Protein Secondary Structures in Water from Second-Derivative Amide I Infrared Spectra†

Aichun Dong, Ping Huang, and Winslow S. Caughey*

Department of Biochemistry, Colorado State University, Fort Collins, Colorado 80523

Received July 11, 1989; Revised Manuscript Received November 8, 1989

ABSTRACT: Infrared spectra have been obtained for 12 globular proteins in aqueous solution at 20 °C. The proteins studied, which vary widely in the relative amounts of different secondary structures present, include myoglobin, hemoglobin, immunoglobulin G, concanavalin A, lysozyme, cytochrome *c*, α-chymotrypsin, trypsin, ribonuclease A, alcohol dehydrogenase, β_2-microglobulin, and human class I major histocompatibility complex antigen A2. Criteria for evaluating how successfully the spectra due to liquid and gaseous water are subtracted from the observed spectrum in the amide I region were developed. Comparisons of second-derivative amide I spectra with available crystal structure data provide both qualitative and quantitative support for assignments of infrared bands to secondary structures. Band frequency assignments assigned to α-helix, β-sheet, unordered, and turn structures are highly consistent among all proteins and agree closely with predictions from theory. α-Helix and unordered structures can each be assigned to only one band whereas multiple bands are associated with β-sheets and turns. These findings demonstrate a method of analysis of second-derivative amide I spectra whereby the frequencies of bands due to different secondary structures can be obtained. Furthermore, the band intensities obtained provide a useful method for estimating the relative amounts of different structures.

The pioneering work of Elliott and Ambrose (1950) gave promise of extensive use of infrared spectroscopy in the study of the substructures of peptides and proteins (Susi, 1972; Parker, 1971; Mendelsohn, 1984; Surewicz & Mantsch, 1988). Nine characteristic vibrational bands or group frequencies which arise from the amide groups of protein have been identified (Krimm & Bandekar, 1986). Among them, the amide I band, which is due almost entirely to the C–O stretch vibration of the peptide linkages that constitute the backbone structure, has been the most useful probe for determining the secondary structures of proteins in solution (Susi, 1972; Koenig & Tabb, 1980; Surewicz & Mantsch, 1988). Each type of secondary structures gives rise, in principle, to a different C–O stretch frequency in the amide I region of the spectrum (Miyazawa & Blout, 1961; Krimm & Bandekar, 1986). The basic theoretical calculations for these vibrations have been extensively examined by Miyazawa et al. (1956, 1961) and Krimm and co-workers (Krimm & Bandekar, 1986; Bandekar & Krimm, 1988). However, the low sensitivity of conventional dispersive infrared instruments and the difficulties encountered in achieving the accurate subtraction of the strong water absorption found in the amide I region severely limited the early experiments to solid state (Parker, 1983) or deuterium oxide solution (Susi & Byler, 1986) studies. The greater sensitivity of the FT-IR instrumentation now available makes it feasible to measure amide I spectra for proteins in water if a short (6–10 μm) path-length cell is used (Koenig & Tabb, 1980; Gorga et al., 1989). Furthermore, recent progress in the development of methods for analysis of spectral data makes it easier to distinguish the individual components within the intrinsically overlapped amide I band contours. Three analytic procedures in current use are Fourier self-deconvolution (enhancement), second derivative, and band curve-fitting (Susi & Byler, 1983, 1986; Yang et al., 1985). Nevertheless, the accurate measurement of the frequencies and intensities of the amide I bands has been elusive.

The potential usefulness of amide I spectra for the determination of types of secondary structure quantitatively, as well as qualitatively, of proteins and smaller polypeptides in aqueous solution makes the accurate measurement of such spectra of great interest. A continuing basic question in the study of polypeptide structure is the relationship between solution and crystal structures. The ability to measure the infrared spectra of crystals as well as solutions, intact tissues, or even a single cell (Dong et al., 1989) makes infrared methods particularly well-suited for comparing the structure of a protein in aqueous solution with its structure in a crystal (Potter et al., 1985; Gorga et al., 1989).

We report here the distribution of secondary structures determined from amide I spectra of aqueous solutions of globular proteins for which crystal structures are available. Effective criteria for subtraction of water absorption were developed. The resultant protein spectra upon second-derivative analysis yielded bands with frequencies characteristic of specific secondary structures that are essentially the same for all proteins. The second-derivative band areas (integrated intensities) gave relative amounts of different types of secondary structure for each protein that are nearly identical with the amounts computed from crystallographic data. We conclude that amide I infrared spectra provide a simple and reliable method for the determination of the secondary structures of proteins and other polypeptides in aqueous solution. A preliminary account of a portion of this study has been reported (Gorga et al., 1989).

Materials and Methods

Protein Sources. Human oxyhemoglobin A ($HbAO_2$)[1] and bovine heart oxymyoglobin (MbO_2) were isolated and purified as described previously (Caughey & Watkins, 1985; Shimada

† This work was supported by U.S. Public Health Service Grant HL-15980.

[1] Abbreviations: IR-SD, infrared second derivative; IgG, immunoglobulin G; Hb, hemoglobin; Mb, myoglobin.

0006-2960/90/0429-3303$02.50/0

& Caughey, 1982). Both proteins were converted to CO-liganded species by passing the CO gas over the top of solutions for 30 min. Final HbCO and MbCO solutions were concentrated to 8 and 5 mM in heme, respectively. Human class I major histocompatibility complex antigen A2 and purified human urinary β_2-microglobulin were obtained as described previously (Gorga et al., 1989).

The following proteins were purchased from Sigma and used without further purification: immunoglobin G (bovine, I-5506); concanavalin A (jack bean, C-2010); lysozyme (chicken egg white, L-6876); cytochrome *c* (horse heart, type VI, C-7752); α-chymotrypsin (bovine pancreas, C-7762); trypsin (bovine pancreas, T-8253); ribonuclease A (bovine pancreas, R-6000); and alcohol dehydrogenase (equine liver, A-6128). These proteins were chosen because their three-dimensional structures are known from X-ray studies (Levitt & Greer, 1977) and have been studied in D_2O solution by FT-IR (Susi & Byler, 1986) and in H_2O solution by circular dichroism (Provencher & Glockner, 1981; Hennessey & Johnson, 1981).

Preparation of Protein Solutions. HbCO and MbCO were in 10 mM sodium phosphate buffer, pH 7.4. The other proteins were at 5% (w/v) in 1% saline at pH 6.5.

Infrared Measurement and Data Manipulation. Protein solutions were prepared for infrared analysis in a Beckman FH-01 cell with CaF_2 windows and 6-μm path length. Infrared spectra were measured with a Perkin-Elmer Model 1800 Fourier transform infrared spectrophotometer at 20 °C. For each spectrum, a 1000-scan interferogram was collected at single beam mode with a 2 cm^{-1} resolution and a 1 cm^{-1} interval from the 4000 to 1000 cm^{-1} region. Reference spectra were recorded under identical conditions with only the media in which the protein was dissolved in the cells. The subtraction of the reference spectrum from the spectrum of the protein solution was carried out in accord with the criteria described below. The resultant protein difference spectra were smoothed with a nine-point Savitsky–Golay function (Savitsky & Golay, 1964) to remove the possible white noise. The Perkin-Elmer ENHANCE function, which is analogous to the method developed by Kauppinen et al. (1981a), using a half-bandwidth of 16 cm^{-1} and a *K* value of 2.3, was used to achieve spectral deconvolution. Second-derivative spectra were obtained with Savitsky–Golay derivative function software for a five data point window. The relative amounts of different protein secondary structures were determined from infrared second-derivative amide I spectra by manually computing the areas under the bands assigned to a particular substructure.

Results and Discussion

Criteria for Subtraction of the Water Spectrum. The strong water absorption in the protein amide I region makes it necessary to use a cell with a very short path length (about 6–10 μm) to get enough energy through the solution to permit measurement of the spectra of proteins in water solution. Such short path lengths severely limit the intensity of vibrational bands from protein amide groups and the signal-to-noise ratio at a given protein concentration. The short path lengths also make it more difficult to match the path lengths of sample and reference cells. Typically, the same cell was used to record spectra of both the sample and the reference medium; nevertheless, the simple drying and reloading steps apparently can alter the cell path length slightly. A particularly difficult problem associated with the accurate determination of protein infrared spectra in aqueous solution is the correct measurement of the spectrum in the amide I region that occurs due to water vapor in the radiation path. Water vapor bands are not easily compensated for by regular subtraction procedures and, furthermore, are easily enhanced by deconvolution and derivative processing (Surewicz & Mantsch, 1988). Artifactual bands and/or incorrect band positions appear in protein spectra if the subtraction of the water vapor spectrum is not done very accurately.

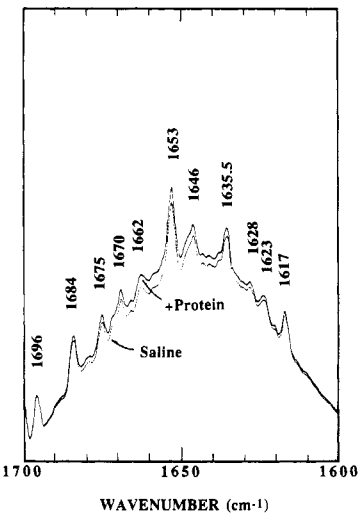

FIGURE 1: Infrared spectra in the amide I region of aqueous solutions of 1% saline with and without cytochrome *c* at 20 °C. The protein solution contained 5% (w/v) horse heart cytochrome *c* in 1% saline, pH 6.5. The ordinate represents absorbance as obtained directly with no manipulation. The inserted wavenumbers denote the bands from water vapor which are superimposed on a broad absorption from liquid water.

The nature of the water vapor problem is illustrated by Figure 1 which presents the infrared spectrum of a solution of cytochrome *c* in 1% saline over the 1700–1600 cm^{-1} region expressed in absorbance as measured without any data manipulation. The bands originating from water vapor are clearly seen even after extensive purging with dry air. The water vapor problem is less important in D_2O, which has much lower absorption in the amide I region than does H_2O, much longer path lengths can be used in D_2O than in H_2O to give greater intensity of the protein spectrum compared to the water vapor spectrum.

We find two criteria are especially critical to the successful subtraction of absorption bands due to liquid water and vaporized water in the atmosphere. To satisfy criterion I, the bands originating from water vapor must be accurately subtracted from the protein spectrum between the 1800 and 1500 cm^{-1} region regardless of the base line. To satisfy criterion II, a straight base line must be obtained from 2000 to 1750 cm^{-1}. The application of these two criteria is illustrated in Figure 2 wherein the infrared spectra of a solution of cytochrome *c* are shown for different approaches to the subtraction of the water bands. Figure 2A satisfies both criteria I and II; Figure 2B satisfies criterion II only whereas Figure 2C satisfies criterion I only. Whether or not the resultant second-derivative spectrum has bands from uncompensated water vapor, the success of water subtraction can be easily judged from an examination of the bands at 1684, 1670, 1662, 1653, 1646, and 1617 cm^{-1} which arise from water vapor. Many investigators have used a straight base line between 2000 and 1750 cm^{-1} as the standard by which to judge the success of water subtraction to obtain protein spectra (Surewicz et al., 1987; Olinger et al., 1986; Haris et al., 1986; Mitchell et al., 1988). It is clear from the remaining water vapor bands revealed by second-derivative analysis, as shown in Figure 2B, that the use of a straight base line from 2000 to 1750 cm^{-1} as the only standard is far from adequate since the bands due

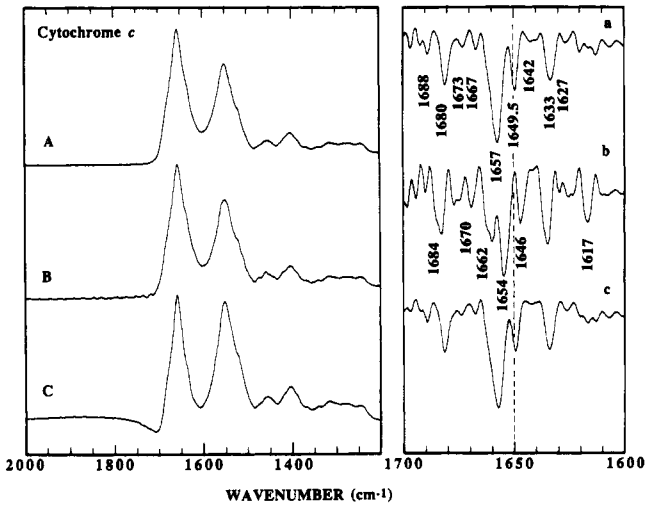

FIGURE 2: Effects of subtraction of spectra due to liquid water and water vapor from the infrared spectrum of cytochrome *c* in 1% aqueous saline. Conditions are as in Figure 1. (A) Difference spectrum resulting from subtractions that follow criteria I and II discussed in text. (B) Difference spectrum resulting from subtraction following criterion II but not criterion I. (C) Difference spectrum resulting from subtraction following criterion I but not criterion II. (a) Second derivative of (A) in the amide I region. (b) Second derivative of (B) in the amide I region. (c) Second derivative of (C) in the amide I region.

Table I: Deconvoluted Amide I Band Frequencies and Assignments for Proteins in Water

mean frequencies (cm^{-1})	assignment	mean frequencies (cm^{-1})	assignment
1624.0 ± 0.5	β-sheet	1650.0 ± 1.0	unordered
1627.0 ± 1.0	β-sheet	1656.0 ± 2.0	α-helix
1632.0 ± 1.0	β-sheet and extended chain	1666.0 ± 1.0	turn
		1672.0 ± 1.0	turn
1638.0 ± 1.0	β-sheet	1680.0 ± 1.0	turn
1642.0 ± 1.0	β-sheet	1688.0 ± 1.0	turn

to water vapor are evident. On the other hand, the second-derivative spectrum from Figure 2C is similar to Figure 2A. Thus, oversubtraction of the spectrum due to liquid water gives a nonlinear base line but does not perturb the wavenumbers of band maxima, if the water vapor spectrum has been correctly subtracted.

A straight base line from 2000 to 1750 cm^{-1} (criterion II) is required if the difference spectrum is to be used for band curve-fitting, e.g., for quantitative comparison of the protein amide I bands of different proteins. The amide I band shapes and areas can be affected by the choice of different base lines. When both criteria I and II were followed and the integral areas of amide I bands had been normalized, it proved possible to use the relative areas of second-derivative amide I bands to directly determine the relative amounts of the different types of secondary structure.

Assignments for Deconvoluted Amide I Bands. Deconvoluted amide I spectra for each protein were generated by both enhancement and second-derivative functions (Susi & Byler, 1983, 1986; Yang et al., 1985). More detailed information was obtained from the second-derivative analysis. Consequently, all band assignments in this study are based on second-derivative infrared spectra (Table I).

Assignments for α-Helical Structures. Figure 3 presents the infrared spectra of human HbACO and bovine MbCO. It has been generally agreed on the basis of X-ray crystallography that Hb is composed of 75–85% α-helix and 15–25% unordered structures and Mb has 70–80% α-helix and 20–30%

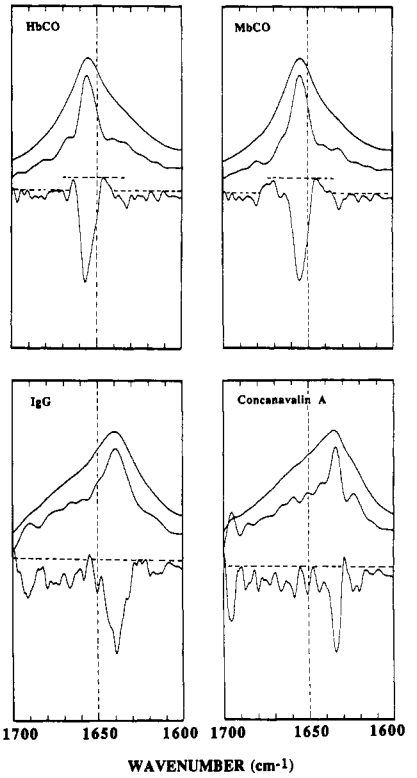

FIGURE 3: Infrared spectra in the amide I region of aqueous solutions of proteins. In each panel, the upper, middle, and lower spectra represent the difference spectrum from subtraction of water (liquid and vapor) spectra in accord with criteria I and II, the enhanced difference spectrum, and the second-derivatie of the difference spectrum, respectively. Hemoglobin carbonyl (8 mM in heme) and myoglobin carbonyl (5 mM) are in 10 mM sodium phosphate buffer, pH 7.4. Immunoglobin G and concanavalin A are each at a concentration of 5% (w/v) in 1% saline, pH 6.5. The horizontal dashed lines represent base lines used in computations of band intensities.

unordered structures (Richardson, 1981; Ladner et al., 1977; Levitt & Greer, 1977); no β-sheet structure was found in either protein. The original infrared spectra of both proteins show a relatively narrow amide I band maximum near 1655 cm^{-1}, which is expected for proteins with secondary structures of largely α-helical character (Keonig & Tabb, 1980; Susi & Byler, 1986; Krimm & Bandekar, 1986). By deconvoluting the original spectra, strong bands centered at 1656 cm^{-1} for Hb and 1654 cm^{-1} for Mb were revealed. On the basis of structural information on Hb and Mb from X-ray studies, these strong bands can be assigned to α-helical structures. These assignments are strongly supported by previous infrared studies (Susi & Byler, 1986; Mitchell et al., 1988; Holloway & Mantsch, 1989).

Assignments for β-Sheet Structure. Figure 3 also shows the spectra of immunoglobulin G (IgG) and concanavalin A. Being classified as all β-structure proteins by X-ray studies (Richardson, 1981; Levitt & Chothia, 1976), both IgG and concanavalin A have 60–75% β-sheet structure and almost no α-helix (only 2–5% α-helix found in both IgG and concanavalin α) (Levitt & Greer, 1977). The original infrared spectra show amide I band maxima near 1640 cm^{-1} for IgG and 1635 cm^{-1} for concanavalin A. Upon deconvolution, five bands at 1642, 1638, 1632, 1627, and 1624 cm^{-1}, respectively, can be assigned to β-sheet components. The second intense band at 1624 cm^{-1} of concanavalin A represents an unusual β-sheet structure among the proteins; the appearance of this band is rare, and it probably arises from a distorted β-strand as suggested by Arrondo et al. (1988). Concanavalin and β-lactoglobulin (Casal et al., 1988) are the only two proteins, so far, which

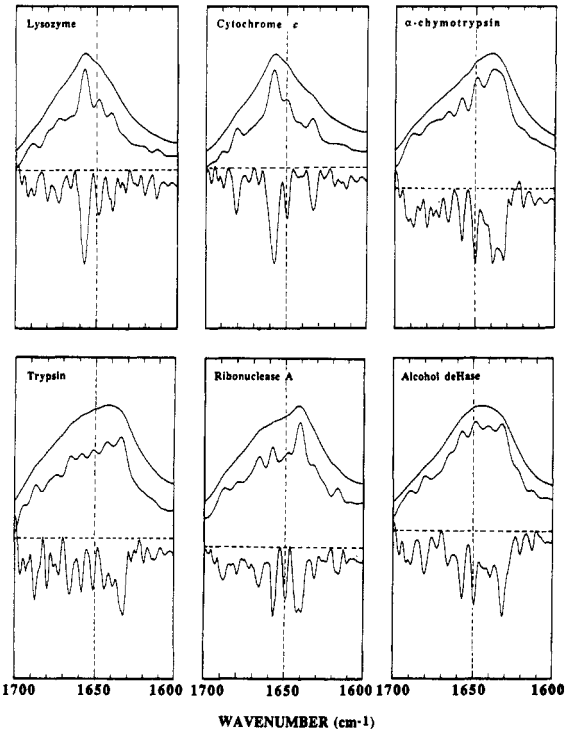

FIGURE 4: Infrared spectra in the amide I region of aqueous solutions of proteins. In each panel, the upper, middle, and lower spectra represent the difference spectrum from subtraction of water (liquid and vapor) spectra in accord with criteria I and II, the enhanced spectrum, and the second derivative spectrum, respectively. Protein concentrations are 5% (w/v) in 1% aqueous saline, pH 6.5. Alcohol deHase represents alcohol dehydrogenase.

have been found to contain this particular β-structure.

Assignments for Unordered Structure. The unordered component, usually referred to as "random coil", was defined in early X-ray crystallographic studies as a secondary structure of neither α-helix nor β-sheet. Its major distinguishing feature is that it is nonrepetitive in polypeptide backbone conformation. Due to the recognition of various types of β-turn conformation (Venkatachalam, 1968; Lewis et al., 1973), the unordered component has been referred to as a structure that is neither α-helix, nor β-sheet, nor turn (Richardson, 1981). As a consequence of this new definition, the percentage value of unordered structure in a given protein is much less than the value given in original X-ray crystallographic studies.

Current opinion has considered the unordered component to be indistinguishable from α-helix by infrared spectroscopy because of the small frequency difference between the amide I C–O stretch of these two structural components (Olinger et al., 1986; Casal et al., 1988; Surewicz & Mantsch, 1988). Others have assigned the band near 1645 cm^{-1} in D_2O solution to the unordered component (Byler & Suzi, 1986; Fry et al., 1988; Holloway & Mantsch, 1989).

Figure 4 shows spectra of lysozyme, cytochrome *c*, α-chymotrypsin, trypsin, ribonuclease A, and alcohol dehydrogenase. Each protein has been classified as an α + β protein, except for alcohol dehydrogenase which contains a parallel α/β domain (Richardson, 1981; Levitt & Chothia, 1976). X-ray crystallography and circular dichroism studies indicate the amounts of α-helix, β-sheet, unordered, and β-turn structures vary widely among these proteins (Levitt & Greer, 1977; Provencher & Glockner, 1981; Hennessey & Johnson, 1981). The spectra in Figure 4 are arranged in order from α-helix-predominant to β-sheet-predominant proteins. The assignments for secondary structures are given in Table I. The characteristic band of an unordered conformation in the amide I region can be assigned to the band located at 1650 ± 1.0 cm^{-1}. This band appears as an asymmetric shoulder on the α-helix band in the deconvoluted spectra of Hb and Mb, but it is clearly separated from the α-helix band in the spectra of low helical content proteins and β-sheet-predominant proteins, due to the intrinsic difference between the amide C–O vibration of these two components.

Assignments for β-Turn Structures. Various turn conformations are additional common secondary structures that generally exist in globular proteins. It has been observed that about 32% of the residues in 29 globular proteins reside in β-turn conformations (Chou & Fasman, 1977). Since α-helix, unordered, and β-sheet can be assigned to other bands, the bands located at 1688, 1680, 1672, and 1666 cm^{-1} are reasonably assigned to various types of β-turn structures. The frequencies for these turn structures are generally similar to values reported by other investigators (Byler & Susi, 1986, 1988; Surewicz et al., 1987; Arrondo et al., 1988), but our frequency values are less variable. Among the band assignments in Table I, there is excellent reproducibility in frequency for the bands among different proteins which indicates that the vibrational force field of the amide C–O stretch remains very consistent for each type of secondary structure. It should also be noted that bands assigned to β-sheet in the β-proteins located near 1638, 1632, and 1627 cm^{-1} also exist in Hb and Mb, in which no β-sheet conformation has been found in X-ray studies (Ladner et al., 1977; Kendrew et al., 1960). Similar bands have also been observed in an infrared study of Hb and Mb in D_2O and appear to be due to short irregular pieces of extended structure similar to β-sheet (Byler & Susi, 1986).

The second-derivative analysis procedure provides detailed information of the protein β-sheet structure. As shown in Figures 3 and 4, the major β-sheet structures in concanavalin A, trypsin, and alcohol dehydrogenase are characterized by a band near 1632 cm^{-1}, whereas the major β-sheet structures in IgG and lysozyme are represented by the band at 1638 cm^{-1}. A structural difference between trypsin and α-chymotrypsin is clearly indicated by a less intense β-structure band at 1638 cm^{-1} in trypsin. The difference of β-sheet composition among the proteins is potentially useful in secondary structure identification. A band located near 1675 cm^{-1}, which has been assigned to a "high component" β-structure (Byler & Susi, 1986; Olinger et al., 1986; Casal et al., 1988), was also observed in this study, but the contribution of this band to the total β-sheet structures is small.

Quantitation of Protein Secondary Structures. The second-derivative analysis of infrared spectra permits direct quantitative analysis of the secondary structural components of proteins in water solution. The intrinsic shape of an infrared absorbance is approximated by a Lorentzian function (Surewicz & Mantsch, 1988). In the second-derivative spectrum, the peak frequency of an absorbance is identical with the original peak frequency, but the half-bandwidth is reduced by a factor of 2.7 (Susi & Byler, 1983, 1986). The height of a second-derivative peak is proportional to the square of the original peak height with an opposite sign, and the half-bandwidth is inversely proportional to the square of the original half-bandwidth (Susi & Byler, 1983, 1986). The areas corresponding to the different types of secondary structure are illustrated with the cytochrome *c* spectrum shown in Figure 5. The close correlation between the relative areas of assigned amide I bands in second-derivative spectra and known three-dimensional structures discussed below supports the use of the second-derivative infrared spectrum for the quantitative as well as qualitative evaluation of protein secondary structure.

Table II: Comparison of Protein Secondary Structures Determined by IR-SD and CD Spectra and X-ray Crystallography

protein	secondary structure (%) α-helix	β-sheet	turn	random	method	ref
hemoglobin	78[a]	12	10	*a*	IR-SD	
	87	0	7	6	X-ray	*b*
	68–75	1–4	15–20	9–16	CD	*c, g*
myoglobin	85[a]	7	8	*a*	IR-SD	
	85	0	8	7	X-ray	*b*
	67–86	0–13	0–6	11–30	CD	*c–e, g–i*
lysozyme	40	19	27	14	IR-SD	
	45	19	23	13	X-ray	*b*
	29–45	11–39	8–26	8–60	CD	*c–i*
cytochrome *c* (oxidized)	42	21	25	12	IR-SD	
	48	10	17	25	X-ray	*b*
	27–46	0–9	15–28	28–41	CD	*c–g, i*
α-chymotrypsin	9	47	30	14	IR-SD	
	8	50	27	15	X-ray	*b*
	8–15	10–53	2–22	38–70	CD	*c–e, g–i*
trypsin	9	44	38	9	IR-SD	
	9	56	24	11	X-ray	*b*
ribonuclease A	15	40	36	9	IR-SD	
	23	46	21	10	X-ray	*b*
	12–30	21–44	11–22	19–50	CD	*c–e, g–i*
alcohol dehydrogenase	18	45	23	14	IR-SD	
	29	40	19	12	X-ray	*b*
concanavalin A	8	58	26	8	IR-SD	
	3	60	22	15	X-ray	*b*
	3–25	41–49	15–27	9–36	CD	*d–f*
immunoglobulin G	3	64	28	5	IR-SD	
	3	67	18	12	X-ray	*b*
major histocompatibility complex antigen A2	17	41	28	14	IR-SD	*j*
	20	42			X-ray	*k*
	8–13	74–77			CD	*j*
β_2-microglobulin	6	52	33	9	IR-SD	*j*
	0	48			X-ray	*k*
	0	59			CD	*j*

[a] The band due to random structure appears as a shoulder on the α-helix band and is too small to be separated from α-helix structure; the random structure is estimated at <5% and is included in the α-helix value. [b] Levitt & Greer (1977). [c] Hennessey & Johnson (1981). [d] Provencher & Glockner (1981). [e] Chang et al. (1978). [f] Brahms & Brahms (1980). [g] Manavalan & Johnson (1987). [h] Greenfield & Fasman (1969). [i] Chen et al. (1972). [j] Gorga et al. (1989). [k] Bjorkman et al. (1987).

The accuracy of measured band areas in amide I second-derivative spectra is dependent upon the correct positioning of the base line. The dashed horizontal lines on second-derivative spectra of Figures 3 and 4 represent the base lines used to compute the relative amounts of each kind of secondary structure given in Table II. Choice of base-line location can be reliably made if it is recognized that on occasion an abnormal excursion above the true base line can occur as in IgG, concanavalin A, and chymotrypsin second-derivative spectra. Fortunately, small deviations in base-line position above or below the true base line only subtly affect the calculated relative percentages of major secondary structures. An appropriate *single base line* cannot be drawn for spectra of proteins with extremely high α-helix contents, e.g., Hb and Mb. The strong α-helix band results in unusually strong positive lobes on each side of the α-helix band (Martin, 1959; Kauppinen et al., 1981b). However, a *two-base-line* approach as shown in the spectra for HbCO and MbCO in Figure 3 permits an approximate distribution among secondary structures to be determined.

The relative amounts of α-helix, β-sheet, turns, and random secondary structures of 12 proteins estimated for aqueous solutions by analysis of the infrared second-derivative spectra (IR-SD) are compared in Table II with relative amounts reported in circular dichroism and X-ray crystallography studies. A close correlation of the amounts of IR-SD-determined secondary structures with those obtained by X-ray crystallography is found.

Conclusions. Methods of processing infrared spectra of 12 proteins in aqueous solution have been developed which permit the measurement of amide I bands in the presence of liquid and gaseous water. A remarkable consistency in amide I stretch frequency for each type of secondary structure is demonstrated. The relative amounts of different secondary

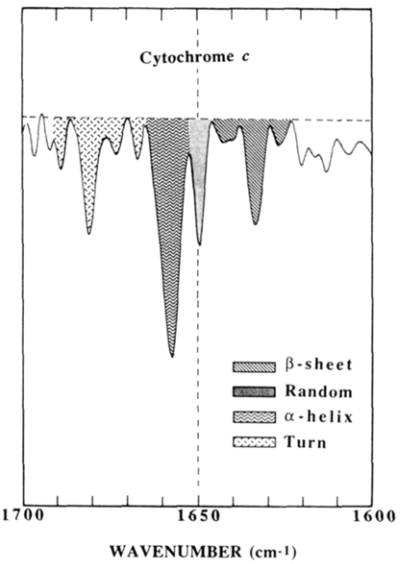

FIGURE 5: Illustration of the band areas assigned to different secondary structures for cytochrome *c*. No attempt is made in this illustration to deconvolute each band completely; e.g., the relative contributions of the α-helix and unordered bands in the area where they overlap are not considered accurately. Also, no attempt is made to determine the areas for each of the β-sheet and turn bands; i.e., only the combined areas of β-sheet bands and of turn bands are measured.

structure as determined from amide I second-derivative band areas agree closely with the amounts determined for crystals by X-ray crystallography. These findings provide support for the utility of the reported IR-SD method for the qualitative and quantitative analysis of protein secondary structure in solution. Since there is presently no experimental method available which permits an absolutely accurate determination of protein secondary structure in solution, it is not possible to assess in absolute terms the accuracy of the amounts we estimate to be present based on second-derivative amide I infrared spectra. However, it is likely that the estimated amounts for major structures are accurate to within ±5% of the value given except for when the α-helix content is greater than 60%, in which case the accuracy is expected to be somewhat lower.

Registry No. Concanavalin A, 11028-71-0; lysozyme, 9001-63-2; cytochrome *c*, 9007-43-6; α-chymotrypsin, 9004-07-3; trypsin, 9002-07-7; ribonuclease A, 9001-99-4; dehydrogenase, 9031-72-5.

References

Arrondo, J. L. R., Young, N. M., & Mantsch, H. H. (1988) *Biochim. Biophys. Acta 952*, 261–268.

Bandekar, J., & Krimm, S. (1988) *Biopolymers 27*, 909–921.

Bjorkman, P. J., Saper, M. A., Samraoui, B., Bennett, W. S., Strominger, J. L., & Wiley, D. C. (1987) *Nature 329*, 506–512.

Brahms, S., & Brahms, J. (1980) *J. Mol. Biol. 138*, 149–178.

Byler, D. M., & Susi, H. (1986) *Biopolymers 25*, 469–487.

Byler, D. M., & Susi, H. (1988) *J. Ind. Microbiol. 3*, 73–88.

Casal, H. L., Kohler, U., & Mantsch, H. H. (1988) *Biochim. Biophys. Acta 957*, 11–20.

Caughey, W. S., & Watkins, J. A. (1985) in *Handbook of Methods for OxyRadical Research* (Greenwald, R., Ed.) pp 95–104, CRC Press, Boca Raton, FL.

Chang, C.-T., Wu, C.-S. C., & Tang, J.-T. (1978) *Anal. Biochem. 91*, 13–31.

Chen, Y.-H., Tang, J.-S., & Martinez, H. M. (1972) *Biochemistry 11*, 4120–4131.

Chou, P. Y., & Fasman, G. D. (1977) *J. Mol. Biol. 115*, 135–175.

Dong, A., Messerschmidt, R. G., Reffner, J. A., & Caughey, W. S. (1988) *Biochem. Biophys. Res. Commun. 156*, 752–756.

Elliott, A., & Ambrose, E. J. (1950) *Nature 165*, 921–922.

Fry, D. C., Byler, D. M., Susi, H., Brown, E. M., Kuby, S. A., & Mildvan, A. S. (1988) *Biochemistry 27*, 3588–3598.

Gorga, J. C., Dong, A., Manning, M. C., Woody, R. W., Caughey, W. S., & Strominger, J. L. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 2321–2325.

Greenfield, N., & Fasman, G. D. (1969) *Biochemistry 8*, 4108–4116.

Haris, P. I., Lee, D. C., & Chapman, D. (1986) *Biochim. Biophys. Acta 874*, 255–265.

Hennessey, J. P., Jr., & Johnson, C., Jr. (1981) *Biochemistry 20*, 1085–1094.

Holloway, P. W., & Mantsch, H. H. (1989) *Biochemistry 28*, 931–935.

Kauppinen, J. K., Moffat, D. J., Mantsch, H. H., & Cameron, D. G. (1981a) *Appl. Spectrosc. 35*, 271–276.

Kauppinen, J. K., Moffatt, D. J., Mantsch, H. H., & Cameron, D. (1981b) *Anal. Chem. 65*, 1454–1457.

Kendrew, J. C., Dickerson, R. E., Strandberg, B. E., Hart, R. G., Davies, D. R., Phillips, D. C., & Shore, V. C. (1960) *Nature 185*, 422–427.

Koenig, J. K., & Tabb, D. L. (1980) in *Analytical Applications of FT-IR to Molecular and Biological Systems* (Durig, J. R., Ed.) pp 241–255, D. Reidel, Boston.

Krimm, S., & Bandekar, J. (1986) *Adv. Protein Chem. 38*, 181–365.

Ladner, R. C., Heidner, E. J., & Perutz, M. F. (1977) *J. Mol. Biol. 114*, 385–414.

Levitt, M., & Chothia, C. (1976) *Nature 261*, 552–558.

Levitt, M., & Greer, J. (1977) *J. Mol. Biol. 114*, 181–239.

Lewis, P. N., Momany, F. A., & Scheraga, H. A. (1973) *Biochim. Biophys. Acta 303*, 211–229.

Manavalan, P., & Johnson, W. C., Jr. (1987) *Anal. Biochem. 167*, 76–85.

Martin, A. E. (1959) *Spectrochim. Acta 14*, 97–103.

Mendelsohn, R. (1984) *Tech. Life Sci.: Biochem. B109/1*, 1–28.

Mitchell, R. C., Haris, P. I., Fallowfield, C., Keeling, D. J., & Chapman, D. (1988) *Biochim. Biophys. Acta 941*, 31–38.

Miyazawa, T., & Blout, E. R. (1961) *J. Am. Chem. Soc. 83*, 712–719.

Miyazawa, T., Shimanouchi, T., & Mizushima, S. I. (1956) *J. Chem. Phys. 24*, 408–418.

Olinger, J. M., Hill, D. M., Jakobsen, R. J., & Brody, R. S. (1986) *Biochim. Biophys. Acta 869*, 89–98.

Parker, F. S. (1971) in *Applications of Infrared Spectroscopy in Biochemistry, Biology and Medicine*, pp 232–270, Plenum, New York.

Parker, F. S. (1983) *Applications of Infrared, Raman and Resonance Raman Spectroscopy in Biochemistry*, Plenum Press, New York.

Potter, W. T., Houtchens, R. A., & Caughey, W. S. (1985) *J. Am. Chem. Soc. 107*, 3350–3352.

Provencher, S., & Glockner, J. (1981) *Biochemistry 20*, 33–37.

Richardson, J. S. (1981) *Adv. Protein Chem. 34*, 167–339.

Savitzky, A., & Golay, J. E. (1964) *Anal. Chem. 36*, 1628–1639.

Shimada, H., & Caughey, W. S. (1982) *J. Biol. Chem. 258*, 11893–11900.

Surewicz, W. K., & Mantsch, H. H. (1988) *Biochim. Biophys. Acta 952*, 115–130.

Surewicz, W. K., Szabo, A. G., & Mantsch, H. H. (1987) *Eur. J. Biochem. 167*, 519–523.

Susi, H. (1972) *Methods Enzymol. 26*, 445–472.

Susi, H., & Byler, D. M. (1983) *Biochem. Biophys. Res. Commun. 115*, 391–397.

Susi, H., & Byler, D. M. (1986) *Methods Enzymol. 130*, 290–311.

Venkatachalam, C. M. (1968) *Biopolymers 6*, 1425–1436.

Yang, W. J., Griffiths, D., Byler, D. M., & Susi, H. (1985) *Appl. Spectrosc. 39*, 282–287.