Gazi University Journal of Science
GU J Sci
27(1):637-644 (2014)

Evaluation of the Protein Secondary Structures Using Fourier Transform Infrared Spectroscopy

Andrei A. BUNACIU[1], Şerban FLESCHIN[2], Hassan Y. ABOUL-ENEIN[3,♠]

[1]*SCIENT - Research Center for Instrumental Analysis, (S.C. CROMATEC_PLUS S.R.L.), 18 Sos. Cotroceni, Bucharest - 6, 060114, ROMANIA*

[2]*Department of Organic Chemistry, Faculty of Chemistry, University of Bucharest, 92-96 Sos. Panduri, Bucharest-5, 050663, ROMANIA*

[3]*Pharmaceutical and Medicinal Chemistry Department, Pharmaceutical and Drug Industries Research Division, Tahrir Street, Dokki, Cairo 12311, EGYPT*

Received:12/03/2013 Revised:06/06/2013 Accepted:08/11/2013

ABSTRACT

Fourier transform infrared (FTIR) spectroscopy is an attractive tool for proteomics research as it can be used to rapidly characterize protein secondary structure in aqueous solution. Fourier transform infrared spectrometry is well known as a powerful tool for determination of secondary structures of proteins. This paper presents the most recent experimental results obtained in our laboratory for this kind of analysis. A PeakFit operation was performed and some observations were made.

Key words: FT-IR analysis, protein secondary structure.

1. INTRODUCTION

As the "second part of the genetic code" proteins play a pivotal role in living organisms, facilitating metabolism, communication, transport, and the maintenance of structural integrity. A true understanding of proteins can only be achieved by exploring the relationship that exists between the unique structure adopted by a protein and its function. The enormous structural diversity of proteins begins with different amino acid sequences (primary structure) of polypeptide chains that fold into complex 3D structures. The final folded arrangement of the polypeptide chain is referred to as its conformation (secondary and tertiary structures). The requirements for hydrogen bond reservation in the folded structure result in the cooperative formation of regular hydrogen bonded secondary structure regions in proteins. The secondary structure specifies regular polypeptide chain folding patterns of helices, sheets, coils, and turns which are combined / folded into tertiary structure. A variety of techniques have been applied to the elucidation of the three-dimensional structure of proteins, ranging from prediction based on the sequence and physico-chemical properties of the constituent amino acids [1,2] to more precise methods for the identification of atoms and the determination of their molecular coordinates, such as NMR spectroscopy [3] and X-ray diffraction [4] .

As the number of sequences identified by the various genome projects increases at a phenomenal rate, it becomes correspondingly necessary to improve methods for predicting their three-dimensional (3D) structures. Understanding structural relationships between proteins, such as whether certain architectures occur more frequently than others, can inform these approaches but requires an appreciation of all the known structures. Therefore, fast and automatic methods are needed to evaluate the relationships between the known structures.

Proteins are frequently referred to as having a certain fraction of helical structure and a certain fraction of extended beta-structure, but there are no generally valid objective criteria to establish numerical values in this

♠Corresponding author, e-mail: haboulenein@yahoo.com

context because, among other things, the endpoints of segments with different secondary structures are not easy to determine in an objective manner.

The secondary structure 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 techniques. α-Helices and pleated β-sheets were predicted in 1951 by Linus Pauling and Robert Corey [5] on the basis of hydrogen-bonding and cooperativity criteria.

Fourier transform infrared (FTIR) spectroscopy has emerged as a useful tool for the characterization of protein secondary structure with a precision lying between that of the purely predictive and the molecular coordinate approaches [6-11], being one of the earliest experimental method for estimating the secondary structure of polypeptides and proteins [12]. That IR spectroscopy could give us informations related to the secondary structure of proteins was first demonstrated by Elliot and Ambrose [12, 13], who showed that there is an empirical correlation between the frequency of the amide I and amide II absorptions of a protein and the predominant secondary structural motif within the protein as determined by X-ray diffraction studies. These studies demonstrated that proteins that were known to be predominantly α-helical in structure exhibited amide I and II absorptions in the spectral range 1652 to 1657 and 1545 to 1551 cm^{-1}, respectively, in aqueous solution, while proteins with a predominantly β-sheet structure exhibited similar absorptions at 1628 to 1635 and 1521 to 1525 cm^{-1}.

These initial observations were later extended to proteins in ^2H_2O solutions [14,15], because was shown the contribution of water with two main bands in the mid-infrared range: the intense OH stretching mode, between 3600 and 3100 cm^{-1} and the HOH bending mode, δ(HOH), at about 1645 cm^{-1}.

The regular secondary structures [16,17], α helices and β sheets, are connected by coil or loop regions of various lengths and irregular shape as can be seen in Figure 1.

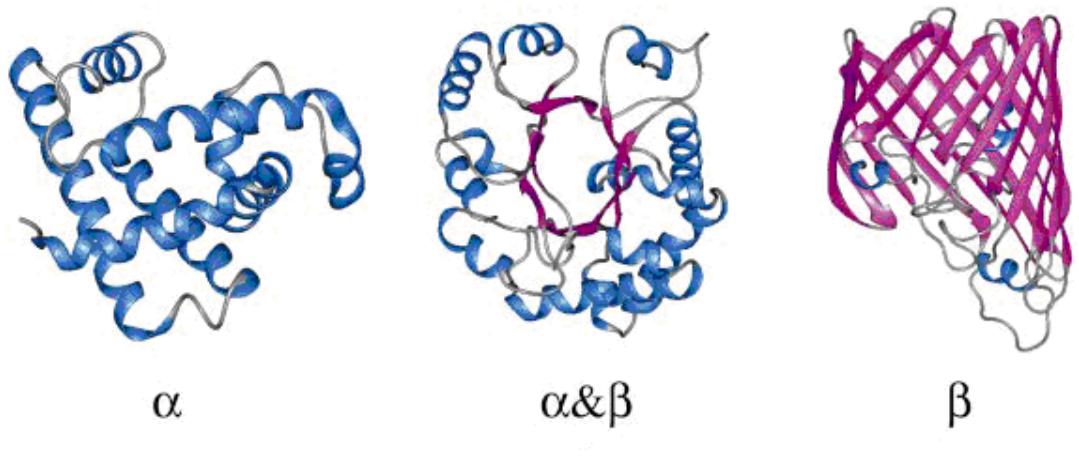

Figure 1. Schematic representation of protein secondary structures.

The structure classification procedure naturally results in a tree hierarchy outlining the relationships between folds. The data is stored in a hierarchical database (CATH - protein class (C), architecture (A), topology (T) and homologous superfamily (H)) [17], with each structure indexed by a CATH number akin to the EC nomenclature for enzymes. CATH has been made accessible in a hypertext form over the World Wide Web for use by text-based or graphical browsers (http://www.biochem.ucl.ac.uk/bsm/cath).

Proteins are frequently referred to as having a certain fraction of helical structure and a certain fraction of extended beta-structure, but there are no generally valid objective criteria to establish numerical values in this context because, among other things, the endpoints of segments with different secondary structures are not easy to determine in an objective manner. In the FTIR spectra of proteins, the secondary structure is most clearly reflected by the amide I and amide II bands, particularly the former, which absorbs around 1620 to 1690 cm^{-1} and is primarily associated with the stretching vibrations of peptide carbonyl groups [18].

A major advantage of FTIR spectroscopy for structural characterization is the lack of dependence on the physical state of the sample. Samples may be readily examined as aqueous or organic solutions, hydrated films, inhomogeneous dispersions, or solids, and proteins have been analyzed by FTIR spectroscopy in all of these physical states. Information may be obtained from analysis of solid proteins either deposited on an IR-transparent substrate, pressed into a KBr pellet, or by using specialized techniques such as photoacoustic IR spectroscopy [19]. However, the information obtained from such an analysis is of limited relevance due to the nonphysiological nature of the measurement. In addition, the structure of the protein or peptide under investigation may depend critically on the medium from which it was dried and the solvent history may not be known. In addition to these traditional methods

for the analysis of solid materials, a number of studies [20] have been reported in which water-soluble proteins have been analyzed as films by attenuated total reflectance (ATR). In such studies, a small volume of dilute protein solution is allowed to dry on the surface of a (usually trapezoidal) crystal composed of an IR transparent material of high-refractive index.

The purpose of this study was to investigate for the first time in our laboratory how FTIR spectroscopy can be used in examination of the secondary structures of some proteins in solid in a qualitative manner. It appears that a judicious application of deconvoluted FTIR spectra can form a useful supplement to the existing methods.

2. EXPERIMENTALS

Reagents

The proteins studied were: Alcohol dehydrogenase was supplied by Fluka - Germany, Hemoglobin and Muramidase were supplied by Merck, Alkaline phosphatase supplied by Reanal - Hungary while Aldehyde dehydrogenase, Cytochrome C, Cytochrome C reductase, Glucose dehydrogenase, Tyrosinase and Xanthine oxidase were supplied by Sigma Aldrich. For fused KBr disk preparation a potassium bromide IR spectral grade was used (Sigma Aldrich).

Apparatus

Data acquisition was performed using a Spectrum Two System FT-IR spectrometer equipped with Spectrum for Windows v.10 (Perkin Elmer Co.). all the ATR-FTIR spectra were recorded between 4000 and 650 cm^{-1}, at room temperature, after 32 scans with a spectral resolution 0f 2 cm^{-1}. This software also provided for a complete processing of the spectra measured. For PeakFit analysis special software was used (PeakFit v. 4.1, Jandel Scientific Software).

3. RESULTS AND DISCUSSION

The ATR-FTIR spectra of the proteins studied in this first research are presented in figure 2. Spectra have been scaled to an identical area under amide I and amide II (1700 – 1400 cm^{-1}) and a linear baseline correction between these two frequencies. Because water absorbs strongly in the most important spectral range, around 1640 cm^{-1}, studies in aqueous solution are difficult to evaluate, so deuterium oxide is suggested as solvent.

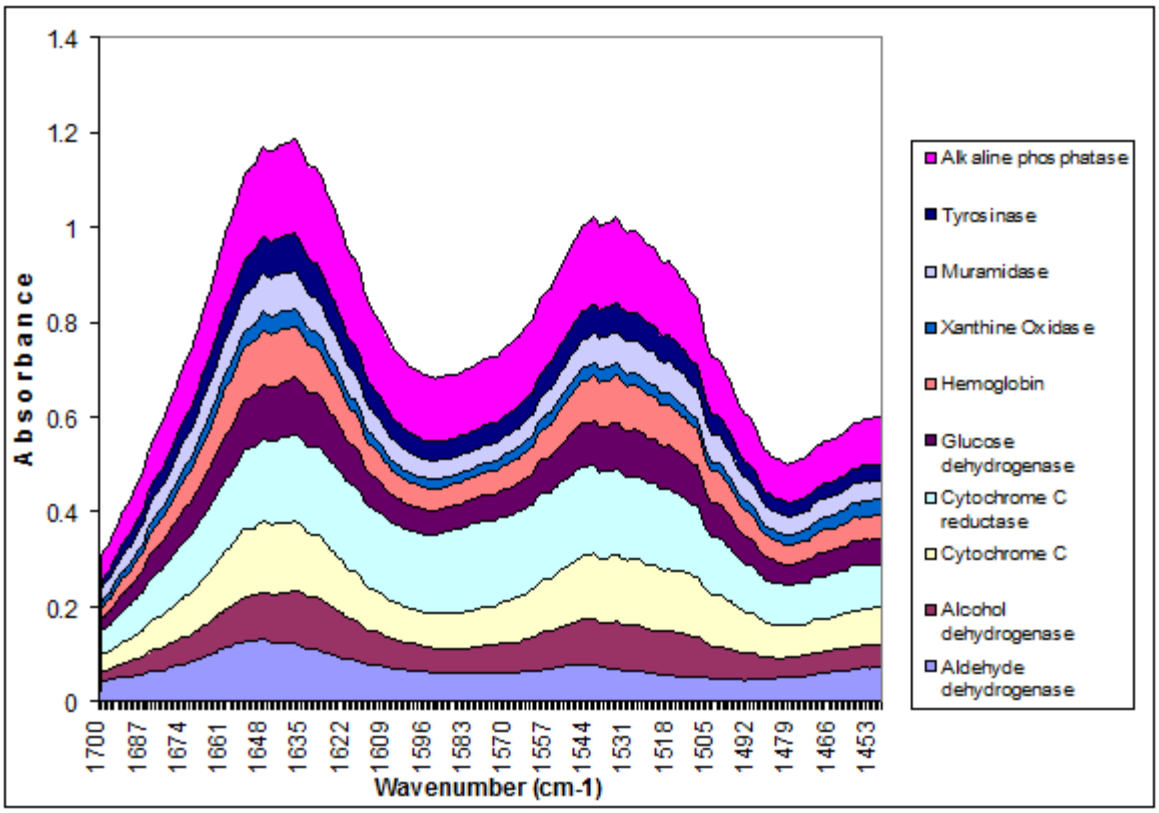

Figure 2. ATR-FTIR spectra of the proteins.

There were performed some more experiments related to other methods used for obtaining FTIR spectra. KBr disks were tried and a better intensity was obtained as can be seen in figure 3, where are presented comparatively.

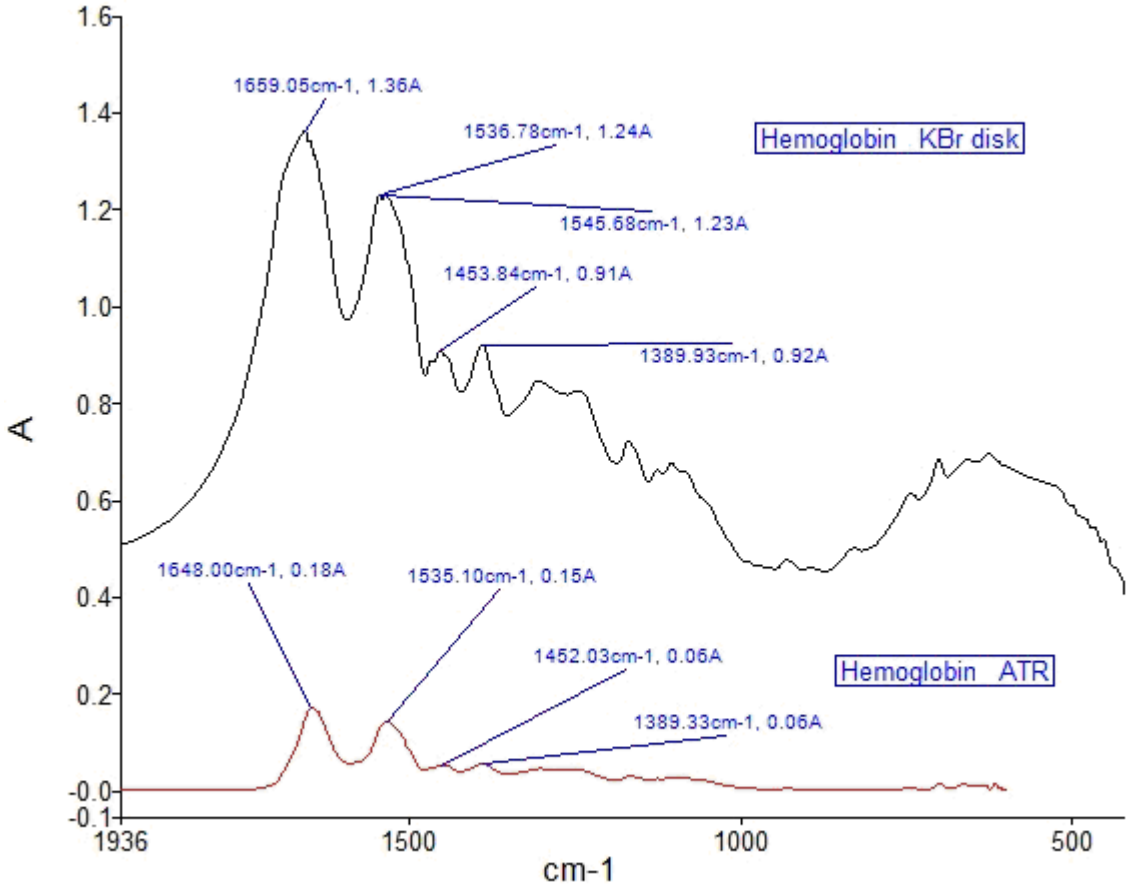

Figure 3. Comparison between KBr and ATR -FTIR spectra of haemoglobin.

As can be seen the spectra are similar as shape, but for KBr disk method we obtained a better signal. The algorithm we used for deconvolution includes both an apodization function and a line-shape one. We find that the deconvoluted spectra can fit reasonably well by assuming that the components have Gaussian line shapes. There are some studies [8,21] that suggest that the sum of all the integrated areas of the "beta-bands " as a fraction of the total amide I band area is closely related to the total "beta-content " of a protein. The same applies to the helices. Otherwise, quantitative estimations would be very difficult or even impossible, because the number and frequency of the components varies from a protein to another.

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%) [22, 23], showing much less protein conformational sensitivity than its amide I counterpart. 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 [24] described the nine amide vibration modes and some standard conformations in detail in his review. The characteristic IR bands of the proteins and peptides are listed in Table 1.

Table 1. Characteristics infrared bands of peptide linkage

Designation	Approximate wavenumber (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

Table 2 presents the principal FTIR peaks for the proteins studied in the range 1600-1500 cm^{-1}, while in Table 3 we present the peaks observed after PeakFit process. Figure 4 presents comparatively the FTIR spectra, in ATR and KBr methods used, after PeakFit process was performed for Hemoglobine protein. Similar spectra were obtained for all the substances used in this study.

Table 2. The principal FTIR peaks for the proteins studied

Proteins studied	**FTIR Peak spectra (cm^{-1})**						
Xanthine oxidase	-	1638.01	-	-	1541.89	-	-
Tyrosinase	-	1638.18	-	-	1542.23	-	-
Alkaline phosphatase	-	1637.31	1630.33	-	1542.33	1535.08	-
Muramidase	1644.80	-	-	-	-	-	1526.73
Haemoglobine	1649.03	-	-	-	-	1535.07	-
Glucose dehydrogenase.	-	1637.68	-	-	-	1534.63	-
Cytochrome C reductase.	-	-	1625.24	1578.07	1545.95	-	-
Cytochrome C	1648.28	1638.01	-	-	1542.11	1534.78	
Aldehyde dehydrogenase	1648.61	-	-	-	1542.20	-	-
Alcohol dehydrogenase	-	1637.02	-	-	-	1534.61	

Table 3. The principal FTIR peaks for the proteins studied after PeakFit process

Proteins studied	FTIR peak spectra after PeakFit process (cm^{-1})											
Xanthine oxidase	-	-	1539.07	-	-	1580.35	-	-	1621.87	1662.24	-	1690.04
Tyrosinase	-	1520.12	-	1548.36	-	1585.52	-	-	1628.28	-	-	1686.68
Phosphatase alkaline	1507.05	1520.57	1536.92	1551.38	-	1575.65	1600.87	1615.55	1627.03	1652.78 - 1659.59	1676.89	1696.64
Muramidase	-	1516.05	1533.02	1548.36	1565.32	-	-	-	1628.28	1644.32	-	-
Hemoglobine	-	1517.08	-	1546.41	-	-	-	-	1628.28	1651.07	-	-
Glucose dehydrogenase	-	1519.47	-	1549.67	-	1583.77	-	-	1622.09	1656.41	-	1688.33
Cytochrome C reductase	-	1519.28	-	1549.93	-	1583.95	-	1619.84	-	1648.62	1678.73	-
Cytochrome C	1512.15	-	1530.79	1550.60	-	1576.25	1605.29	-	1627.78	1664.23	1673.30	1686.83
Aldehyde dehydrogenase	1506.04	-	-	1543.70	-	1577.70	-	1613.16	-	1655.26	-	-
Alcohol dehydrogenase	1512.2	-	1531.38	1556.15	-	-	-	-	1625.32	1654.98	-	1683.99

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. A component centered between approximately 1658 and 1650 cm^{-1} has been assigned to the α-helix, which is consistent with both theoretical calculation[24] and the observation of bands in the spectra of α-helical proteins[7, 26]. Bands near 1663 cm−1 are assigned to 3_{10} helices [7, 24, 25], 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 [7, 23-27]. Theoretical calculation of β-sheets also predicts an IR active mode between approximately 1695 and 1670 cm^{-1}. 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. 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}.

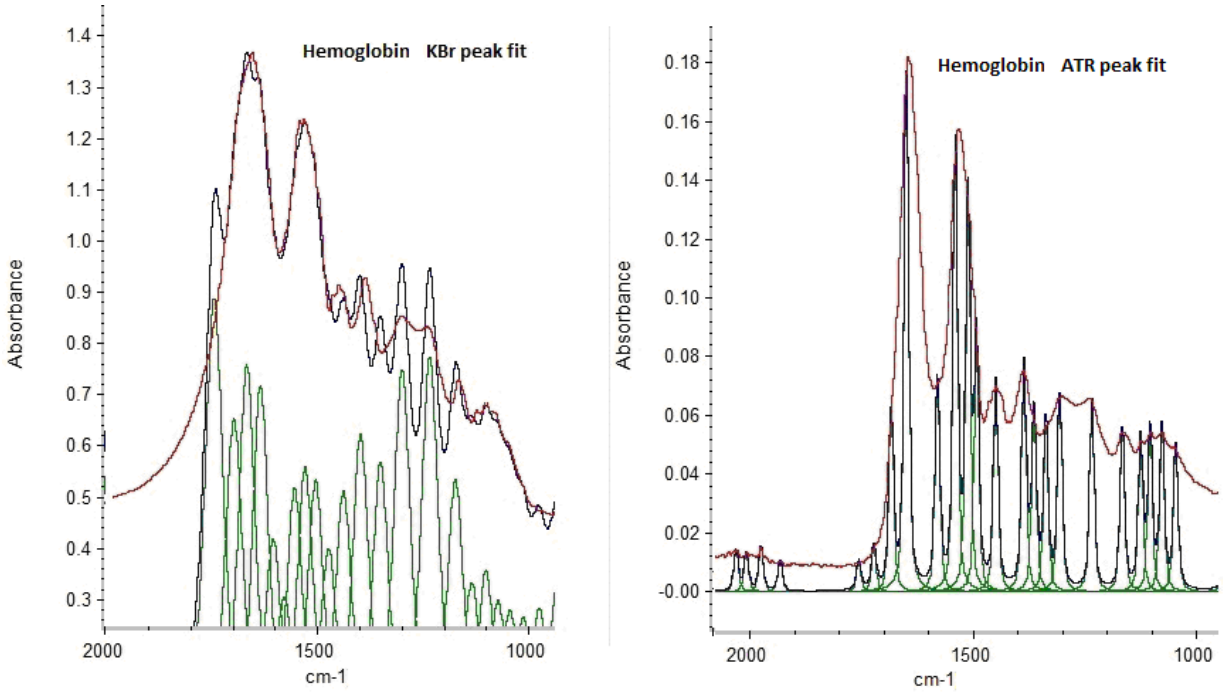

Figure 4. Comparison between KBr and ATR -FTIR spectra of haemoglobine after PeakFit process.

Although its potential usefulness has been recognized for some time, the recognition of the different components of conformation-sensitive amides bands (particularly amide I band) was difficult due to the resulting overlap of bands originating from the different secondary structures such as α-helix, β-sheet, β-turns and random. Pointing out the different components of amide I have been made possible by the advent of numerical analysis of the spectra.

4. CONCLUSIONS

FTIR spectroscopy is a well-established experimental technique for studying the secondary structural composition. Empirically, most workers have found the amide I absorption to be more useful for protein secondary structure determination than the amide II absorption, probably due to the fact that it effectively arises from only one of the amide functional groups, in contrast to the amide II mode. The correlations between IR spectra and protein secondary structures have been established. 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. 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.

REFERENCES

[1]. Chou and Fasman, G. D., "Prediction of the secondary structure of proteins from their amino acid sequence", ***Adv. Enzymol***, 47: 45-148 (1978).

[2]. Fasman, G.D., "Protein conformational prediction", ***Trends Biochem. Sci.,*** 14: 295-299 (1989).

[3]. Braun, W., "Distance geometry and related methods for protein structure determination from

NMR data", ***Quarterly Rev. Biophys.,*** 19: 115-157 (1987)

[4]. Wlodawer. A., Bott, R. and Sjoh, L., "The refined crystal structure of ribonuclease A at 2.0 A resolution", ***J. Biol. Chem.,*** 2S7: 1325-1332 (1982)

[5]. Pauling, L. and Corey, R. B., "Configurations of Polypeptide Chains With Favored Orientations Around Single Bonds: Two New Pleated Sheets", ***Proc. Natl. Acad. Sci. USA***, 37: 729-740 (1951).

[6]. Arrondo, J. L. R.; Muga, A.; Castresana, J., and God. F. M., "Quantitative studies of the structure of proteins in solution by fourier-transform infrared spectroscopy", ***Prog. Biophys. Mol. Biol.,*** 59: 23-56 (1993).

[7]. Susi, H. and Byler, D. M., "Resolution-enhanced Fourier transform infrared spectroscopy of enzymes", ***Methods Enzymol.,*** 130: 290-311 (1986).

[8]. Byler, D. M. and Susi, H., "Examination of the secondary structure of proteins by deconvolved FTIR spectra", ***Biopolymers***, 25: 469-487 (1986).

[9]. Surewicz, W. K. and Mantsch, H. H., ***Biochim. Biophys. Acta***, 952: 115-130 (1988).

[10]. Jackson, M.; Haris, P., and Chapman, D., "Fourier transform infrared spectroscopic studies of lipids, polypeptides and proteins", ***J. Mol. Structure,*** 214: 329- 355 (1989).

[11]. Jackson, M. and Mantsch, H. H., "The Use and Misuse of FTIR Spectroscopy in the Determination of Protein Structure", ***Critical Reviews in Biochemistry and Molecular Biology***, 30(2): 95-120 (1995).

[12]. Elliot, A. and Ambrose. E. J., "Structure of Synthetic Polypeptides", ***Nature,*** 165: 921-922 (1950).

[13]. Ambrose, E. J. and Elliot, A., "Infra-Red Spectroscopic Studies of Globular Protein Structure", ***Proc. R. Soc. London, Ser A***, 208: 75-90 (1951).

[14]. Susi, H.; Timasheff, S. N. and 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.,*** 242: 5460-5466 (1967).

[15]. Timasheff, S. N.; Susi, H. and Stevens, L., "Infrared Spectra and Protein Conformations in Aqueous Solutions: II. Survey of Globular Proteins", ***J. Biol. Chem.,*** 242: 5467-5473 (1967).

[16]. Levitt, M. and Chothia, C., "Structural patterns in globular proteins", ***Nature,*** 261: 552–558 (1976).

[17]. Orengo, C.A.; Michie, A.D.; Jones, S.; Jones, D.T.; MB Swindells, M.B., and JM Thornton, J.M**.,** **"**CATH – a hierarchic classification of protein domain structures**",** ***Structure***, 5(8): 1093-1109 (1997).

[18]. Ruegg, M.; Metzger, V. and Susi, H., "Computer analyses of characteristic infrared bands of globular proteins", ***Biopolymers,*** 14: 1465-1471 (1975).

[19]. Wang, J.; Sowa, M. G.; Ahmed M. K. and Mantsch, H. H., "Studies of photoionization dynamics of methylidyne, imidogen, and hydroxyl radicals at near-threshold photoelectron kinetic energies", ***J. Phys. Chem.,*** 98(2): 4748-4755 (1994).

[20]. Goormaghtigh, E.; Cabiaux Veronique and Ruysschaert, Jean-Marie, "Secondary structure and dosage of soluble and membrane proteins by attenuated total reflection Fourier-transform infrared spectroscopy on hydrated films", ***Eur. J. Biochem.***, 193(2): 409-420 (1990).

[21]. Levitt, M. and Greer, J., "Automatic identification of secondary structure in globular proteins", ***J. Mol. Biol.,*** 114: 181-239 (1977).

[22]. Kong Jilie and Yu Shaoning, "Fourier Transform Infrared Spectroscopic Analysis of Protein Secondary Structures", ***Acta Biochimica et Biophysica Sinica,*** 39(8): 549–559 (2007).

[23]. Krimm, S. and Bandekar, J., "Vibrational spectroscopy and conformation of peptides, polypeptides, and proteins", ***Adv. Proteins Chem.,*** 38: 181-364 (1986).

[24]. Banker, J., "Amide Modes and Protein Confimation", ***Biochim. Biophys. Acta***, 1120: 123-143 (1992).

[25]. Koenig, J.K. and Tabb, D.L., Analytical Application of FT-IR to Molecular and Biological Systems. D. Reidel: Boston 1980

[26]. Dong, A.; Huang, P. and Caughey, W.S., "Protein secondary structures in water from second-derivative amide I infrared spectra", ***Biochemistry,*** 29: 3303−3308 (1990).

[27]. Lee, D.C.; Haris, P.I.; Chapman, D. and Mitchell, R.C., "Determination of protein secondary structure using factor analysis of infrared spectra", ***Biochemistry,*** 29: 9185−9193 (1990).