Anal. Chem. **2002,** *74,* 4076−4080

Single-Pass Attenuated Total Reflection Fourier Transform Infrared Spectroscopy for the Analysis of Proteins in H_2O Solution

Brandye M. Smith and Stefan Franzen*

Department of Chemistry, North Carolina State University, Raleigh, North Carolina 27695-8204

The application of single-pass attenuated total reflection Fourier transform infrared (ATR-FT-IR) microscopy was investigated for secondary structure analysis of 15 representative proteins in H_2O solution. This is the first reported application of single-pass ATR-FT-IR for protein analysis; thus, the method was validated using transmission FT-IR and multipass ATR-FT-IR as referee methods. The single-pass ATR-FT-IR technique was advantageous since the single-pass geometry permits rapid secondary structure analysis on small volumes of protein in H_2O solution without the use of demountable thin path length sample cells. Moreover, the fact that H_2O backgrounds were small allowed the simultaneous observation of the amide I−III, A, and B regions without having to perform H_2O subtraction. A comparison of replicate protein spectra indicated that the single-pass ATR-FT-IR method yields more reproducible data than those acquired by transmission FT-IR. The observed trends for the amide I−III and A bands obtained by single-pass ATR-FT-IR agreed with those in the literature for conventional transmission FT-IR.

The interpretation of genetic data is dependent upon the ability to establish the structure and function of a vast number of proteins known only by their primary sequence. Large-scale screening of proteins to identify structural motifs relies on the development of techniques that rapidly provide information on protein secondary structure. Such techniques can be combined with primary sequence information and homology modeling to obtain three-dimensional structures.[1,2] During the course of the last thirty years, Fourier transform infrared spectroscopy (FT-IR) has been developed as a procedure for protein secondary structure determination.[3−12] In all, proteins have nine characteristic absorption amide bands, labeled amide A, B, and I−VII; in the mid-infrared, that can be interpreted in terms of structure.[4,13] Presently, protein secondary structure determination is largely based on the examination of amide I as acquired by transmission FT-IR or multipass attenuated total reflectance (ATR)-FT-IR.[4,5,8−11,13,14] Amide I consists principally of the C═O stretching internal coordinate with a minor contribution from C−N−H in-plane bending within the peptide backbone. Transmission FT-IR and multipass ATR-FT-IR have disadvantages associated with solvent interference and sample cell assembly. This paper introduces the use of single-pass ATR-FT-IR to overcome such disadvantages as well as to maximize spectral information and to improve the efficiency of protein analysis.

H_2O is the ideal solvent for biological samples; however, the intense H_2O bending mode (1640 cm^{-1}) coincides with the amide I region (1630−1690 cm^{-1}).[3−6,13,15] To overcome H_2O interference, the majority of FT-IR investigations have focused on the amide I′ band in D_2O solution. The deuterium-exchanged amide bands are referred to as amide I′, amide II′, etc., in the spectroscopy literature. As with H_2O, there are problems associated with the use of D_2O that include the following: (1) the exchange of proteins into D_2O is tedious and can compromise the integrity of some samples, (2) hydrogen atoms within the protein exchange with deuterium over a wide range of time scales unless the protein is fully denatured,[13,15−18] and (3) absorption bands due to H−O−D, D−O−D, and H−O−H are all present upon the introduction of D_2O[5] and the presence of H−O-D in solution overlaps the amide II′, amide A′, and amide B′ regions. Thus, limitations exist for either solvent. Despite the potential for highly informative spectra, the problems associated with interference from solvent bands combined with involved sample preparation have made transmission FT-IR less attractive as a routine technique for secondary structure determination.[3−6,13,15]

Rather than D_2O exchange, H_2O interference can be minimized by a reduction in sample path length. Specifically, a path length reduction resulting in less than 0.3 AU for H_2O would result in reliable protein spectra in H_2O solution. Transmission FT-IR

* Corresponding author. Telephone: (919) 515-8915. Fax: (919) 515-8909. E-mail: Stefan_Franzen@ncsu.edu.

(1) Teichmann, S. A.; Chothia, C.; Church, G. M.; Jong, P. *Bioinformatics* **2000**, *16*, 117−124.
(2) Zhang, C. T.; Zhang, R. *J. Biomol. Struct. Dyn.* **2000**, *17*, 829−842.
(3) Byler, D.; Susi, H. *Biopolymers* **1986**, *25*, 469−487.
(4) Susi, H.; Byler, D. *Methods Enzmol.* **1986**, *130*, 290−311.
(5) Jencks, W. *Methods Enzmol.* **1986**, *6*, 914−929.
(6) Douseeau, F.; Pezolet, M. *Biochemistry* **1990**, *29*, 8771−8779.
(7) Purcell, J.; Susi, H. *J. Biochem. Biophys. Methods* **1984**, *9*, 193−199.
(8) Miyazawa, T.; Shimanouchi, T.; Mizushima, S. *J. Chem. Phys.* **1956**, *24*, 408−418.
(9) Miyazawa, T. *J. Chem. Phys.* **1960**, *32*, 1647−1652.
(10) Krimm, S. *J. Mol. Biol.* **1962**, *4*, 528−540.
(11) Krimm, S.; Abe, Y. *Proc. Natl. Acad. Sci. U.S.A.* **1972**, *69*, 2788−2792.
(12) Griffiths, P. R. *Science* **1983**, *222*, 297−302.
(13) Susi, H. *Methods Enzymol.* **1986**, *26*, 22.
(14) Baello, B.; Pancoska, P.; Keiderling, T. *Anal. Biochem.* **2000**, *280*, 46−57.
(15) Jongh, H.; Goormaghtigh, E.; Ruysschaert, J. *Anal. Biochem.* **1996**, *242*, 95−103.
(16) Nabet, A.; Pezolet, M. *Appl. Spectrosc.* **1997**, *51*, 466−469.
(17) Powell, J.; Wasacz, F.; Jakobsen, R. *Appl. Spectrosc.* **1986**, *40*, 339−344.
(18) Dousseau, F.; Therrien, M.; Pezolet, M. *Appl. Spectrosc.* **1989**, *43*, 538−542.

10.1021/ac020103v CCC: $22.00
Published on Web 06/27/2002

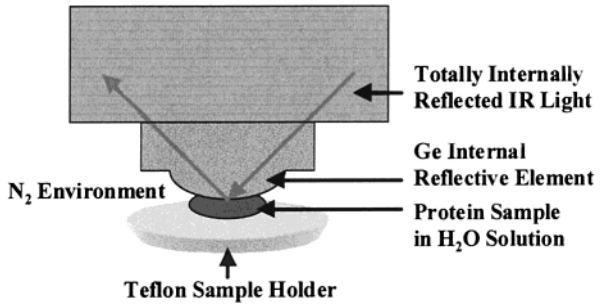

Figure 1. Geometry of the single-pass attenuated total reflection cell. The protein sample is placed in a depression below the Ge crystal.

requires the placement of proteins between two calcium fluoride (barium fluoride, zinc selenide, etc.) windows with thin path length spacers (3−25 μm). Even with a path length of 3−6 μm, there is a potential for solvent interference. ATR-FT-IR is the desired method to overcome solvent masking since the penetration depth of infrared light is inherently limited to a fraction of the wavelength estimated to be $\lambda/10$. For example, when considering a wavelength of 1650 cm^{-1}, the corresponding penetration depth is ∼0.4 μm. Therefore, the effective path length for ATR-FT-IR is sufficiently short to enable the analysis of protein in H_2O solution. This in turn has led to the use of algorithms for subtraction of H_2O from protein spectra that are provided in the literature.[15,18,19] Subtraction of H_2O from the protein sample spectrum has provided a means to simultaneously observe the amide I and II bands for the first time.[6] However, there are potential difficulties with subtraction algorithms, particularly artifacts that are caused by the large size of the H_2O signals relative to the protein bands. In addition, subtraction algorithms do not account for H_2O−protein interactions.

Transmission FT-IR and multipass ATR-FT-IR techniques require a cell to confine the protein solution. The cell geometry in a transmission or multipass experiment can change upon disassembly leading to variations in path length, especially in the transmission mode. It has even been stated in the literature that an external method for sample loading and unloading needs to be developed to avoid changes in path length and deviations in the angle of incidence.[17] The single-pass technique presented here attempts to provide such a method for sample deposition and recovery.

To circumvent problems associated with solvent interference and sample cell assembly, the single-pass technique developed by Harrick a decade ago[20] can be used. This paper is the first in a two-part study where the focus is on the use of single-pass ATR-FT-IR for protein secondary structure determination. The method of single-pass ATR-FT-IR is valuable since it is a rapid technique that does not require protein exchange into D_2O or sample cell assembly. The single-pass ATR-FT-IR sample geometry shown in Figure 1 exposes the sample to a N_2 environment and allows for the easy deposition and retrieval of protein samples from the Teflon reservoir. Since the protein is in an N_2 environment, it slowly dehydrates into a concentrated gel state. Rapid scanning can continuously monitor the changes from a fully hydrated state to a concentrated gel state. Concentrating the protein to a gel state yields spectral enhancement such that protein amide bands can be observed simultaneously without performing H_2O subtraction. This paper marks the first time that single-pass ATR-FT-IR has been reported for H_2O solutions of proteins. Therefore, this paper will focus on the validation of the method for protein secondary structure analysis.

Table 1. A List of All the Proteins Used in This Research

protein	company protein purchased from	Catalog No.
bovine serum albumin	Sigma	A-7638
α-casein	Sigma	C-6780
α-chymotrypsin	ICN	100461
chymotrypsinogen	ICN	100477
concanavalin A	ICN	150710
cytochrome *c*	Aldrich	10,520-1
hemoglobin	Sigma	H-2500
lysozyme	ICN	195303
lysozyme	Sigma	L-6876
myoglobin	Sigma	M-1882
papain	ICN	100921
pepsin	Sigma	P-6887
ribonuclease A	ICN	193980
trypsin	ICN	153571
trypsin inhibitor	ICN	100612
trypsinogen	Sigma	T-1143

EXPERIMENTAL SECTION

The proteins listed in Table 1 were prepared without further purification to a final concentration of ∼3 mM in H_2O. In the experimental apparatus, the Ge crystal is at the focus of a Cassagranian objective in a UMA500 microscope (Digilab). The sample was injected onto a cylindrical sample well that was milled in a Teflon block (refer to Figure 1). More specifically, 10−20 μL of the sample was injected onto the Teflon block using a Wheaton automatic pipet. The protein spectra were recorded at ambient temperature and averaged over 64 scans on a Digilab FTS 6000 FT-IR spectrometer equipped with a liquid nitrogen-cooled MCT detector in a single-pass ATR mode and in a multipass mode.

The protein spectra were recorded with a resolution of 2 cm^{-1}. Single-pass ATR-FT-IR spectra were recorded immediately after the sample was deposited onto the Teflon block. A steady stream of N_2 gas over the Teflon block was used to gently dehydrate the protein samples. Subsequent collection of spectra allowed for the observation of three states that are referred to throughout this paper as the hydrated state, the intermediate state, and the gel state (refer to Figure 2). As the names imply, the hydrated state refers to the protein sample when it is first analyzed whereas the intermediate and gel states refer to the protein sample as it is allowed to concentrate in the N_2 environment in which it is placed. The gel state is differentiated from the intermediate state in that when in the gel state, no more spectral changes are evident. In both single-pass and multipass mode, the Ge crystal was rinsed with solvent and allowed to dry prior to loading another protein sample.

Denaturation of a protein due to gel formation or to an interaction with the internal reflection element has been discussed in the literature and is a concern for all ATR-FT-IR techniques.[21]

(19) Chittur, K. *Biomaterials* **1998**, *19*, 357−369.
(20) Sommer, A. J.; Hardgrove, M. *Vib. Spectrosc.* **2000**, *24*, 93−100.
(21) Oberg, K.; Fink, A. *Anal. Biochem.* **1998**, *256*, 92−106.

In the single-pass ATR method, acquiring protein solution spectra as the protein slowly concentrated enabled any protein denaturation to be observed and distinguished from the native state. The ability to differentiate those proteins that suffer from denaturation under the conditions of the experiment is shown in comparison spectra given in Figures 1 and 2 in the Supporting Information. Any spectra that showed evidence of protein denaturation were discarded.

Transmission FT-IR was used as a referee method for the validation of the single-pass ATR-FT-IR technique. Since H_2O is not an adequate solvent for amide I analysis via transmission FT-IR, the proteins listed in Table 1 were also prepared to a final concentration of 3 mM in D_2O. When in transmission mode, a liquid nitrogen-cooled wideband MCT detector was used. The sample cell consisted of CaF_2 windows separated by a either a 6- or 12.5-μm spacer with a partition to yield a compartment for sample and a compartment for solvent. Approximately 5-μL aliquots of both sample and solvent were loaded into the cell.

All spectral data were acquired using the software package Win-IR-Pro v2.97 manufactured by Digilab. The spectral range of 600−4200 cm^{-1} was used for protein analysis. For the validation process, single-pass ATR-FT-IR protein data was compared to transmission FT-IR protein data where D_2O was used as the solvent. When the accuracy of the single-pass ATR-FT-IR method being validated, a D_2O subtraction was used to allow for the observation of the amide III′ region. When the secondary structure analysis was being performed, protein spectra were acquired by single-pass ATR-FT-IR where H_2O was used as the solvent. Data analysis was performed using the software package Igor-Pro v3.1. There was no need for H_2O subtraction when the single-pass ATR-FT-IR technique was used since the protein sample dehydrated and formed a concentrated gel on the Ge IRE. The only time a H_2O subtraction algorithm was used in this study was to compare the dehydrated protein spectra to those of water-subtracted protein spectra to ensure that no protein denaturation was occurring in the single-pass ATR-FT-IR configuration.

RESULTS AND DISCUSSION

The objective of this study was to validate the single-pass ATR-FT-IR technique for secondary structure determination. Statistical analysis was performed to test the accuracy and reproducibility of the single-pass ATR-FT-IR method for protein analysis. This involved the spectral comparison of proteins in D_2O solution acquired by single-pass ATR-FT-IR to that acquired via transmission FT-IR. Two lots of the protein lysozyme were prepared in D_2O solution and allowed to exchange overnight. All respective replicate spectra were superimposable; thus, a day-to-day study was performed over three consecutive days. It was predicted that the single-pass ATR-FT-IR method would be more reproducible since variations in sample cell path length would not be an issue when this technique was used. As seen in Table 2, the single-pass ATR-FT-IR technique yielded smaller standard deviations for a Gaussian fit of the spectral range from 1600 to 1700 cm^{-1}, which included the amide I′ region. The Gaussian fitting function used was as follows

$$L_j(\omega) = \frac{A_j}{\sqrt{2\pi}\sigma_j} \exp\left\{-\frac{(\omega - \omega_{0j})^2}{2\sigma_j^2}\right\} \qquad (1)$$

Table 2. Comparison of Multi-Gaussian Fits for Six Sets of Lysozyme Spectra

	Single-Pass ATR-FT-IR Microscope			
	lot 1, day 1	lot 1, day 2	lot 1, day 3	SD[a]
χ^2	2.32E−05	5.93E−06	4.54E−06	
Gaussian 1	1603.15	1604.58	1604.58	0.83
Gaussian 2	1641.48	1641.83	1641.83	0.20
Gaussian 3	1654.01	1655.14	1655.14	0.65
	lot 2, day 1	lot 2, day 2	lot 2, day 3	SD
χ^2	4.71E−06	8.02E−06	2.18E−06	
Gaussian 1	1606.22	1605.54	1606.35	0.44
Gaussian 2	1645.48	1647.11	1646.74	0.85
Gaussian 3	1661.78	1665.55	1665.12	2.06
	Transmission FT-IR			
	lot 1, day 1	lot 1, day 2	lot 1, day 3	SD
χ^2	0.0008906	0.0011660	0.0006502	
Gaussian 1	1608.53	1601.44	1605.98	3.59
Gaussian 2	1646.27	1647.07	1646.62	0.40
Gaussian 3	1655.71	1660.15	1661.28	2.94
	lot 2, day 1	lot 2, day 2	lot 2, day 3	SD
χ^2	0.0009040	0.0005801	0.0007299	
Gaussian 1	1602.79	1596.64	1597.73	3.28
Gaussian 2	1650.43	1649.02	1648.69	0.92
Gaussian 3	1655.00	1641.57	1646.21	6.82

[a] SD, standard deviation.

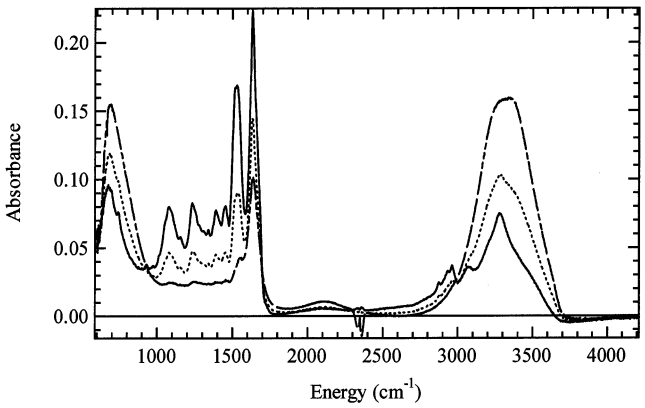

Figure 2. Example of spectral enhancement upon protein gel formation. The spectra shown are of the protein chymotrypsin from a liquid sample (dashed), to an intermediate state (dotted), to a gel state (solid).

where j represents each spectral component. A_j represents the amplitude, ω_{0j} the frequency, and σ_j the variance of the jth Gaussian. A significant reduction in χ^2 was achieved when a three-Gaussian fit was used; thus, three frequencies are reported for each sample spectrum. The results given in Table 2 also demonstrate that the single-pass ATR-FT-IR technique yields amide I′ positions comparable to those protein spectra acquired by transmission FT-IR in D_2O solution. The results presented in Table 2 are representative in that several proteins were analyzed and all yielded standard deviations less than those obtained from transmission FT-IR and had amide I′ positions comparable to those acquired by transmission FT-IR (See Supporting Information Figures 4 and 5.).

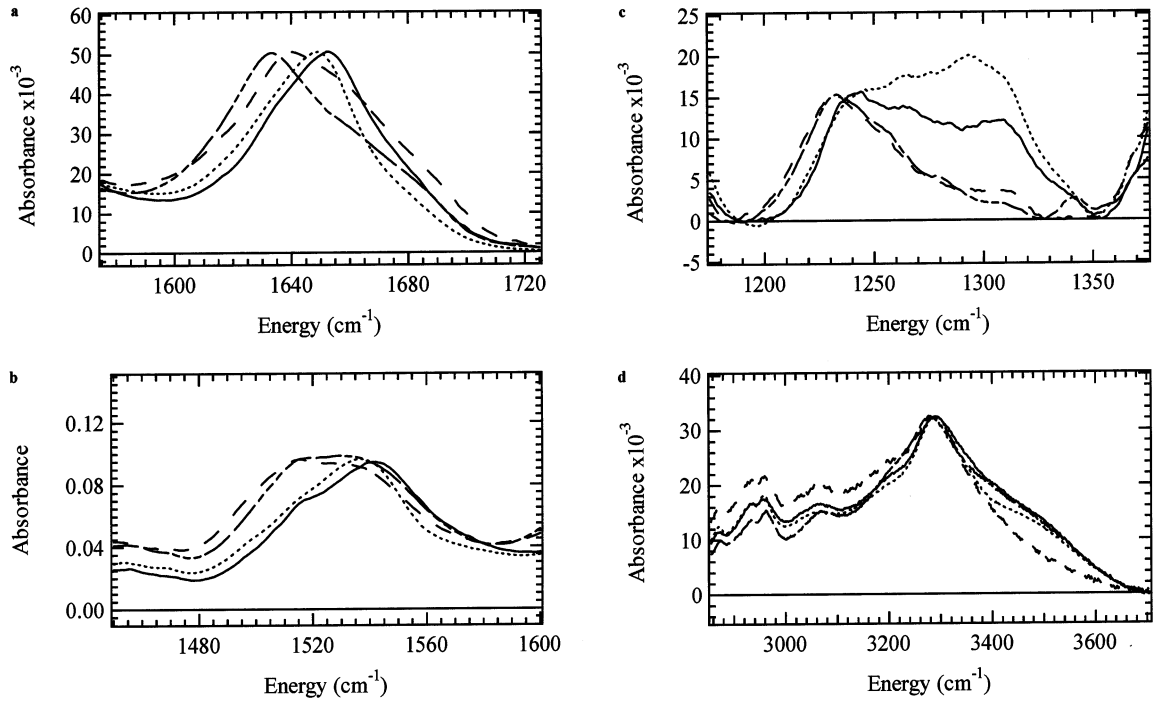

Figure 3. (a) An enlargement of the amide I region, (b) an enlargement of the amide II region, (c) an enlargement of the amide III region, and (d) the enlarged amide A and B regions of four protein spectra, where myoglobin is represented by the dotted spectrum, cytochrome *c* is represented by the solid spectrum, ribonuclease A is represented by the dashed spectrum (– –), and chymotrypsin is represented by the second dashed spectrum (– – – –).

To further validate the accuracy of the single-pass ATR-FT-IR method, the mid-infrared spectra of a representative set of proteins in H_2O solution were acquired using transmission FT-IR, multipass ATR-FT-IR, and single-pass ATR-FT-IR (example spectra are given in Figure 6 of the Supporting Information). This study showed that the amide II regions were comparable for the three techniques. Even with a 6-μm spacer, the transmission FT-IR spectra had enough H_2O interference to obscure the amide I region. The amide I regions in the multipass and single-pass ATR-FT-IR spectra were similar. The most noticeable difference is that the single-pass ATR-FT-IR method allows for the simultaneous observation of amides I–III, A, and B. Even with H_2O subtraction, the multipass ATR-FT-IR spectra do not have as distinctive amide A and B regions as the single-pass ATR-FT-IR spectra.

The trend for amide I is that β-sheet structures have a maximum near 1633 cm^{-1} with a shoulder at 1685 cm^{-1} and α-helical structures have a maximum near 1650 cm^{-1}.[22] Structures with a mixture of β-sheet and α-helical structure exhibit linear combinations of these two basic spectral forms. Myoglobin and cytochrome *c* are primarily α-helical in structure and ribonuclease A and chymotrypsin have significant amounts of β-structure. As seen in Figure 3a, these band shapes were in agreement with previous studies.[6,13,21–23]

Surprisingly, from observation of all the amide bands in the ATR-FT-IR spectra, the amide II and amide III band shapes showed a significant dependence on secondary structure as well. Many investigators maintain that the amide II region is considerably less significant than amide I in distinguishing secondary structure.[14,22,24] Thus, reported trends for the amide II region are limited. Enlarged amide II bands are displayed in Figure 3b. The amide II line shape appears to show clear differences between the two primarily α-helical proteins and the primarily β-sheet proteins. The reported trend is that a strong amide II component occurs in the region of 1540–1550 cm^{-1} and a weak component occurs in the region of 1510–1525 cm^{-1}.[22] The spectrum for cytochrome *c* followed this trend, and the spectrum of myoglobin was very similar. The strong components for these primarily α-helical proteins were at 1541 cm^{-1} for cytochrome *c* and at 1536 cm^{-1} for myoglobin. Ribonuclease A and chymotrypsin amide II bands have strong components in the region from 1530 to 1540 cm^{-1}. The weak component predicted in the region of 1510–1525 $cm^{-1}$was present for the four amide II bands of all of the proteins studied. However, for the primarily β-sheet proteins, the component in the region of 1510–1525 cm^{-1} appeared just as strong as the component in the region of 1530–1540 cm^{-1}. Amide III bands for the same four proteins (Figure 3c) exhibited substantial changes in line shape as a function of varying secondary structure as well (refer to Figure 3c). The amide III band is usually observed by Raman spectroscopy since the amide III band is so weak in the infrared. As with the amide II region, proteins with high β-sheet content yielded strong bands at a lower frequency than proteins with high α-helical content. This is in agreement with the trend that the following secondary structure motifs; turns, β-sheet, α-helix, and random coils yield bands in respective order in the amide III region.[24]

Examination of Figure 3d revealed trends in the amide A and B bands that correspond to protein secondary structure. The

(22) Pelton, J. T.; McLean, L. R. *Anal. Biochem.* **2000**, *277*, 167–176.

(23) Parker, F. S. *Applications of Infrared Spectroscopy in Biochemistry, Biology, and Medicine*; Plenum Press: New York, 1971.

(24) Vedantham, G.; Sparks, H. G.; Sane, S. U.; Tzannis, S.; Przybycien, T. M. *Anal. Biochem.* **2000**, *285*, 33–49.

pattern for amide A was similar to that observed in amide I (Figure 3a) where bands for cytochrome *c*, myoglobin, ribonuclease A, and chymotrypsin occur in order of decreasing frequencies. The lowest frequency (3277 cm^{-1}) corresponded to the β-sheet structure of chymotrypsin. The intermediate frequency of ribonuclease A (3279 cm^{-1}) reflects a mixed content (21% α-helix, 34.7% β-sheet). Myoglobin and cytochrome *c* had maximum frequencies of 3292 and 3293 cm^{-1}, respectively. No correlations were found in the literature for the amide B region. The only noticeable trend in this study was that the amide B bands are broader for cytochrome *c* and myoglobin, the primarily α-helical proteins, than those for ribonuclease A and chymotrypsin, the primarily β-sheet proteins.

CONCLUSIONS

Single-pass attenuated total reflection Fourier transform infrared microscopy provides a novel sample geometry that has distinct advantages over current technology for obtaining protein spectra. Statistical analysis indicates that the single-pass ATR-FT-IR method is more reproducible than transmission FT-IR spectroscopy since the path length in ATR is fixed (albeit wavelength dependent) while the path length in transmission geometry is dependent upon the spacer used. Moreover, experiments performed in H_2O solution will be more reproducible than those that require D_2O exchange since the extent of the exchange can vary from sample to sample. The spectral range for observation of protein spectra shown here is greater than any previous FT-IR study. This is the first report of simultaneous observation of amides A, B, I, II, and III demonstrating a correlation in each band. The increase in information content due to the observation of multiple bands requires further study for application to protein secondary structure prediction that is discussed in the second paper in this series. The enhancement of spectral information from using this technique has already proven useful for the study of protein and peptide conformation.[25]

SUPPORTING INFORMATION AVAILABLE

A listing of figures showing the validity and reproducibility of the method as well as information supporting the claim that protein denaturation can be detected and that water subtraction is not necessarily required for application of the technique. This material is available free of charge via the Internet at http://pubs.acs.org.

Received for review February 15, 2002. Accepted June 3, 2002.

AC020103V

(25) Pop, C.; Chen, Y.; Smith, B.; Bose, K.; Bobay, B.; Tripathy, A.; Franzen, S.; Clark, C. *Biochemistry* **2001**, *40*, 14224−14235.