

Three-dimensional structure of CheY, the response regulator of bacterial chemotaxis. Nature 337:745–749, 1989.
50. Fasman, G.D., Park, K., Schlesinger, D.H. Conformational analysis of the immunodominant epitopes of the circumsporozoite protein (CS) of Plasmodium falciparum and knowlesi. Biopolymers 29:123–130, 1990.
51. Mandel, K., Bose, S.K., Chakrabarti, B., Siezen, R.J. Structure and stability of γ-crystallins. I. Spectroscopic evaluation of secondary and tertiary structure in solution. Biochim. Biophys. Acta 832:156–164, 1985.
52. Reeke, Jr., G.N., Becker, J.W., Edelman, G.M. Covalent and three-dimensional structure of Concanavalin A. Atomic coordinates, hydrogen bonding and quaternary structure. J. Mol. Biol. 250:1525–1547, 1975.
53. Perlman, G.E. The optical properties of pepsinogen. J. Mol. Biol. 6:452–464, 1961.
54. Sodek, J., Hoffmann, T. Large-scale preparation and some properties of Penicillopepsin. Can. J. Biochem. 48:425–431, 1970.
55. Bränden, C.-I., Jones, T.A. Between objectivity and subjectivity. Nature 343:687–689, 1990.
56. Kabsch, W., Sander, C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 22:2577–2637, 1983.
57. Greenfield, N., Fasman, G.D. Computed circular dichroism spectra for the evaluation of protein conformation. Biochemistry 8:4108–4116, 1969.
58. Kelly, M.M., Pysh, E.S., Bonora, G.M., Toniolo, C. Vacuum ultraviolet circular dichroism of protected homo-oligomers derived from Leucine. J. Am. Chem. Soc. 99:3264–3266, 1977.
59. Hollósi, M., Köver, K.E., Holly, S., Fasman, G.D. β-Turns in serine-containing linear and cyclic models. Biopolymers 26:1527–1533, 1987.
60. Hollósi, M., Köver, K.E., Holly, S., Radics, L., Fasman, G.D. β-Turns in bridged proline-containing cyclic models. Biopolymers 26:1555–1572.
61. Woody, R.W. Studies of theoretical circular dichroism of polypeptides: Contributions of β-turns. In "Peptides, Polypeptides and Proteins." Blout, E.R., Bovey, F.A., Goodman, M., Lotan, N., eds. New York: John Wiley and Sons, 1974: 338–350.
62. Chou, P.Y., Fasman, G.D. β-Turns in proteins. J. Mol. Biol. 115:135–175, 1977.
63. Woody, R.W. Aromatic side-chain contributions to the far ultraviolet circular dichroism of peptides and proteins. Biopolymers 17:1451–1467, 1978.
64. Auer, H.E. Far-ultraviolet absorption and circular dichroism spectra of L-Tryptophan and some derivatives. J. Am. Chem. Soc. 95:3003–3011, 1967.
65. Day, L.A. Circular dichroism and ultraviolet absorption of a deoxyribonucleic acid binding protein of filamentous bacteriophage. Biochemistry 12:5329–5339, 1973.
66. Hider, R.C., Kupryszewski, G., Rekowski, P., Lammek, B. Origin of the positive 225–230 nm circular dichroism band in proteins: Its application to conformational analysis. Biophys. Chem. 31:45–51, 1988.
67. Bolotina, I.A., Lugasuskas, Y.Yu. Molek. Biol. (USSR) 19: 1409–1421, 1985.