Baleja, J. D., Pon, R. T., & Sykes, B. D. (1990) *Biochemistry 29*, 4828-4839.
Borgias, B. A., & James, T. L. (1988) *J. Magn. Reson. 79*, 493-512.
Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983) *J. Comput. Chem. 4*, 187-217.
Brünger, A. T. (1990) X-PLOR *version 2.1*, user manual, Yale University, New Haven, CT.
Brünger, A. T., Clore, G. M., Gronenborn, A. M., & Karplus, M. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3801-3805.
Clore, G. M., Brünger, A. T., Karplus, M., & Gronenborn, A. M. (1986a) *J. Mol. Biol. 191*, 523-551.
Clore, G. M., Nilges, M., Sukumaran, D. K., Brünger, A. T., Karplus, M., & Gronenborn, A. M. (1986b) *EMBO J. 5*, 2729-2735.
Grollman, A. P. (1989) *Proc. Am. Assoc. Cancer Res. 30*, 682.
Gronenborn, A. M., & Clore, G. M. (1989) *Biochemistry 28*, 5978-5984.
Kalk, A., & Berendsen, J. C. (1976) *J. Magn. Reson. 24*, 343-366.
Keepers, J. W., & James, T. L. (1984) *J. Magn. Reson. 57*, 404-426.
Kouchakdjian, M., Marinelli, E., Gao, X., Johnson, F., Grollman, A., & Patel, D. J. (1989) *Biochemistry 28*, 5647-5657.
Kouchakdjian, M., Eisenberg, M., Live, D., Marinelli, E., Grollman, A., & Patel, D. J. (1990) *Biochemistry 29*, 4456-4465.
Lefevre, J. F., Lane, A. N., & Jardetzky, O. (1987) *Biochemistry 26*, 5076-5090.
Marion, D., Genest, M., & Ptak, M. (1987) *Biophys. Chem. 28*, 235-244.
Nilges, M., Clore, G. M., & Gronenborn, A. M. (1987a) *Biochemistry 26*, 3718-3733.
Nilges, M., Clore, G. M., & Gronenborn, A. M. (1987b) *Biochemistry 26*, 3734-3744.
Nilges, M., Habazettl, J., Brünger, A. T., & Holak, T. A. (1991) *J. Mol. Biol. 219*, 499-510.
Nilsson, L., & Karplus, M. (1986) *J. Comput. Chem. 7*, 591-616.
Nilsson, L., Clore, G. M., Gronenborn, A. M., Brünger, A. T., & Karplus, M. (1986) *J. Mol. Biol. 188*, 455-476.
Powell, M. J. D. (1977) *Math. Program. 12*, 241-254.
Ravishanker, R., Swaminathan, S., Beveridge, R. L., & Sklenar, H. (1989) *J. Biomol. Struct. Dyn. 6*, 669-699.
Ryckaert, J. P., Cicotti, G., & Berendsen, H. J. C. (1977) *J. Comput. Phys. 23*, 327-337.
Singer, B., & Grunberger, D. (1983) in *Molecular Biology of Mutagens and Carcinogenesis*, Plenum Press, New York.
Singer, B., & Bartsch, H. (1986) *The Role of Cyclic Nucleic Acid Adducts in Carcinogenesis and Mutagenesis*, IARC Scientific Publications 70, International Agency for Research on Cancer, Lyon, France.
Stout, G. H., & Jensen, L. H. (1989) *X-ray Structure Determination*, Wiley, New York.
Verlet, L. (1967) *Phys. Rev. 159*, 98-105.
Yip, P., & Case, D. A. (1989) *J. Magn. Reson. 83*, 643-648.

Secondary Structure of Streptokinase in Aqueous Solution: A Fourier Transform Infrared Spectroscopic Study

Heinz Fabian,*,‡ Dieter Naumann,§ Rolf Misselwitz,‡ Otto Ristau,‡ Dieter Gerlach,‖ and Heinz Welfle‡

Max Delbrück Center for Molecular Medicine, Robert-Rössle-Strasse 10, D-1115 Berlin-Buch, FRG, Robert Koch Institute of the Federal Health Office of Germany, Nordufer 20, D-1000 Berlin 65, FRG, and Institute of Experimental Microbiology, Beutenbergstrasse 11, D-6900 Jena, FRG

Received January 10, 1992; Revised Manuscript Received April 6, 1992

ABSTRACT: The secondary structure of streptokinase (Sk) in aqueous solution was quantitatively examined by using Fourier transform infrared (FT-IR) spectroscopy. Resolution enhancement techniques, including Fourier deconvolution and derivative spectroscopy, were combined with band curve-fitting procedures to quantitate the spectral information from the amide I bands. Nine component bands were found under the broad, nearly featureless amide I bands which reflect the presence of various substructures. The relative areas of these component bands indicate an amount of β-sheet between 30 and 37% and an α-helix content of only 12-13% in Sk. Further conformational substructures are assigned to turns (25-26%) and to "random" structures (15-16%). Additionally, the correlation of a pronounced component band near 1640 cm^{-1} (10-16% fractional area) with the possible presence of 3_{10}-helices is discussed.

The bacterial protein streptokinase (Sk)[1] can activate human plasminogen, and it is in clinical use as a thrombolytic agent (Collen & Gold, 1990). For understanding its structure–function relationships and to improve further its therapeutic properties by techniques of genetic engineering, knowledge on the secondary structure of Sk in solution is desirable.

Sk has a molar mass of about 47 000 $g \cdot mol^{-1}$, and its primary structure was estimated by protein sequencing (Jackson & Tang, 1982) and from the nucleotide sequence of the cloned Sk gene (Malke et al., 1985). From CD studies, a content of 17% α-helices, 28% β-sheet, 21% turns, and 34% disordered structures was derived, in agreement with results obtained by secondary structure prediction methods (Radek & Castellino,

* Author to whom correspondence should be addressed.
‡ Max Delbrück Center for Molecular Medicine.
§ Robert Koch Institute of the Federal Health Office of Germany.
‖ Institute of Experimental Microbiology.

[1] Abbreviations: CD, circular dichroism; FT-IR, Fourier transform infrared; IR, infrared; Sk, streptokinase.

0006-2960/92/0431-6532$03.00/0

1989). Studies on the solution structure of Sk applying techniques of vibrational spectroscopy and circular dichroism were performed in our laboratory. Values of 14–23% α-helices, 38–46% β-structures, 10–30% turns, and 12–23% residual structures were estimated by CD, dispersive IR, and Raman studies (Welfle, H., et al., 1992). Whereas agreement with the data described by Radek and Castellino is fairly good for the α-helix contents, the deviations for the contents of β-structure are disappointing. Furthermore, data evaluation of CD spectra of Sk obtained at increasing temperatures indicated an apparent conservation of secondary structure elements up to temperatures as high as 85 °C (Welfle, H., et al., 1992), but a complete thermal unfolding of Sk at this temperature was indicated by differential scanning calorimetric studies (Welfle, K., et al., 1992). Therefore, we are interested in a more detailed analysis of the conformational properties of Sk in its native, folded state as well as in a description of its unfolding behavior. Here we present a secondary structure analysis of Sk by FT-IR spectroscopy as a first step in this direction.

FT-IR spectroscopy is an established method for studying the secondary structure of polypeptides and proteins in aqueous solution (Parker, 1983). The structural information is predominantly derived from the analysis of the conformation-sensitive amide bands, particularly the amide I band. The numbers, frequencies, and relative intensities of the various amide I band components have been shown to be related to the molecular geometry and hydrogen-bonding pattern of the peptide backbone (Miyazawa & Blout, 1961; Krimm & Bandekar, 1986) which, ultimately, determine the specific protein substructures, i.e., α-helix, β-sheet, turns, etc. In principle, a protein containing different types of secondary structures should give rise to different amide I maxima. However, the component bands are typically broad and lie in close proximity to one another, and the observed amide I band contour often appears featureless. Recent progress in the development of methods for the visualization of overlapping bands, such as Fourier deconvolution and Fourier derivation, makes it feasible to distinguish between the individual components of the intrinsically broad and overlapping individual amide I band contours. Most, if not all, of the amide I component bands may be assigned unambiguously to specific conformational types. Thus, together with band curve-fitting procedures, the quantitative determination of the secondary structure of proteins is attainable (Byler & Susi, 1986; Surewicz & Mantsch, 1988).

FT-IR spectra of Sk were measured in H_2O and D_2O and analyzed by resolution enhancement techniques. Nine component bands were observed under the amide I bands of Sk. These bands were assigned to secondary structure elements according to known spectrum structure correlations. The extent of amino acid side-chain absorptions in the amide I region was estimated, and its effect on the determination of the contents of secondary structure elements was discussed.

Experimental Procedures

Materials. Sk was isolated from group C streptococci (*Streptococcus equisimilis*) as described earlier (Gerlach & Koehler, 1977). Sk was further purified by gel filtration on Sephadex G-100 (2.5 × 100 cm) in 10 mM sodium phosphate, pH 7.5, concentrated up to about 50 mg/mL, and lyophilized. Before the measurements, Sk was dissolved in the appropriate volumes of H_2O or D_2O, in order to yield the buffer condition of 10 mM sodium phosphate at pH 7.5 or pD 7.5.

Infrared Spectra. Infrared spectra were recorded using a Bruker IFS-66 FT-IR spectrometer equipped with a DTGS detector. The protein solutions were placed at ambient temperature in a home-made special detachable cell with CaF_2 windows. The path lengths were 8.4 μm for H_2O samples and 45 or 82.4 μm for D_2O samples. For proper compensation of H_2O and D_2O absorption, the buffer solutions were placed in the same kind of cells but with slightly lower path lengths. The protein solutions and the buffers were measured using identical scanning parameters. Typically, the spectra were recorded at a resolution of 4 cm^{-1} by co-adding 1024 interferograms which were Fourier-transformed using a Happ–Genzel apodization function.

The protein spectrum was obtained by subtracting the buffer spectrum from the spectrum of the protein solution according to procedures as described, e.g., by Dong et al. (1990). The spectrometer and sample chamber were continuously purged with dry air. Spectral contributions from uncompensated water vapor bands were eliminated using a set of water vapor spectra measured under identical conditions at different time intervals after opening the sample chamber. The subtraction factor was varied until the second derivative of the protein spectrum between 1800 and 1700 cm^{-1} was featureless. In this way, artificial bands and/or incorrect band positions in the amide I and amide II region of the protein spectrum are avoided.

Resolution Enhancement and Band Curve-Fitting. Resolution enhancement of the spectra was obtained using the Bruker software which is basically identical to that of Kauppinen et al. (1981). To calculate the second-derivative spectra, the first-derivative function was applied twice with a nine-point smoothing according to Savitzky–Golay. Also the fourth-derivative spectra were calculated because they permitted the detection of additional components.

Band curve-fitting of the deconvoluted spectra was performed using a nonlinear least-squares curve-fitting procedure based on the Marquardt algorithm (Marquardt, 1963). This program (GAUSSFIT) written in Turbo–Pascal by O. Ristau in Berlin–Buch was installed on a 32-bit PC Dell 325D. The number of bands and their positions were taken from the derivative and the deconvoluted spectra. Initial values for peak heights and widths were estimated visually from the deconvoluted spectra. A Gaussian band profile was used as a parameter for the envolope of the deconvoluted bands. For the final fits, the heights, widths, and positions of all bands were varied simultaneously.

Results

FT-IR Spectra of Sk in H_2O and D_2O. Figure 1 shows Fourier transform infrared spectra of Sk as measured from H_2O and D_2O solutions after digital subtraction of the respective buffer spectra. In H_2O buffer (Figure 1, spectrum 1), two main bands are seen in the spectral region between 1800 and 1400 cm^{-1}: the amide I band [essentially C=O stretching vibration of the amide functional group (Miyazawa & Blout, 1961; Krimm & Bandekar, 1986)] around 1644 cm^{-1} and the amide II band around 1551 cm^{-1}. Upon deuteration in D_2O buffer (Figure 1, spectrum 2), the amide I band is shifted to about 1637 cm^{-1} (amide I'). The amide II band (N–H bending vibration strongly coupled to C–N stretching) shifts to 1445 cm^{-1} (amide II'). The rate and extent of hydrogen–deuterium exchange in proteins may be obtained by measuring the decrease in intensity of the amide II band near 1550 cm^{-1}. The spectrum of Sk in D_2O buffer (Figure 1, spectrum 2) obtained after a 3-day equilibrium period at room temperature represents the protein spectrum after completion of H–D exchange because the bands around 1550 cm^{-1} (amide II) and around 3300 cm^{-1} [amide A band, essentially N–H

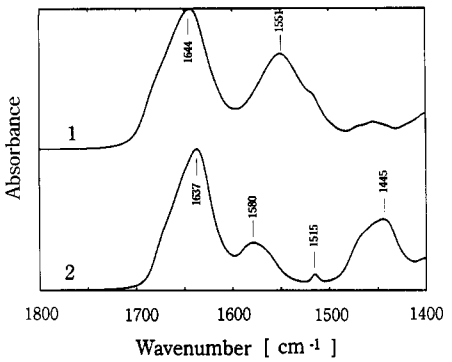

FIGURE 1: Infrared spectra of Sk in H_2O buffer (curve 1) and D_2O buffer (curve 2) after digital subtraction of the buffer spectra in the spectral region of the amide I and amide II bands.

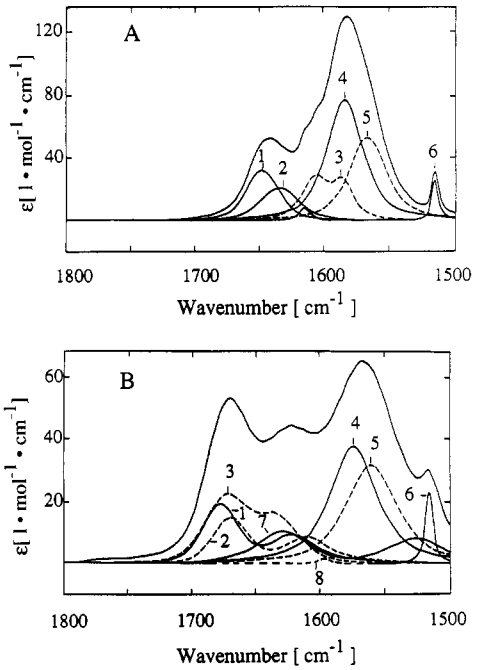

FIGURE 2: Amino acid side-chain group absorptions of Sk in D_2O (A) and in H_2O (B) calculated from the amino acid composition of Sk and a set of individual amino acid side-chain group spectra (Chirgadze et al., 1975; Venyaminov & Kalnin, 1990a). Assignments of the individual absorptions are as follows: asparagine (1); glutamine (2); arginine (3); aspartic acid (4); glutamic acid (5); tyrosine (6); lysine (7); histidine (8).

stretching (Parker, 1983)] were absent. Spectra recorded after about 1 and 6 h of H–D exchange in D_2O buffer (spectra not shown) indicated that already 90% and 95% of the amide protons were exchanged within these time intervals, respectively. The bands persisting between 1600 and 1500 cm^{-1} (Figure 1, spectrum 2) were due to amino acid side-chain absorptions (tyrosine at 1515 cm^{-1}; overlapping bands of aspartic acid, glutamic acid, and arginine around 1580 cm^{-1}) as can be seen by a comparison with Figure 2A, which shows the spectrum of the amino acid side-chain absorptions of Sk calculated from its amino acid composition (Jackson & Tang, 1982; Malke et al., 1985) and from a set (spectra 1–6 in Figure 2A) of individual amino acid spectra in D_2O (Chirgadze et al., 1975). Figure 2B shows the spectrum of the amino acid side-chain absorptions of Sk in H_2O calculated from the amino acid composition of Sk and a set (spectra 1–8 in Figure 2B) of individual amino acid side-chain spectra in H_2O (Venyaminov & Kalnin, 1990a).

The second-derivative spectra of Sk in H_2O and D_2O are shown in Figure 3. Seven- and six-component bands are clearly visible in the amide I region (1700–1620 cm^{-1}), in H_2O

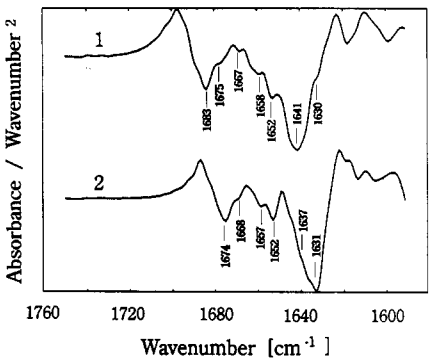

FIGURE 3: Second-derivative infrared spectra of Sk in H_2O buffer (curve 1) and D_2O buffer (curve 2) in the amide I and amide II regions.

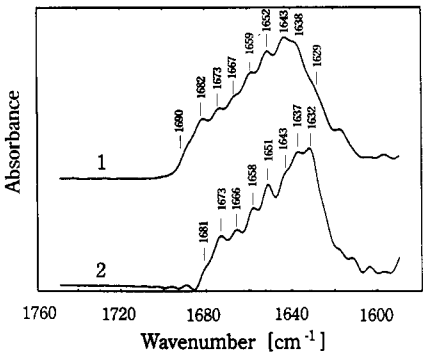

FIGURE 4: Deconvoluted infrared spectra of Sk in H_2O buffer (curve 1) and D_2O buffer (curve 2) in the amide I region. Deconvolution was performed by using a Lorentzian of 16-cm^{-1} half-bandwidth and a resolution enhancement factor (k value) of 2.9.

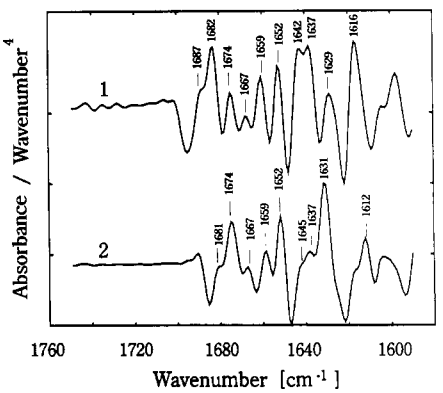

FIGURE 5: Fourth-derivative infrared spectra of Sk in H_2O buffer (curve 1) and D_2O buffer (curve 2).

and D_2O, respectively. Two more features could be resolved by Fourier deconvolution of the H_2O and D_2O spectra (Figure 4). These additional components were also visualized in the fourth-derivative spectra (Figure 5).

Assignment of Component Bands to Secondary Structure Elements. The quantitative contribution of each band to the total amide I contour was obtained by band curve-fitting of the deconvoluted H_2O (Figure 6A) and D_2O spectra (Figure 6B). Table I lists the peak positions and relative areas of the individual amide I band components as obtained by these techniques. The assignment of the individual bands to secondary structure elements is essentially based on FT-IR data of proteins published recently (Krimm & Bandekar, 1986; Surewicz & Mantsch, 1988; Olinger et al., 1986; Prestelski et al., 1991). In D_2O, the bands at 1632 and 1625 cm^{-1} were assigned to different β-structures. The low-frequency component near 1625 cm^{-1} was interpreted as arising from a particular β-structure. It has been suggested that such a low frequency is indicative of a particularly tightly hydrogen-

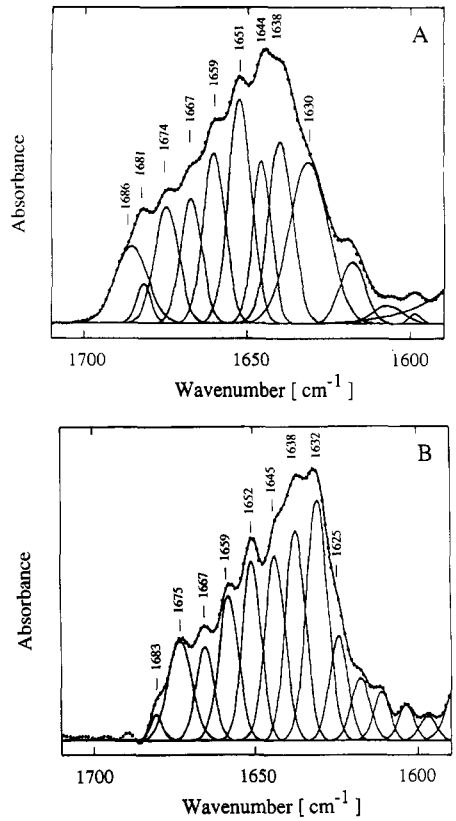

FIGURE 6: Deconvoluted amide I band of Sk in H_2O buffer (A) and D_2O buffer (B) with the best-fitted individual component bands. The figure contains the experimental spectrum (· ·), the individual Gaussian components, and their sum (—). Note: Several bands arising from side-chain absorptions (see Figure 2) are fitted below 1620 cm^{-1} which are not amide I components. These are added to the analysis to avoid the approximation otherwise incurred with the addition of a sloping base-line parameter.

Table I: Peak Positions and Relative Integrated Intensities of the Amide I Bands of Sk

H_2O		D_2O		
band position (cm^{-1})	band area (%)	band position (cm^{-1})	band area (%)	assignment
1686	8	1683	1	turn/β-sheet
1681	2	1675	9.5	turn/β-sheet
1674	9	1667	6.5	turn
1667	8.5	1659	11.5	turn
1659	12	1652	13	α-helix
1651	16.5	1645	15	"random"
1644	10	1638	16	3_{10}-helix, "open loop", turn (β-sheet)
1638	14	1632	20	β-sheet
1630	20	1625[a]	7.5	β-sheet

[a] To fit the spectrum of Sk in D_2O, an additional component around 1625 cm^{-1}, although not revealed by the second-derivative spectrum (Figure 3, spectrum 2) or the deconvoluted spectrum (Figure 4, spectrum 2), was added to improve the agreement between the observed spectrum (Figure 4, spectrum 2) and the calculated spectrum. Without this additional component, no satisfactory fit in the region 1635–1600 cm^{-1} was obtained.

bonded β-sheet (Arrondo et al., 1988; Casal et al., 1988). In H_2O, two bands observed at 1638 and 1630 cm^{-1} were assigned to different β-structures, too.

The components found at 1652 and 1645 cm^{-1} in D_2O were assigned to α-helical segments and to "random" conformations, respectively (Byler & Susi, 1986; Surewicz & Mantsch, 1988). For proteins measured in H_2O, the amide I bands of α-helical and "random" structures are generally thought to exhibit very similar positions and, therefore, cannot be resolved into distinct components by resolution enhancement procedures. In our spectra of Sk measured in H_2O, however, two bands could be resolved in the spectral region typical for α-helical and "random" conformations (around 1655 cm^{-1}), similar to that observed for Sk in D_2O. The component at 1659 cm^{-1} was assigned to α-helical segments, and the component at 1651 cm^{-1} was assigned to "random" structures on the basis of a report of very accurate measurements of infrared spectra of a set of proteins in H_2O published very recently (Dong et al., 1990).

The bands located at 1667 and 1659 cm^{-1} in the spectrum of Sk in D_2O and at 1674 and 1667 cm^{-1} in H_2O were assigned to turn structures (Byler & Susi, 1986; Surewicz & Mantsch, 1988; Olinger et al., 1986; Arrondo et al., 1988; Casal et al., 1988; Holloway & Mantsch, 1989; Prestelski et al., 1991).

There is some disagreement in the literature concerning the assignment of band components between 1675 and 1690 cm^{-1} in the infrared spectra of proteins. Theoretical calculations for antiparallel β-sheets predict an infrared-active component in this region (Krimm & Bandekar, 1986). One band near 1675 cm^{-1} (Byler & Susi, 1986) or even all bands in the region 1689–1675 cm^{-1} (Holloway & Mantsch, 1989) were assigned to the high-frequency component as diagnostic for an antiparallel β-structure. However, contributions from β-structure and turns seem to be more realistic (Dong et al., 1990; Olinger et al., 1986; Prestelski et al., 1991). Thus, the unambiguous identification and quantification of the high-frequency β-component(s) are often problematic. The high-frequency β-component is only about one-tenth as intense as the low-frequency component in the spectra of model polypeptides in the β-sheet conformation as well as in the spectra of proteins with very high β-sheet contents (Byler & Susi, 1986; Arrondo et al., 1988, Casal et al., 1988; Dousseau & Pezolet, 1990; Kalnin et al., 1990). Considering the intensities of the low-frequency β-components in the spectra of Sk and the intensities of the component bands in the region 1689–1675 cm^{-1}, it becomes obvious that in the region above 1675 cm^{-1} also contributions from structures other than β-structure have to be considered. The band components in this region contain contributions from turns and β-structures, and we expect the β-sheet contributions to be small.

Striking is the presence of a pronounced component band at 1638 and 1644 cm^{-1} in the spectra of Sk in D_2O and H_2O, respectively. Most investigators usually assign all bands in the region from 1642 to 1620 cm^{-1} solely to β-sheet structures (Byler & Susi, 1986; Dong et al., 1990). Very recently, however, additional assignments were suggested. Proteins containing, e.g., a relatively high amount of another type of helix, termed 3_{10}-helix, exhibited a band component near 1639 cm^{-1} (Holloway & Mantsch, 1989; Prestelski et al., 1991). On the basis of a good correlation between the secondary structure derived from the infrared data and the results of X-ray diffraction, the authors concluded that the band near 1639 cm^{-1} should be assigned to 3_{10}-helices. The assignment of a band component near 1640 cm^{-1} to 3_{10}-helices seems to be unsettled, however, because FT-IR studies of model 3_{10}-helical oligopeptides showed the main amide I band near 1665 cm^{-1} (Yasui et al., 1986; Kennedy et al., 1991).

Furthermore, a band near 1640 cm^{-1} was observed which could be assigned most satisfactorily to β-turns (H. H. Mantsch, personal communication) or "open loop" structures (fully hydrated, not interacting with proximate amide functional groups) on the basis of FT-IR studies on the proteins RNase T1 and acetylcholine esterase in solution (D. Naumann et al., unpublished results). From these facts, it can be concluded that for the assignment of the component bands near

Table II: Contents of Secondary Structure Elements of Sk

	α-helix (%)	β-sheet (%)	turn (%)	"random" (%)	others (%)
this work[a] ("FT-IR method")	12–13	30–37	25–26	15–16	10–16
CD[b]	17	28	21	34	
CD[c]	21	46	10	23	
"IR-method"[d]	16 (14+2) (o+d)	40 (17+23) (o+d)	26	19	

[a] For the given numbers, the mean values of the H_2O and D_2O measurements were considered. For the relative estimation of the β-sheet and turn contributions to the high-frequency components, see the text. [b] CD spectra in the region 185–260 nm were analyzed by the approach of Hennessey and Johnson (Radek & Castellino, 1989). [c] CD spectra in the region 198–240 nm were analyzed by the CONTIN algorithm (Welfle, H., et al., 1992). [d] Infrared spectra of Sk in H_2O were analyzed by two methods: a constrained statistical regularization method and a procedure using an orthogonal set of basic spectra according to Kalnin et al. (1990). Both methods gave very similar results (Welfle, H., et al., 1992), and the averaged values are given. As suggested by Kalnin et al. (1990), helical elements are divided into disordered (d) and ordered (o) helices, comprising two residues on each end of the helical segments and the rest of the segments, respectively. From the total amount of residues assigned to β-sheets, those having "classical" H-bonds with at least one neighbor β-strand were defined to be ordered (o); the rest of the β-sheets were considered to be disordered (d).

1640 cm^{-1} in the spectra of Sk, probably not β-sheets but 3_{10}-helices, "open loops", or β-turn structures have to be considered. An unequivocal assignment seems to be impossible at the present stage of knowledge, but the possible existence of 3_{10}-helices or "open loops" in Sk suggested by the component band near 1640 cm^{-1} is an interesting aspect for further elucidation of the structure of Sk.

Effects of Side-Chain Absorptions in the Amide I Region on the Secondary Structure Determination. The quantitative analysis of the secondary structure composition from the amide I band requires accurate determination of the infrared amide I band that is due solely to the pure peptide absorption. It is known that not only the amide functional groups but also the amino acid side-chain groups can absorb in the amide I region (Chirgadze et al., 1975; Venyaminov & Kalnin, 1990a). To estimate the contribution of side-chain absorptions to the spectra of Sk, the amino acid side-chain group spectra calculated for Sk in D_2O (Figure 2A) and H_2O (Figure 2B) were subtracted from the infrared spectra of Sk. On the basis of the molar absorptivities of the experimental spectra of Sk in H_2O (Welfle, H., et al., 1992) and in D_2O, a contribution of amino acid side-chain groups to the spectra of 14% and 10% of the total integral intensity in the amide I region of Sk in H_2O and D_2O, respectively, was estimated. These contributions could be neglected if they would come from a continuous background absorption, but unfortunately, the contributions of the side-chain absorptions are not balanced in the amide I region (Figure 2A,B). Therefore, accurate subtraction of the side-chain absorptions from the experimental protein spectrum would be desirable before any further spectra treatment such as, e.g., Fourier deconvolution and band curve-fitting. Unfortunately, the side-chain absorption spectra of Sk shown in Figure 2A,B are not very useful for this purpose. First, due to the low signal-to-noise ratio of the spectra of the side-chain groups measured earlier on a dispersive infrared instrument (Chirgadze et al., 1975; Venyaminov & Kalnin, 1990a), also difference spectra (peptide absorptions) of only low signal-to-noise ratio were obtained, thus preventing a satisfying Fourier deconvolution which is a prerequisite for the detailed band curve-fitting. Second, the spectra of individual free amino acids, which have been used to calculate the amino acid side-chain group absorptions (Figure 2A,B), may differ from the spectra of the side-chain groups localized in specific protein microenvironment(s).

Nevertheless, an estimation of side-chain contributions to the amide I band contour and thus to the final secondary structure determination can be approximated on the basis of the spectra shown in Figure 2A,B. Only a minor effect can be expected for protein D_2O spectra. Side chains of asparagine and glutamine residues absorb at specific frequencies only in the region from 1650 to 1635 cm^{-1} (Figure 2A) (Chirgadze et al., 1975). Taking into account the total contribution of side-chain groups of only 10% to the total integral intensity in the amide I region (1700–1620 cm^{-1}) of Sk in D_2O and the different frequencies of the side-chain absorptions, it can be concluded that individual contributions of asparagine (near 1649 cm^{-1}) and glutamine (near 1635 cm^{-1}) residues disturb the area estimation of deconvoluted bands in this region. This gives rise to an overestimation of the band areas in the region 1650–1630 cm^{-1} in comparison to the areas of the band components above 1650 cm^{-1}.

The spectral characteristics of the side-chain absorptions of Sk in H_2O (Figure 2B) are different from those in D_2O (Figure 2A). Side chains of asparagine (near 1678 cm^{-1}), glutamine (near 1670 cm^{-1}), and arginine residues (near 1672 cm^{-1}) absorb in the region 1680–1670 cm^{-1} and less intensely absorb arginine (near 1638 cm^{-1}), lysine (near 1630 cm^{-1}), and asparagine residues (near 1623 cm^{-1}) in the region from 1640 to 1620 cm^{-1} (Figure 2B). These absorptions contribute 14% to the total integral amide I intensity and influence the spectral characteristics and the component band areas in these regions. Therefore, the use of the uncorrected spectra of Sk in H_2O for the secondary structure determination may be the source of small differences from values expected for corrected spectra. This means the data given in Table I can be considered as fairly realistic but improvable to some extent when amino acid side-chain absorption FT-IR spectra would become available. The different contributions of the absorption spectra of the side-chain groups to the experimental spectra of Sk in H_2O and D_2O, respectively, could explain some differences in the calculated relative band areas (Table I), e.g., between the band components 1638/1630 cm^{-1} in H_2O and 1632/1625 cm^{-1} in D_2O or in the region 1680–1670 cm^{-1}.

Discussion

Assignment of the various amide I band components of Sk and the calculated intensities permits a quantitative estimation of the secondary structure of Sk. These data are summarized in Table II, together with the results obtained by other methods (Radek & Castellino, 1989; Welfle, H., et al., 1992).

An important assumption common to methods using Fourier deconvolution followed by band curve-fitting (Byler & Susi, 1986; Surewicz & Mantsch, 1988; Holloway & Mantsch, 1989; Prestelski et al., 1991), second-derivative analysis of the amide I spectra (Dong et al., 1990), and partial least-squares methods (Dousseau & Pezolet, 1990) as well as factor analysis of infrared spectra (Lee et al., 1990) is that the integrated molar absorptivities are equal for each type of structure. The good agreement observed between the data obtained by FT-IR spectroscopy (Byler & Susi, 1986; Surewicz & Mantsch, 1988; Dong et al., 1990; Dousseau & Pezolet, 1990, Lee et al., 1990) and the secondary structures derived from single-crystal X-ray

analysis of many proteins supports this assumption. Recent arguments against the generalization of this assumption come from infrared studies on polypeptides and proteins which can exist in different conformations (Dousseau & Pezolet, 1990; Jackson et al., 1989; Mantsch et al., 1989; Venyaminov & Kalnin, 1990b). These studies indicate that particularly the content of "random" structures may be underestimated by assuming identical integrated molar absorptions for each type of structure. Such possible uncertainties should be kept in mind when relative band areas in the amide I region are used for quantitative estimation of secondary structure elements in proteins.

By summing up the band areas (Table I) assigned to specific secondary structure elements, contents of 12–13% α-helix, 15–16% "random" structure, 30–37% β-sheet, 25–26% turns, and 10–16% of other structure(s) in Sk can be derived from the FT-IR studies (Table II). For the calculation of the β-sheet content, we assumed that the high-frequency β-component(s) is (are) only one-tenth as intense as the low-frequency component(s). In this way, 30% (27.5% in the region below 1635 cm^{-1} plus 3% above 1675 cm^{-1}) and 37% (34% in the region below 1640 cm^{-1} plus 3% above 1675 cm^{-1}) were calculated for the total amount of β-sheet from the band areas in D_2O and H_2O, respectively. By summing up the band areas above 1660 cm^{-1} (27.5% in H_2O and 28.5% in D_2O, respectively) minus 3%, a portion of turns between 25 and 26% was obtained. Thus, a β-sheet content of approximately 34% and a content of turns of about 25% seem to be in agreement with our experimental data.

The amounts of the various secondary structures present in Sk as estimated by FT-IR spectroscopy and CD spectroscopy (CD^a and CD^b in Table II) are similar, but not identical. The general limits of the quantitative determination of the secondary structure of proteins from CD spectra (Johnson, 1990; Venyaminov et al., 1991), especially the specific problems of the CONTIN algorithm in the evaluation of the β-sheet contents of proteins from their CD spectra (Bobba et al., 1990), might be reasons for some differences. Table II also gives the results of an estimation of the contents of secondary structure elements in Sk (Welfle, H., et al., 1992) as obtained by the fitting of the shape of the amide I and amide II bands by a set of spectra of proteins of known X-ray structure according to a procedure introduced very recently (Kalnin et al., 1990), named "IR-method" here. These data are compared with the values obtained in this work by the described FT-IR techniques ("FT-IR method"). Before comparison of the results in detail, we would like to mention some problems inherent to these procedures. When FT-IR spectra of proteins are analyzed by Fourier deconvolution and band curve-fitting, contributions of amino acid side-chain groups may introduce uncertainties, "random" structures are possibly underestimated, and the amide I component bands must be assigned unambiguously to specific conformational types. Using the "IR-method" introduced by Kalnin and Venyaminov (1990), the effect of side-chain absorptions is effectively eliminated by calculating the pure peptide absorption spectrum. Furthermore, different molar absorptivities of different types of structures are taken into account by comparing the spectrum of the protein studied with the spectra of the reference proteins on the basis of molar extinction coefficients. Here, errors may only be introduced by an inaccurate estimation of the concentration of the protein solution and/or of the thickness of the infrared cell. Problems may arise, however, when spectra of proteins with conformations strongly deviating from the structures of the reference proteins are analyzed. Proteins exhibiting "unusual" secondary structures should give poor fits to the spectra of the reference data base comprising only "normal" protein structures, and thus inaccurate (Lee et al., 1990) or worse results are obtained.

Taking into account these limitations, then the results obtained for α-helix (12–13% by the "FT-IR method" versus 16% by the "IR-method"), β-sheet (about 34% by the "FT-IR method" versus 40% by the "IR-method), turns (about 25% by the "FT-IR method" versus 26% by the "IR-method"), and "random" (15–16% by the "FT-IR method" versus 19% by the "IR-method") are in good agreement. Structures other than α-helices, turns, β-sheets, and "remainders", as derived from the X-ray structures of the reference proteins according to the algorithm used by Kalnin et al. (1990) cannot be recognized by the "IR-method". The infrared bands near 1640 cm^{-1} assigned to other structures by the FT-IR studies are located between the position of the infrared bands typical for "random" structures (1650–1645 cm^{-1}) and β-structures (1638–1620 cm^{-1}). Therefore, an influence on the fractions of β-sheet and "random" estimated by the "IR-method" can be expected. This may account for a higher content of β-sheet plus "random" (59%) as estimated by the "IR-method" in comparison to about 50% as derived from the "FT-IR method" used in this work.

The main advantage of the Fourier transform infrared spectroscopic technique as applied in this work is not only that there is additional support for the quantification of the relative fractions of the various secondary structures present in Sk but also that it provides a more precise correlation on experimentally observed spectral characteristics to defined structural elements. Further studies on the solution structure of Sk, especially on structural changes as a function of temperature and other effectors, will surely benefit from these experiments.

ACKNOWLEDGMENTS

We thank Drs. S. Yu. Venyaminov and N. N. Kalnin for providing the infrared spectra of the amino acid side-chain groups, Mrs. B. Kannen for skillful technical assistance, and Mrs. A. Staeuble for her help in preparation of the figures.

REFERENCES

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

Bobba, A., Cavatorta, P., Attimonelli, M., Riccio, P., Masotti, L., & Quagliariello, E. (1990) *Protein Sequences Data Anal. 3*, 7–10.

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

Casal, H. L., Köhler, U., & Mantsch, H. H. (1988) *Biochim. Biophys. Acta 757*, 11–20.

Chirgadze, Yu. N., Fedorov, O. V., & Trushina, N. P. (1975) *Biopolymers 14*, 679–694.

Collen, D. C., & Gold, H. K. (1990) *Thromb. Res., Suppl. X*, 10–131.

Dong, A., Huang, D., & Caughey, W. S. (1990) *Biochemistry 29*, 3303–3308.

Dousseau, F., & Pezolet, M. (1990) *Biochemistry 29*, 8771–8779.

Gerlach, D., & Koehler, W. (1977) *Zentralbl. Bakteriol., Mikrobiol. Hyg., Abt. 1, Orig. A 238*, 336–349.

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

Jackson, K. W., & Tang, J. (1982) *Biochemistry 21*, 6620–6625.

Jackson, M., Haris, P. I., & Chapman, D. (1989) *Biochim. Biophys. Acta 998*, 75–79.

Johnson, W. C., Jr. (1990) *Proteins: Struct., Funct., Genet. 7*, 205–214.

Kalnin, N. N., Baikalov, I. A., & Venyaminov, S. Yu. (1990) *Biopolymers 30*, 1273-1280.
Kauppinen, J. K., Moffat, D. J., Mantsch, H. H., & Cameron, D. G. (1981) *Appl. Spectrosc. 35*, 271-276.
Kennedy, D. F., Crisma, M., Toniolo, C., & Chapman, D. (1991) *Biochemistry 30*, 6541-6548.
Krimm, S., & Bandekar, J. (1986) *Adv. Protein Chem. 38*, 181-364.
Lee, D. C., Haris, P. I., Chapman, D., & Mitchell, R. C. (1990) *Biochemistry 29*, 9185-9193.
Malke, H., Roe, B., & Ferretti, J. J. (1985) *Gene 34*, 357-362.
Mantsch, H. H., Surewicz, W. K., Muga, A., Moffatt, D. J., & Casal H. L. (1989) *Proc. SPIE—Int. Soc. Opt. Eng. 1145*, 580-581.
Marquardt, D. W. (1963) *J. Soc. Ind. Appl. Math. 11*, 431-441.
Miyazawa, T., & Blout, E. R. (1961) *J. Am. Chem. Soc. 83*, 712-719.
Olinger, J. M., Hill, D. M., Jakobson, R. J., & Brody, R. S. (1986) *Biochim. Biophys. Acta 869*, 89-98.
Parker, F. S. (1983) *Applications of Infrared, Raman, and Resonance Raman Spectroscopy in Biochemistry*, Plenum Press, New York.
Prestrelski, S. J., Byler, D. M., & Thompson, M. P. (1991) *Int. J. Pept. Protein Res. 37*, 508-512.
Radek, J. T., & Castellino, F. J. (1989) *J. Biol. Chem. 264*, 9915-9922.
Surewicz, W. K., & Mantsch, H. H. (1988) *Biochim. Biophys. Acta 952*, 115-130.
Venyaminov, S. Yu., & Kalnin, N. N. (1990a) *Biopolymers 30*, 1243-1257.
Venyaminov, S. Yu., & Kalnin, N. N. (1990b) *Biopolymers 30*, 1259-1271.
Venyaminov, S. Yu., Baikalov, I. A., Wu, Ch.-S. C., & Yang, J. T. (1991) *Anal. Biochem. 198*, 250-255.
Welfle, H., Misselwitz, R., Fabian, H., Hoelzer, W., Damerau, W., Gerlach, D., Kalnin, N. N., & Venyaminov, S. Yu. (1992) *Int. J. Biol. Macromol. 14*, 9-18.
Welfle, K., Pfeil, W., Misselwitz, R., Welfle, H., & Gerlach, D. (1992) *Int. J. Biol. Macromol. 14*, 19-22.
Yasui, S. C., Keiderling, T. A., Bonora, G. M., & Toniolo, C. (1986) *Biopolymers 25*, 79-89.

Penetration of Analogues of H_2O and CO_2 in Proteins Studied by Room Temperature Phosphorescence of Tryptophan†

Wayne W. Wright,‡ C. S. Owen,§ and J. M. Vanderkooi*,‡

Department of Biochemistry and Biophysics, School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, and Biochemistry Department, School of Medicine, Jefferson Medical College, Philadelphia, Pennsylvania 19107

Received February 4, 1992; Revised Manuscript Received April 27, 1992

ABSTRACT: The influence of the protein matrix on the reactivity of external molecules with a species buried within the protein interior is considered in two general ways: (1) there may be structural fluctuations that allow for the diffusive penetration of the small molecules and/or (2) the external molecule may react over a distance. As a means to study the protein matrix, a reactive species within the protein can be formed by exciting tryptophan to the triplet state, and then the reaction of the triplet-state molecule with an external molecule can be monitored by a decrease in phosphorescence. In this work, the quenching ability (i.e., reactivity) was examined for H_2S, CS_2, and NO_2^- acting on tryptophan phosphorescence in parvalbumin, azurin, horse liver alcohol dehydrogenase, and alkaline phosphatase. A comparison of charged versus uncharged quenchers (H_2S vs SH^- and CS_2 vs NO_2^-) reveals that the uncharged molecules are much more effective than charged species in quenching the phosphorescence of fully buried tryptophan, whereas the quenching for exposed tryptophan is relatively independent of the charge of the quencher. This is consistent with the view that uncharged triatomic molecules can penetrate the protein matrix to some extent. The energies of activation of the quenching reaction are low for the charged quenchers and higher for the uncharged CS_2. A model is presented in which the quenchability of a buried tryptophan is inversely related to the distance from the surface when diffusion through the protein is the rate-limiting step. Using this model, upper limit values of the diffusion coefficient in the protein for these molecules can be estimated to be about 7×10^{-11} cm^2 s^{-1} through the protein matrix. Finally, since the quencher molecules H_2S and CS_2 resemble H_2O and CO_2 in size and charge, it is suggested that these molecules would show similar diffusion behavior.

How a molecule that is external to a protein reacts with a species buried within the protein interior will necessarily be influenced by the properties of the protein. As a means to study this process, we have introduced the use of phosphorescence quenching of intrinsic tryptophan by quenchers that are free to diffuse in solution. It is now well recognized that tryptophans in a rigid protein environment exhibit long-lived phosphorescence at room temperature, provided that external quenchers, notably O_2, are removed from the solution [Saviotti & Galley, 1974; Kai & Imakubo, 1979; Strambini, 1987; Vanderkooi et al., 1987; Vanderkooi & Berger, 1989; reviewed by Papp and Vanderkooi (1989)]. Phosphorescence

† This work was supported by NIH Grant GM 34448 (J.M.V.) and by NSF Grant DCB-9118267 (C.S.O.).
* To whom correspondence should be addressed.
‡ University of Pennsylvania.
§ Jefferson Medical College.

0006-2960/92/0431-6538$03.00/0