ISSN 1672-9145 Acta Biochimica et Biophysica Sinica 2007, 39(8): 549–559 CN 31-1940/Q

Minireview

Fourier Transform Infrared Spectroscopic Analysis of Protein Secondary Structures

Jilie KONG and Shaoning YU*

Department of Chemistry, Fudan University, Shanghai 200433, China

Abstract Infrared spectroscopy is one of the oldest and well established experimental techniques for the analysis of secondary structure of polypeptides and proteins. It is convenient, non-destructive, requires less sample preparation, and can be used under a wide variety of conditions. This review introduces the recent developments in Fourier transform infrared (FTIR) spectroscopy technique and its applications to protein structural studies. The experimental skills, data analysis, and correlations between the FTIR spectroscopic bands and protein secondary structure components are discussed. The applications of FTIR to the secondary structure analysis, conformational changes, structural dynamics and stability studies of proteins are also discussed.

Keywords FTIR; protein structure; protein dynamic

Infrared (IR) spectroscopy is one of the oldest and well established experimental techniques for the analysis of secondary structure of polypeptides and proteins [1–5]. The use of stable and powerful laser has led to the development of the Fourier transform (FT) method for IR data acquisition and reliable digital subtraction. The availability of modern computers has enabled the rapid and powerful FTIR data processing and conversion. FTIR spectroscopy is recognized as a valuable tool for the examination of protein conformation in H_2O-based solution, as well as in deuterated forms and dried states, resulting in a greatly expanded use in studies of protein secondary structure and protein dynamics in the past decade [4–17]. Although X-ray crystallography provides the most detailed information concerning positions of individual atoms in the protein structure, it is not, however, possible for all proteins to form a quality crystal for such analysis. In addition, the crystallographic data of a protein cannot be easily extrapolated to the dynamic properties of the proteins in solutions. Nuclear magnetic resonance spectroscopy can be an alternative to X-ray crystallography in solution, but the interpretation of nuclear magnetic resonance spectra of a large protein is a very cumbersome process [18]. Thus, the vibrational spectroscopies, such as FTIR and circular dichroism (CD), are still important and commonly used techniques for protein structure and dynamics studies.

FTIR spectroscopy is a measurement of wavelength and intensity of the absorption of IR radiation by a sample. The IR spectral data of high polymers are usually interpreted in terms of the vibrations of a structural repeat unit [2,3,19]. The polypeptide and protein repeat units give rise to nine characteristic IR absorption bands, namely, amide A, B, and I–VII. Of these, the amide I and II bands are the two most prominent vibrational bands of the protein backbone [3–5]. The most sensitive spectral region to the protein secondary structural components is the amide I band (1700–1600 cm^{-1}), which is due almost entirely to the C=O stretch vibrations of the peptide linkages (approximately 80%). The frequencies of the amide I band components are found to be correlated closely to the each secondary structural element of the proteins. The amide II band, in contrast, derives mainly from in-plane NH bending (40–60% of the potential energy) and from the CN stretching vibration (18–40%) [3], showing much less protein conformational sensitivity than its amide I coun-

Received: March 15, 2007 Accepted: April 29, 2007

This work was supported by the grants from the National Natural Science Foundation of China (No. 20745001 and 20525519), the National Basic Research Program of China (2007CB914304), and the Educational Foundation of Fudan University (EYH1615002)

* Corresponding author: Tel, 86-21-55664974; Fax, 86-21-65641740; E-mail, yushaoning@fudan.edu.cn

DOI: 10.1111/j.1745-7270.2007.00320.x

©Institute of Biochemistry and Cell Biology, SIBS, CAS

Downloaded from https://academic.oup.com/abbs/article/39/8/549/420 by guest on 19 April 2024

terpart [3]. Other amide vibrational bands are very complex depending on the details of the force field, the nature of side chains and hydrogen bonding, which therefore are of little practical use in the protein conformational studies. Banker described the nine amide vibration modes and some standard conformations in detail in his review [20]. The characteristic IR bands of the proteins and peptides are listed in **Table 1**.

Table 1 Characteristic infrared bands of peptide linkage

Designation	Approximate frequency (cm^{-1})	Description
Amide A	3300	NH stretching
Amide B	3100	NH stretching
Amide I	1600–1690	C=O stretching
Amide II	1480–1575	CN stretching, NH bending
Amide III	1229–1301	CN stretching, NH bending
Amide IV	625–767	OCN bending
Amide V	640–800	Out-of-plane NH bending
Amide VI	537–606	Out-of-plane C=O bending
Amide VII	200	Skeletal torsion

Data are from Elliott and Ambrose [2], Krimm and Bandekar [3], Banker [20] and Miyazawa *et al.* [21].

In 1980's, the practical use of FTIR was severely limited by factors such as low sensitivity of the instrument, interfering absorption from aqueous solvent, lack of understanding of the correlations between specific backbone folding types and individual component bands. It was considered back then that the IR spectrum was difficult, if not impossible, to be used in an aqueous solution unless deuterium oxide was used as a solvent [7], because water absorbs strongly in the most important spectral region at approximately 1640 cm^{-1}. Even in D_2O solution, usually only qualitative information was obtained because the components of absorption bands associated with specific substructures, such as α-helix and β-sheet, can not be resolved. Later, as more protein structures were solved by X-ray crystallography, and the computational procedures for the resolution enhancement of broad IR bands of protein were developed, the IR spectra of polypeptides and proteins in D_2O solution could be assigned [4–7]. Later the amide I band assignments for proteins in H_2O solution were also made [8–11]. The computerized FTIR instrumentation has improved the signal-to-noise ratio and allowed extensive data manipulation. The new band-narrowing methods, Fourier self-deconvolution (FSD) and second derivative, have not only enriched the qualitative interpretation of the IR spectra, but also provided a basis for the quantitative estimation of protein secondary structure [17]. Now, for most globular proteins, people can acquire IR spectra quantitatively under certain conditions and use the FTIR technique to study the structure and function of proteins and peptides [11–17].

The singular advantages of FTIR over other techniques are that spectra can be obtained for proteins in a wide range of environments, requiring less time and sample, and direct correlations between the IR amide I band frequencies and the secondary structure components can be found. The goal of this paper is to review the recent developments in FTIR technique and its applications to protein secondary structure analysis and conformational study. Bypassing the basic theory of molecule vibration, we will focus on the experimental skills, data analysis, IR band assignment and applications to protein structural study. In this work, all spectra presented were measured by BOMEM FTIR spectrophotometer (ABB Bomen Inc., St-Laurent, Canada) and Beckman FH-01 IR cell (Beckman Coulter Inc., Fullerton, USA) with CaF_2 windows. The cell path-length is 7.5 μm for H_2O and 50 μm for D_2O solution. BOMEM GRAMS/32 software (ABB Bomen Inc.) was used for data acquisition and analysis.

Protein FTIR Data Analysis and Band Assignment

Data analysis

High sensitivity to small variations in molecular geometry and hydrogen bonding patterns makes the amide I band uniquely useful for the analysis of protein secondary structural composition and conformational changes [4,11]. In the amide I region (1700–1600 cm^{-1}), each type of secondary structure gives rise to a somewhat different C=O stretching frequency due to unique molecular geometry and hydrogen bonding pattern. However, other than the distinctive absorbance maxima, the observed amide I bands of proteins are usually featureless, due to the extensive overlap of the broad underlying component bands, which lie in close proximity to one another and are instrumentally unresolvable. Thus, mathematical methods such as resolution-enhancement technique are necessary to resolve the individual band component corresponding to specific secondary structure. **Fig. 1(A)** is the IR spectrum of cAMP receptor protein (CRP) in buffer. The individual underlying

©Institute of Biochemistry and Cell Biology, SIBS, CAS

Downloaded from https://academic.oup.com/abbs/article/39/8/549/420 by guest on 19 April 2024

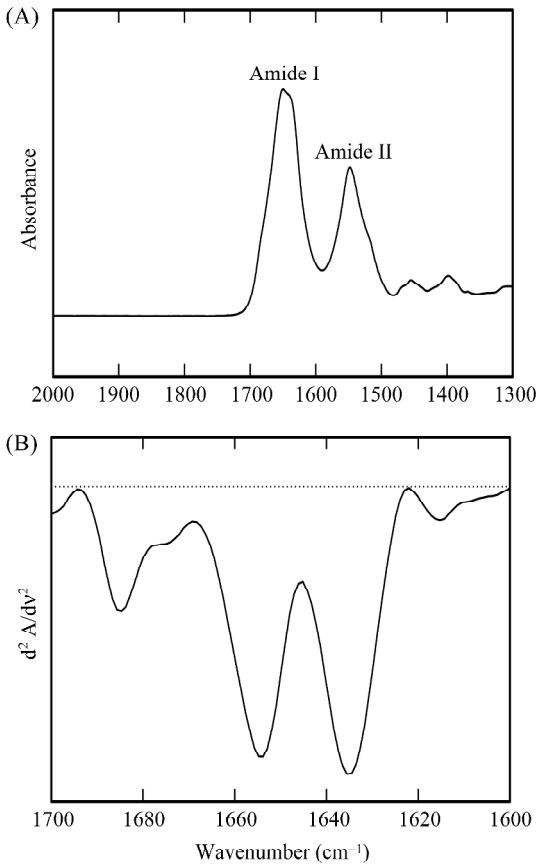

Fig. 1 Resolution enhancement of a complex band contour of overlapping infrared bands

(A) Original spectrum. (B) Second derivative spectrum.

components can not be visualized without resolution enhancement.

Mathematical data analysis methods can be used to "enhance" the resolution of the protein spectrum, allowing the intrinsically broad components to be narrowed and separated beyond the instrument resolution [1,22]. The mathematical band-narrowing process does not actually increase the instrumental resolution, but rather increases the degree of separation by narrowing the half-bandwidth of individual components for easier visualization. This band-narrowing process is achieved at the expense of spectral quality of the original band, which leads to a degradation of signal-to-noise ratio. Several methods have been developed to estimate quantitatively the relative contributions of different types of secondary structures in proteins from their IR amide I spectra in solution, including FSD-curve fitting [7,22], second derivative analysis [4,7,11], partial least-squares analysis [23], and data basis analysis [24]. The FSD-curve fitting and second derivative analysis are the two most popularly used methods.

The theoretical backgrounds of IR data handling techniques, including FSD and second derivative, have been discussed in detail by Susi and Byler [4,6]. The FSD method was also reviewed by Surewicz and Mantsch [18]. The key meaningfulness of the FSD method is to select the conditions that achieve the maximum band narrowing while keeping the increase in noise and the appearance of side-lobes at minimum [6,14,22]. This method is based on the assumption that a spectrum of single bands (each narrow band is characteristic of a secondary structure) is broadened in the liquid or solid state. Therefore, the bands overlap and can not be distinguished within the amide envelope. An alternative approach to "enhance" the resolution of overlapping IR bands is based on the generation of *n*th order derivative band files. This can be carried out in the frequency (wavenumber) domain of the spectrum. Second derivative spectra allow the identification of various secondary structures present in the protein [6]. Most of the peak positions are easily found in the second derivative spectra. An improved method for carrying out second derivative analysis has established the utility of the method for obtaining quantitative as well as qualitative determination of α-helix, β-sheet, random and turn structures [11]. A curve fitting procedure can be applied to calculate quantitatively the area of each component representing a type of secondary structure [4,8–11]. **Fig. 1(B)** shows the second derivative spectrum of CRP, which was obtained according to the methods of Dong *et al.* [11]. Both FSD and second derivative methods were included in BOMEM GRAMS/32 software. Second derivative analysis was used in this work. Manipulation procedures were carried out in accordance with Dong *et al.* [11].

The overlap of secondary structural components is significant in the amide I region, even after mathematical resolution enhancement. Some investigators considered that the valley between two adjacent peaks of equal intensity must be 20% lower than the peak tops that could be resolved [1,25].

Band assignment

Quantitative analysis of protein secondary structure is based on the assumption that protein can be considered as a linear sum of a few fundamental secondary structural elements. Comparisons of IR spectra with high-resolution X-ray crystal structures of proteins could establish necessary spectra-structure correlations. Over the years, many correlations between IR spectra and particular protein

Downloaded from https://academic.oup.com/abbs/article/39/8/549/420 by guest on 19 April 2024

structure have been established. The amide I band components can be assigned by studying their frequency behavior in which the protein secondary structure is known by other techniques. As mentioned above, the amide I band (1700–1600 cm^{-1}) is due mainly to the C=O stretching vibration (approximately 80%) of the amide groups coupled with little in-plane NH bending (<20%) [3]. The extract frequency of this vibration band depends on the nature of hydrogen bonding involving the C=O and NH moieties [3]. In turn, this is determined by the secondary structure adopted by the polypeptide chain, reflecting the backbone conformation and hydrogen-bonding pattern. Strictly speaking, the observed amide I band contours of proteins or polypeptides consist of overlapping component bands, representing α-helices, β-sheets, turns and random structures. It is needed to establish a correspondence between IR spectra and the various types of protein secondary structure. Assignments of the amide I band component to each secondary structure element are available for proteins in both D_2O and H_2O media [4–16].

In D_2O solution, it has been revealed that the broad protein amide I band contours can be decomposed into a number of components. A component centered between approximately 1658 and 1650 cm^{-1} has been assigned to the α-helix, which is consistent with both theoretical calculation [3] and the observation of bands in the spectra of α-helical proteins [4,26]. Bands near 1663 cm^{-1} are assigned to 3_{10} helices [3,4,11], although this structure is rarely found in proteins. More than one β-component has been observed in the spectra of many β-sheet proteins. Bands in the regions of 1640–1620 cm^{-1} and 1695–1690 cm^{-1} have been assigned to β-sheet by many authors [3,4, 23]. Theoretical calculation of β-sheets also predicts an IR active mode between approximately 1695 and 1670 cm^{-1} [3]. These β-components are often complicated by the presence of more than one band above 1670 cm^{-1}. The assignment of bands around 1670, 1683, 1688 and 1694 cm^{-1} to β-turns has been proposed [3]. Turns are also associated with a characteristic band around 1665 cm^{-1}. The unordered conformation (usually referred to as random coil) is usually associated with the IR band between 1640 and 1648 cm^{-1} [3]. The distinguishing feature of random coil is that it is non-repetitive. Such assignment is supported by the position of a prominent band in the spectra of apparently order-less proteins [7].

H_2O as a solvent is much more preferable than D_2O for studying protein structure [8–11]. D_2O changes the protein properties somewhat in comparison with the native ones. In H_2O solution, the bands between 1654 and 1658 cm^{-1} are assigned to α-helix, which is supported by human hemoglobin A and bovine Myoglobin proteins, and are expected for proteins with α-helical structure [11]. The bands between 1642 and1624 cm^{-1} are assigned to β-sheets components through the IR spectra of immunoglobulin G and concanavalin A, which contain more than 60–70% β-sheet structures and almost no α-helix. In globular protein, it has been observed that approximately 30% of amino acid residues reside in β-turn conformations [11,28]. The bands located at 1688, 1680, 1672, and 1666 cm^{-1} are assigned to β-turn structures. The characteristic band for random coil conformation can be assigned to the band located at 1648±2 cm^{-1}.

Some minor or rare structures might interfere with the band assignments discussed above. For example, the β-turn band at approximately 1665 cm^{-1} is near the characteristic IR band representing 3_{10}-helices. Vibrations of some amino acid side chains might make small contributions to the intensity of characteristic protein amide bands. In addition, the experimental procedure might also bring spectral error. All these complications indicate that there is no simple correlation between the IR spectra and secondary structural components. Caution has to be exercised in the interpretation of IR spectra of proteins. **Table 2** shows the deconvoluted amide I band frequencies and assignments to secondary structure for proteins in D_2O and H_2O.

Subtraction of background

The accuracy of subtraction of large H_2O bands is

Table 2 Deconvoluted amide I band frequencies and assignments to secondary structure for protein in D_2O and H_2O media

H_2O†		D_2O‡	
Mean frequencies	Assignment	Mean frequencies	Assignment
1624±1.0	β-sheet	1624±4.0	β-sheet
1627±2.0	β-sheet		
1633±2.0	β-sheet	1631±3.0	β-sheet
1638±2.0	β-sheet	1637±3.0	β-sheet
1642±1.0	β-sheet	1641±2.0	3_{10} Helix
1648±2.0	Random	1645±4.0	Random
1656±2.0	α Helix	1653±4.0	α-Helix
1663±3.0	3_{10} Helix	1663±4.0	β-Turn
1667±1.0	β-Turn	1671±3.0	β-Turn
1675±1.0	β-Turn	1675±5.0	β-sheet
1680±2.0	β-Turn	1683±2.0	β-Turn
1685±2.0	β-Turn	1689±2.0	β-Turn
1691±2.0	β-sheet	1694±2.0	β-Turn
1696±2.0	β-sheet		

† Data are from Dong *et al*. [12,13]; ‡ Data from Susi *et al*. [4,7].

©Institute of Biochemistry and Cell Biology, SIBS, CAS

Downloaded from https://academic.oup.com/abbs/article/39/8/549/420 by guest on 19 April 2024

always a concern. H_2O has strong IR absorbance with three prominent bands around 3400 (O-H stretching), 2125 (water association), and 1645 cm^{-1} (H-O-H bending). The amide I vibration for proteins absorbs between 1600 and 1700 cm^{-1}, overlapping directly with the H_2O bending vibrational band at 1645 cm^{-1}. The intensity of the water absorbance at 1645 cm^{-1} is approximately an order of magnitude higher than the amide I absorbance of proteins. For IR study of protein in H_2O solution, water absorption in the 1600–1700 cm^{-1} region might be the biggest problem. It is much easier if the spectroscopic study is carried out in D_2O solution because there is no absorption spectrum of D_2O in the region where the amide I and II bands are observed. A large path-length (i.e. 50 μm) IR cell can be used and results in a higher signal-to-noise ratio, a factor of particular importance for proteins of low solubility. However, the amide I band frequencies are strongly affected by the H-D exchanges in the peptide linkages [4, 7]. The effect of these exchanges on protein structural properties is not fully understood, especially under incomplete H-D exchange conditions. Furthermore, because the exchange of D for H can affect the strength and length of hydrogen bonds, it is possible that protein secondary structures might be altered by the replacement of H_2O by D_2O. Therefore, H_2O-based media have the advantage of providing a more native environment.

The greater sensitivity of the FTIR instrument and new H_2O subtraction program make it possible to obtain a good protein FTIR spectrum in H_2O solution [23,29,30]. The problem of water absorption can also be limited by using an IR cell of sufficiently small path-length (6–10 μm) to permit IR radiation passing through the material under observation. By using the same cell to record both the reference spectrum and the spectrum of protein solution under identical scan conditions, the aqueous water contribution can be removed from the spectrum of the protein solution by digital subtraction. Furthermore, progress in the development of methods for spectral data analysis makes it easier to distinguish the individual components within the intrinsically overlapped amide I band contours [11,12]. Nevertheless, the accurate measurement of frequency and intensity of the amide bands is elusive [12]. The short cell path-length makes it more difficult to match the path-lengths of sample and reference cells [12,31]. Typically, the same cell was used to record the spectra of both sample and reference, although the IR cell drying and reloading steps might slightly alter the cell path-length. In order to get a successful subtraction of absorption bands due to liquid water and gaseous water in the atmosphere, one must have a criterion for determining whether or not the absorption by water is correctly compensated, that is, the region of the spectrum where no absorption by the sample but absorption by water is present [31]. Two criteria have been established to judge whether the spectrum is good or not [12]. First, the bands originating from water vapor must be subtracted accurately from the protein spectrum between 1800 and 1500 cm^{-1}. Second, a straight baseline must be obtained from 2000 to 1750 cm^{-1}. Many investigators have used a straight baseline between 2000 and 1750 cm^{-1} as the standard for judging the success of water subtraction to obtain protein spectra [12,32,33]. By using two criteria, the average experimental errors of the spectra at the amide I and II band maxima could be less than 3% and 1.5%, respectively [8].

Empirically, in order to obtain a high quality protein spectrum, high protein concentration (>10 mg/ml in H_2O) [34] and small cell path-length (6–10 μm) are needed. The identical scan conditions are used to record buffer reference spectrum and the spectrum of protein solution. In addition, water vapor, as well as water (or buffer) itself, must be subtracted accurately from the protein spectrum. As mentioned above, using a straight base line between 2000 and 1750 cm^{-1} as the standard to judge the successfulness of water subtraction would lead to a higher quality of protein spectra.

Application of FTIR to Protein Secondary Structure

FTIR spectroscopy has been used to study the secondary structure composition, structural dynamics, conformational changes (effects of ligand binding, temperature, pH and pressure), structural stability and aggregation of proteins.

Estimation of protein secondary structure

Proteins are frequently referred to as having a certain fraction of structural components (α-helix, β-sheet, etc.). The secondary structural composition is some of the most important information for a structure-unknown protein. Therefore estimation of protein secondary structure is one of the major applications of the FTIR technique.

The quantitative estimation of protein secondary structure is based on the assumption that any protein can be considered as a linear sum of a few fundamental secondary structural elements, and the percentage of each element is only related to the spectral intensity (the molar absorptivity of C=O stretching vibration for each secondary structural element is essentially same). To analyze the amide I band

Downloaded from https://academic.oup.com/abbs/article/39/8/549/420 by guest on 19 April 2024

component, FSD or second derivative spectra need to be curve fitted. The areas of FSD or second derivative spectra correspond to different types of secondary structure components. It is considered that the accuracy of measured band areas in the second derivative amide I spectra depends upon the correct positioning of the baseline [11]. For the globular protein, the correlation coefficient between IR and X-ray estimates of α-helices, β-sheets, β-turns and remainder was 0.98, 0.99, 0.90 and 0.92, respectively [8]. Dong and colleagues reported the distribution of secondary structures determined from the amide I spectra of globular proteins in aqueous solutions, which was nearly identical to the amount computed from crystallographic data [12].

Fig. 2 shows the analysis procedure of second derivative FTIR spectra of CRP. The percentage of α-helices was 41, calculated by the relative area at 1653 cm^{-1}. The percentage of β-sheets was 36, calculated by the relative area near 1637 cm^{-1}. The percentage of β-turns was 18, calculated by adding the areas of all β-turn bands between 1670 and 1690 cm^{-1}. The band area at 1648 cm^{-1} was assigned to random coil. This data is consistent with the X-ray crystallography data. **Table 3** lists some results calculated from the FTIR method compared with the X-ray crystallographic method, including the data of CRP and pyruvate kinase. From these data, the results from FTIR and X-ray crystallography methods match well.

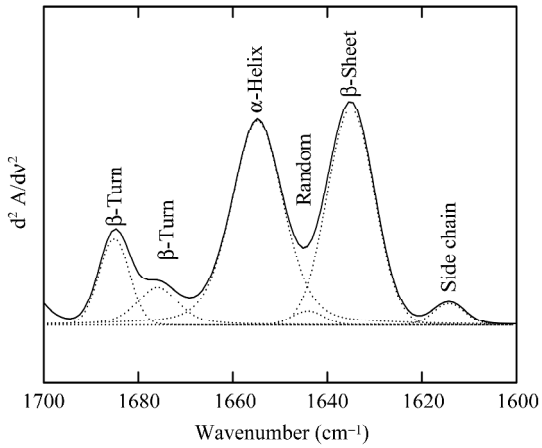

Fig. 2 Second derivative infrared spectrum of cAMP receptor protein in Tris buffer

Curve-fitted second derivative spectrum was carried out by factoring–1. The curve fitting was carried out by BOMEM GRAMS/32 software.

Kinetics of H-D exchange related to conformation

Conformation and structural dynamics are essential to proteins for carrying out their proper biological function. The structural dynamics of a given conformation is likely to influence the activity of the protein. H-D exchange has been used extensively in studies of structural dynamics of

Table 3 Comparison of protein secondary structures determined by Fourier transform infrared (FTIR) spectroscopy-self-deconvolution and X-ray crystallography †

Protein	Secondary structure (%)							
	α-Helix		β-Sheet		β-Turn		Random	
	FTIR	X-ray	FTIR	X-ray	FTIR	X-ray	FTIR	X-ray
Hemoglobin	78	87	12	0	10	7	0	8
Myoglobin	85	85	7	0	8	8	0	7
Lysozyme	40	45	19	19	27	23	14	13
Cytochrome c (oxidized)	42	48	21	10	25	17	12	25
α-Chymotrypsin	9	8	47	50	30	27	14	15
Trypsin	9	9	44	56	38	24	9	11
Ribonuclease A	15	23	40	46	36	21	9	10
Alcohol dehydrogenase	18	29	45	40	23	19	14	12
Concanavalin A	8	3	58	60	26	22	8	15
Immunoglobulin G	3	3	64	67	28	18	5	12
β_2-Microglobulin	6	0	52	48	33	–	9	–
cAMP receptor protein	41	37	36	38	20	19	3	6
Pyruvate kinase	40	41	37	33	13	14	10	12

† X-ray data are from Levitt and Greer [35] and FTIR data are from Dong *et al.* [12].

©Institute of Biochemistry and Cell Biology, SIBS, CAS

Downloaded from https://academic.oup.com/abbs/article/39/8/549/420 by guest on 19 April 2024

proteins [36–38]. It has been established that the rates at which the amide proton exchange with solvent deuterium reflect the structural dynamics of proteins [39] and they are sensitive to the secondary structural composition and experimental conditions such as pH, temperature and pressure [36–39]. At constant experimental conditions, a more rapid rate of exchange implies a greater flexibility and motion in the structural region of the exchange.

FTIR spectroscopy has been used in connection with H-D exchange in polypeptides and proteins. The intensity changes of the amide I and II bands and the intensity change of the predominant secondary structural elements (α-helix and β-sheet) can be determined. It is suggested that it is more convenient to base H-D exchange investigations on apparent intensity changes of the amide II band [4], because it was not adversely interfered by the absorption bands of H_2O, HOD, or D_2O. Barksdale and Rosenburg suggested that the protein amide H-D exchange ratio could be presented as the fraction of unexchanged amide proton [40]. A calculation equation was established and used by some investigators [16,41]. **Fig. 3** shows the H-D exchange ratio of pyruvate kinase in Tris buffer as a function of time, indicated by both original and second derivative FTIR spectra.

Most IR structural dynamics studies are focused on the H-D exchange of proteins on a global scale [16,36–38]. One may also focus on the exchange rate of secondary structure components. As a successful example, Dong *et al*. focused on the exchange rate of CRP secondary structural components, and obtained some useful information. CRP is a dimer with two functional domains in each subunit. There is a bias in the distribution of secondary structural elements between the two functional domains, namely extensive β-sheets in the cAMP binding domain and predominant α-helices in the DNA binding domain [42,43]. They took advantage of this feature and facilitated the interpretation of the ligand-induced conformational and structural dynamics changes [16].

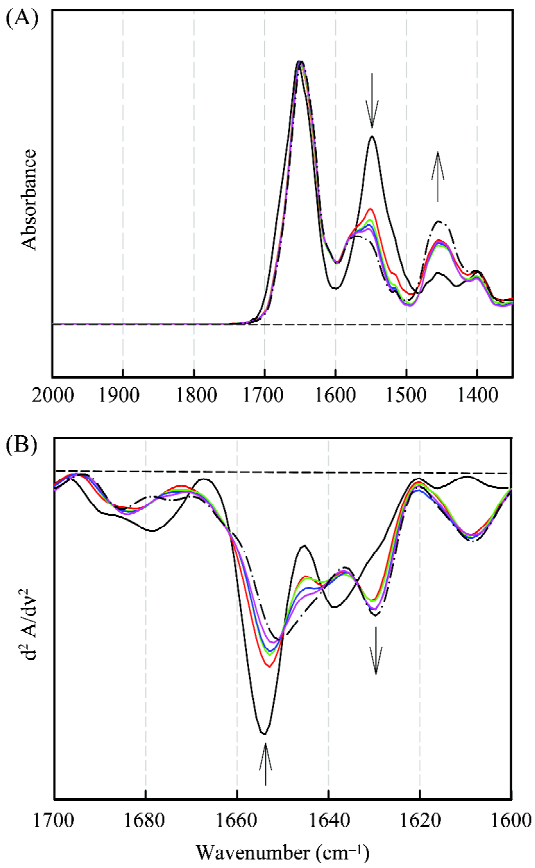

Fig. 3 Original spectra (A) and second derivative amide I spectra (B) of pyruvate kinase (PK) in Tris buffer

H-D exchange as a function of time: spectra of PK in H_2O solution (solid black); H-D exchange 1 min (red); H-D exchange 10 min (green); H-D exchange 30 min (blue); H-D exchange 60 min (Pink); in D_2O solution (dashed black). The arrows indicate the directions in which the intensities of the bands are shifting with time.

Protein stability

Generally, protein stability studies can provide information on the folding/unfolding and structural stability of protein molecules. The advantage of FTIR amide I spectroscopy over other techniques is that the IR method can, in principle, monitor the folding/unfolding process of all parts of the protein simultaneously [44]. However, FTIR amide I spectroscopy has not been widely used as a method of choice in studies of protein folding. The reason is that obtaining good IR spectra of proteins in the presence of chemical denaturant, which must be used at high concentration to ensure protein unfolding, is not simple and has had only limited success [15]. First, IR cell with short path-lengths is required for preventing a saturation of IR detector by the absorption of chemical denaturant (Hg/Te/Cd detector saturates near 4.0 M GdnHCl at the cell path-length used), which currently prevents application of this method to proteins that are fairly resistant to GdnHCl denaturation [15,45]. Second, the major IR bands of denaturant (urea or GdnHCl) mask weaker amide I bands. Previous works on GdnHCl-induced denaturation of proteins have shown that accurate subtraction of GdnHCl absorption is difficult above 1660 cm^{-1} [46]. ^{13}C urea in H_2O shifts the urea IR absorption to 1624 cm^{-1}, leading to a similar problem. To circumvent this, one might resort to ^{13}C-labeled urea in D_2O, which shifts the C=O band of urea from 1618 cm^{-1} to 1562 cm^{-1} [44,45], leaving a clear window in the IR spectrum for observing the protein amide

Downloaded from https://academic.oup.com/abbs/article/39/8/549/420 by guest on 19 April 2024

I band above 1600 cm^{-1}. The combination of IR cells with a short path-length and ^{13}C-labeled urea allows us to measure the IR spectra of protein in the presence of high concentrations of chemical denaturant and gain useful information from the conformation-sensitive amide I band of protein [44,45].

The use of a wide range concentration of denaturant GdnHCl in FTIR spectroscopy has been reported. Bowler and colleagues successfully monitored GdnHCl-induced cytochrome *c* denature by FTIR spectroscopy [15]. The first point to be considered regarding the spectra of denatured proteins in the presence of GdnHCl is the reliability of the subtraction procedure. Protein-bound water molecules do not seem to have an altered vibrational spectrum from unbound water molecules [8,9]. The effects of over- and under-subtraction are best evaluated after resolution enhancement by second derivative methods [15]. It is apparent that the major effects of inaccurate subtraction are limited to the wavelength region between 1690 and 1665 cm^{-1} [15]. One important criterion is subtraction of GdnHCl must be done so as to produce a smooth amide I band shape with no discontinuities and to maintain the amide I to amide II peak intensity ratio close to that observed for the native protein. Improper subtraction is evidenced by a doubled peak in the amide I region or an amide I peak with a concave discontinuity [15]. When the amount of amide I and GdnHCl intensity becomes comparable during the subtraction, the spectrum shows two peaks, allowing the GdnHCl intensity to be removed with a good degree of certainty. Although one cannot be certain of the absolute accuracy of the denaturant subtraction, by using consistent methods a high degree of precision in subtraction can be achieved allowing comparison of differences in the spectra of denatured states of closely related proteins, even in the region between 1690 and 1665 cm^{-1} [15]. The instrument saturates near 4.0 M GdnHCl at the cell path-length used, which is another limitation that cannot be neglected. The highest GdnHCl concentration that is actually used should be less than 3.5 M by FTIR spectroscopy. **Fig. 4** shows the subtraction of GdnHCl. In the figure, A represents the absorbance spectrum of CRP with 0.8 M GdnHCl, B represents the absorbance spectrum of 0.8 M GdnHCl, and C represents the absorbance spectrum of CRP after proper subtraction. Compared with **Fig. 1(A)**, this IR absorbance spectrum is smooth and continuous with a normal intensity ratio of the amide I over the amide II band, so the subtraction of 0.8 M GdnHCl is acceptable.

Unlike chemical induced unfolding, thermal induced protein unfolding followed by FTIR spectroscopy can avoid limitations in saturation and subtraction. This technique has been widely used in protein aggregation studies [46,47]. Current theoretical and experimental evidence for protein aggregation suggests that aggregates are formed from partially folded intermediates [48]. Aggregation of proteins is a problem with serious medical implications and economic importance. To develop strategies for preventing protein aggregation, the mechanism and pathways by which protein aggregate must be characterized. In contrast to far-ultraviolet CD, IR spectroscopy is insensitive to light scattering, thus providing a valuable tool for protein aggregation studies. The thermally-induced aggregation processes of the majority of proteins studied by FTIR can be described with a two-state model, the predominant secondary structural element (α-helix or β-sheet) decreases as a function of temperature and is concomitantly replaced by intermolecular β-sheet aggregates [47]. As a successful example, Dong *et al.* adopted FTIR to investigate protein aggregation by a combination of thermal and chemical denaturation, thus providing a means to populate and characterize aggregation intermediates. They suggested that this method is valuable for studying the aggregation processes of a wide range of proteins. The identification and characterization of aggregation intermediates might lead to new interdiction strategies for amyloidogenic human disease, as well as improvements in industrial processing, storage, and delivery of therapeutic proteins.

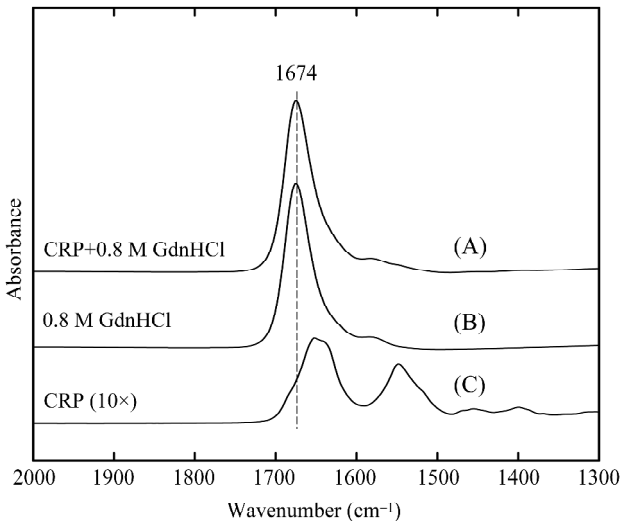

Fig. 4 Subtraction of GdnHCl absorbance

(A) Absorbance spectra of cAMP receptor protein (CRP) in 0.8 M GdnHCl. (B) Absorbance spectra of 0.8 M GdnHCl. (C) Absorbance spectrum of CRP in 0.8 M GdnHCl after subtraction of all background.

Downloaded from https://academic.oup.com/abbs/article/39/8/549/420 by guest on 19 April 2024

©Institute of Biochemistry and Cell Biology, SIBS, CAS

Secondary Structures of Proteins Adsorbed onto Aluminum Hydroxide

Recently, a new IR technique was developed for studying protein structure at low concentration in solution [50]. Adsorbing proteins onto alhydrogel provides a means of obtaining FTIR spectra to study secondary structure and conformational changes of proteins in aqueous solution at very low concentration. This new procedure effectively lowers the concentration requirement for FTIR studies of proteins in aqueous solutions by at least 40-fold, as compared with the conventional FTIR method. This technique permits FTIR study of proteins to be carried out in the same concentration range as those used for CD and fluorescence, making it possible to compare structural information obtained by three commonly used techniques in the biophysical characterization of proteins.

Complexity of FTIR Spectrum: Side-chain Absorption

The FTIR studies of peptides and proteins have made efforts in identifying characteristic frequencies and determining their relations to the structures of protein molecules. The data analysis depended on empirical correlations of the spectra of chemically similar molecules, and occasionally yielded significant insights into the dependence of the spectrum on the conformation of the polypeptide chains [3]. It is very important to recognize that there is no simple correlation between IR spectra and the secondary structure elements of proteins. FTIR spectra are complicated in that every single spectrum has its own characteristics due to different micro-circumstances. Different ambient circumstances, including side-chain absorbance, make it difficult to precisely assign the secondary structure and frequency. The band assignment shown in Table 2 is only empirical. The estimation of secondary structure with the FTIR method should be considered only as a good approximation [8].

The main difficulty in a quantitative study of the protein spectra is the estimation of side-chain absorption, which must be taken into account in the analysis of protein spectra. The contribution of side-chain in the globular protein spectra is 10%–30% of the overall absorption [8]. This absorption is superimposed on the peptide absorption and can somewhat change the amide I band spectral parameters. Some amino acid residues, especially arginine, asparagine, glutamine, aspartic and glutamic acids, lysine, tyrosine, histidine and phenylalanine have intensive absorption in the amide spectral region [8,49]. Venyaminov and Kalnin [8] and Chirgadze *et al*. [49] have established the absorption parameters of these residues in the amide spectral region in H_2O solution (**Table 4**). The quantitative estimation of amino acid side-chain groups permits more refined analysis of the secondary structure of polypeptides and proteins by FTIR, allows the protein spectra to be investigated in detail and achieves a better interpretation of the observed spectral effects. This should be kept in mind when analyzing protein spectra, especially when the content of these residues is high [8].

Summary

FTIR spectroscopy is a well-established experimental technique for studying the secondary structural composition and structural dynamics of proteins. Armed with mathematic resolution enhancement techniques,

Table 4 Spectral parameters for absorbance bands of backbone and side chain in the 1700–1600 cm^{-1} range in H_2O†

Absorbance band assignment		ν_0 (cm^{-1})	ε_0 (M^{-1}cm^{-1})
Asparagine (amide)	C=O ($CONH_2$) stretching	1678±3	310
	NH ($CONH_2$) bending	1622±2	160
Glutamine (amide)	C=O ($CONH_2$) stretching	1670±4	360
	NH ($CONH_2$) bending	1610±4	220
Arginine (guanidine)	C=N ($CN_3H_5^+$) asymmetric stretching	1673±3	420
	C=N ($CN_3H_5^+$) symmetric stretching	1633±3	300
Lysine (amine)	NH^+ (NH_3^+) asymmetric bending	1629±1	130
Tyrosine (phenyl)	Ring vibration	1602±2	160
N-terminal group (amine)	NH^+ (NH_3^+) asymmetric bending	1631±3	210

† ν_0, band position; ε_0, absorptivity of mole residues. Data are from Venyaminov and Kalnin [9].

Downloaded from https://academic.oup.com/abbs/article/39/8/549/420 by guest on 19 April 2024

various methods of FTIR data analysis have been well developed. The correlations between IR spectra and protein secondary structures have been established. The amide I band component assignments to protein secondary structure elements, such as α-helix, β-sheet, β-turn and random structures, are available for proteins in H_2O as well as D_2O media. One could adopt those techniques to study various protein systems for different purposes. Sample backgrounds have to be subtracted accurately from the protein spectrum with strict procedure(s). To obtain high-quality IR spectra, relatively high concentrations of proteins (e.g. >10 mg/ml) and small cell path-length (6–10 μm) are needed for proteins in H_2O solution unless protein was adsorbed onto alhydrogel.

The singular advantage of FTIR over other techniques is convenience. IR spectra can be obtained for proteins in a wide range of environments with a small amount of sample. Other than estimating the content of protein secondary structures, IR can also provide information on protein structural stability and dynamics. The FTIR spectrum is also complex, and some characteristic bands of secondary structure elements might overlap. The background subtraction procedure could also bring experimental error. Estimation of side-chain absorption must be taken into account in the analysis of protein spectra. All the complications indicate that caution has to be exercised in the interpretation of IR spectra of proteins. When estimating the percentage contents of protein secondary structures, a combination of FTIR and CD is recommended for increasing prediction accuracy.

Acknowledgements

The authors appreciate Dr. James C. LEE at the University of Texas Medical Branch (Texas, USA) for suggesting writing this review and Dr. Aichun DONG at the University of Northern Colorado (Colorado, USA) for critical review of the manuscript.

References

1 Smith BC. Fundamentals of Fourier Transform Infrared Spectroscopy. CRC Press: Boca Raton, 1996
2 Elliott A, Ambrose EJ. Structure of synthetic polypeptides. Nature 1950, 165: 921–922
3 Krimm S, Bandekar J. Vibrational spectroscopy and conformation of peptides, polypeptides, and proteins. Adv Protein Chem 1986, 38: 181–364
4 Susi H, Byler DM. Resolution-enhanced fourier transform infrared spectroscopy of enzymes. Methods Enzymol 1986, 130: 290–311
5 Surewicz WK., Mantsch HH. New insight into protein secondary structure from resolution-enhanced infrared spectra. Biochim Biophys Acta 1988, 952: 115–130
6 Susi H, Byler DM. Protein structure by fourier transform infrared spectroscopy: Second derivative spectra. Biochem Biophys Res Comm 1983, 115: 391–397
7 Byler DM, Susi H. Examination of the secondary structure of proteins by deconvolved FTIR spectra. Biopolymer 1986, 25: 469–487
8 Venyaminov SY, Kalnin NN. Quantitative IR spectrophotometry of peptide compounds in water (H_2O) solution. I. Spectral parameters of amino acid residue absorption bands. Biopolymers 1990, 30: 1243–1257
9 Venyaminov SY, Kalnin NN. Quantitative IR spectrophotometry of peptide compounds in water (H_2O) solution. II. Amide absorption bands of polypeptide and fibrous protein in α-, β- and random coil conformations. Biopolymers 1990, 30: 1259–1271
10 Kalnin, Baikalov IA, Venyaminov SY. Quantitative IR spectrophotometry of peptide compounds in water (H_2O) solution. III. Estimation of the protein secondary structure. Biopolymer 1990, 30: 1273–1280
11 Dong A, Huang P, Caughey WS. Protein secondary structure in water from second-derivative amide I infrared spectra. Biochemistry 1990, 29: 3303–3308
12 Dong A, Huang P, Caughey WS. Redox-dependent changes in β-extended chain and turn structures of cytochrome c in water solution determined by second derivative amide I infrared spectra. Biochemistry 1992, 31: 182–189
13 Dong A, Caughey B, Caughey WS, Bhat KS, Coe JE. Secondary structure of the pentraxin female protein in water determined by infrared spectroscopy: Effects of calcium and phosphorylcholine. Biochemistry 1992, 31: 9364–9370
14 Dong A, Caughey WS. Infrared methods for study of hemoglobin reaction and structures. Methods Enzymol 1994, 232: 139–175
15 Bowler BE, May K, Zaragoza T, York P, Dong A, CaugheyWS. Destabilizing effects of replacing a surface lysine of cytochrome c with aromatic amino acids. Biochemistry 1993, 32: 183–190
16 Bowler BE, Dong A, CaugheyWS. Characterization of the guanidine hydrochloride-denatured state of iso-1-cytochrome *c* by infrared spectroscopy. Biochemistry 1994, 33: 2402–2408
17 Dong A, Malecki JM, Lee L, Carpenter JF, Lee JC. Ligand-induced conformational and structural dynamics changes in *Escherichia coli* cyclic AMP receptor protein. Biochemistry 2002, 41: 6660–6667
18 Surewicz WK, Mantsch HH. New insight into protein secondary structure from resolution-enhanced infrared spectra. Biochim Biophys Acta 1988, 952: 115–130
19 Liang CY, Krimm S, Sutherland GBBM. Infrared spectra of high polymers, I. Experimental methods and general theory. J Chem Phys 1956, 25: 534–549
20 Banker J. Amide modes and protein conformation. Biochim Biophys Acta 1992, 1120: 123–143.
21 Miyazawa T, Shimanouchi T, Mizushima S. Characteristic infrared bands of monosubstituted amides. J Chem Phys 1956, 24: 408
22 Kauppinen JK, Moffatt DJ, Mantsch HH, Cameron DG. Fourier self-deconvolution: A method for resolving intrinsically overlapped bands. Appl Spectrosc 1986, 35: 271–276
23 Lee DC, Haris PI, Chapman D, Mitchell RC. Determination of protein secondary structure using factor analysis of infrared spectra. Biochemistry 1990, 29: 9185–9193
24 Sarver RW, Krueger WC. Protein secondary structure from fourier transform infrared spectroscopy: A data base analysis. Anal Biochem 1991, 194: 89–100
25 Griffiths PR, deHaseth JA. Fourier Transform Infrared Spectroscopy. Wiley Interscience: New York 1986

©Institute of Biochemistry and Cell Biology, SIBS, CAS

Downloaded from https://academic.oup.com/abbs/article/39/8/549/420 by guest on 19 April 2024

26 Koenig JK, Tabb DL. Analytical Application of FT-IR to Molecular and Biological Systems. D. Reidel: Boston 1980
27 Holloway PW, Mantsch HH. Structure of cytochrome b5 in solution by Fourier-transform infrared spectroscopy. Biochemistry 1989, 28: 931–935
28 Chou PY, Fasman GD. β-turns in proteins. J Mol Biol 1977, 115: 135–175
29 Susi H, Timasheff SN, Stevens L. Infrared spectra and protein conformations in aqueous solutions. I. The amide I band in H_2O and D_2O solutions. J Biol Chem 1967, 242: 5460–5466
30 Gorga JC, Dong A, Manning MC, Woody RW, Caughey WS, Strominger JL. Comparison of the secondary structures of human class I and class II major histocompatibility complex antigens by Fourier transform infrared and circular dichroism spectroscopy. Proc Natl Acad Sci USA 1989, 86: 2321–2325
31 Kalnin NN, Venyaminov SY. Quantitative measurement of the IR spectra of water solution. J Appl Spectrosc 1988, 49: 1028–1032
32 Surewicz WK, Szabo AG, Mantsch HH. Conformational properties of azurin in solution as determined from resolution-enhanced Fourier-transform infrared spectra. Eur J Biochem 1987, 167: 519–523
33 Mitchell RC, Haris PI, Fallowfield C, Keeling DJ, Chapman D. Fourier transform infrared spectroscopic studies on gastric H^+/K^+-ATPase. Biochim Biophys Acta 1988, 941: 31–38
34 Dong A, Randolph TW, Carpenter JF. Entrapping intermediates of thermal aggregation in α-helical proteins with low concentration of guanidine hydrochloride. J Biol Chem 2000, 275: 27689–27693
35 Levitt M, Greer J. Automatic identification of secondary structure in globular proteins. J Mol Biol 1977, 114: 181–239
36 Hvidt A, Nielsen SO. Hydrogen exchange in protein. Adv Protein Chem 1966, 21: 287–386
37 Woodward CK, Hilton BD. Hydrogen exchange kinetics and internal motions in proteins and nucleric acids. Annu Rev Biophys Bioeng 1979, 8: 99–127
38 Englander SW, Kallenbach NR. Hydrogen exchange kinetics and structural dynamics of proteins and nucleic acids. Q Rev Biophys 1983, 16: 521–655
39 Englander SW. Protein folding intermediates and pathways studied by hydrogen exchange. Annu Rev Biophys Biomol Struct 2000, 29: 213–238
40 Barksdale AD Rosenberg A. Acquisition and interpretation of hydrogen exchange data from peptides, polymers, and proteins. Meth Biochem Anal 1982, 28: 1–114
41 Yu S, Lee LL, Lee JC. Effects of metabolites on the structural dynamics of rabbit muscle pyruvate kinase. Biophys Chem 2003, 103: 1–11
42 Weber IT, Steiz TA. Structure of a complex of catabolite gene activator protein and cyclic AMP refined at 2.5 Å resolution. J Mol Biol 1987, 198: 311–326
43 Passner JM, Schultz SC, Steiz TA. Modeling the cAMP-induced allosteric transition using the crystal structure of CAP-cAMP at 2.1 Å resolution. J Mol Biol 2000, 304: 847–859
44 From NB, Bowler BE. Urea denaturation of staphylococcal nuclease monitored by Fourier transform infrared spectroscopy. Biochemistry 1998, 37: 1623–1631
45 Fabian H, Mantsch HH. Ribonuclease A revisited: Infrared spectroscopic evidence for lack of native-like secondary structures in the thermally denatured state. Biochemistry 1995, 34: 13651–13655
46 Sosnick TR, Trewhella J. Denatured states of ribonuclease A have compact dimensions and residual secondary structure. Biochemistry 1992, 31: 8329–8335
47 Dong A, Randolph TW, Carpenter JF. Entrapping intermediates of thermal aggregation in α-helical proteins with low concentration of guanidine hydrochloride. J Biol Chem 2000, 275: 27689–27693
48 Dong A, Kendrick B, Kreilgard L, Matsuura J, Manning MC, Carpenter JF. Spectroscopic study of secondary structure and thermal denaturation of recombinant human factor XIII in aqueous solution. Arch Biochem Biophys 1997, 347: 213–220
49 Chirgadze YN, Fedorov OV, Trushina NP. Estimation of amino acid residue side-chain absorption in the infrared spectra of protein solution in heavy water. Biopolymers 1975, 14: 679–695
50 Dong A, Jones LS, Kerwin BA, Krishnan S, Carpenter JF. Secondary structure adsorbed onto aluminum hydroxide: Infrared spectroscopic analysis of proteins from low solution concentrations. Anal Biochem 2006, 351: 282–289

Edited by
Minghua XU

Downloaded from https://academic.oup.com/abbs/article/39/8/549/420 by guest on 19 April 2024