ELSEVIER

Journal of Molecular Catalysis B: Enzymatic 7 (1999) 207–221

JOURNAL OF MOLECULAR CATALYSIS B: ENZYMATIC

www.elsevier.com/locate/molcatb

FTIR spectroscopic characterization of protein structure in aqueous and non-aqueous media

Parvez I. Haris [a,*], Feride Severcan [b]

[a] *Department of Biological Sciences, De Montfort University, Hawthorn Building, The Gateway, Leicester LE1 9BH, UK*
[b] *Molecular Biology–Biotechnology Research Unit, Department of Biology, Middle East Technical University, 06531 Ankara, Turkey*

Abstract

With increasing use of proteins in many different applications, ranging from phramaceuticals to biosensors and biomaterials, there has emerged a need for protein structural characterisation in diverse environments. In many cases it is not sufficient to just have the three-dimensional structure of a protein in H_2O or in the crystalline state. Often information on the structural properties of a protein is required in the presence of organic solvents, detergent micelles, phospholipid membranes and so on. Fourier transform infrared spectroscopy (FTIR) has been identified as one of the few techniques that can be applied for structural characterisation of proteins in such environments. Here we discuss how this technique is being used to obtain information on protein structure and stability in both aqueous and non-aqueous media. Examples are drawn from our studies of water soluble proteins and membrane proteins.

Keywords: Infrared spectroscopy; FTIR spectroscopy; Protein conformation; Protein stability; Protein folding; Protein secondary structure; Membrane proteins

1. Introduction

With the progress in sequencing of the human genome, the disparity between the number of known sequences and the number of experimentally determined protein structures continues to increase. Consequently, there is a clear demand for development of techniques for rapid structural characterisation of the encoded proteins. X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy provide the complete three-dimensional structure of a protein and are by far the most powerful techniques available to structural biochemists. Although X-ray crystallography is an excellent technique for the determination of three-dimensional structure of proteins it has the following disadvantages: crystallographic studies require high-quality single crystals which are not available for many proteins such as most of the membrane proteins, and the structure of a protein in a crystal may not always relate to its structure in solution. X-ray diffraction data presents a static picture of protein structure which does not represent the protein conformation with its dynamic nature in biological systems. The slowness of the procedure is the other disadvantage of the technique. NMR spectroscopy has better flexibility to study protein structure in solution. However, the interpretation of NMR

Abbreviations: FTIR, Fourier transform infrared spectroscopy; NMR, Nuclear magnetic resonance spectroscopy

* Corresponding author. E-mail: pharis@dmu.ac.uk

1381-1177/99/$ - see front matter
PII: S1381-1177(99)00030-2

spectra of larger proteins is very complex, and the technique is presently limited to small proteins (30 kDa).

It is important to point out that no one technique in the present arsenal of protein structural methods is able to provide information on all aspects of protein structure. Therefore, a rational strategy is to employ a concerted approach in which the protein is examined using several structural techniques. Information obtained from different techniques can be cross-correlated to provide a more complete picture of the chemical and physical state and/or bioactivity of the protein under different conditions. One of the techniques which has recently become very popular for structural characterisation of proteins is Fourier transform infrared (FTIR) spectroscopy. Here, we discuss the application of infrared spectroscopy for structure and stability studies of proteins in aqueous and non-aqueous media.

Infrared spectroscopy is based on molecular vibrations. Chemical bonds undergo various forms of vibrations such as stretching, twisting and rotating. The energy of most molecular vibrations corresponds to that of the infrared region of electromagnetic spectrum. Many of the vibrations can be localised to specific bonds or groupings, such as the C=O and O–H groups. This has led to the concept of characteristic group frequencies. Typical group frequencies of interest to biochemists include C=O, –COOH, COO^-, O–H and S–H. There are many vibrational modes that do not represent a single type of bond oscillation but are strongly coupled to neighbouring bonds. For example, the infrared spectrum of a protein is characterised by a set of absorption regions known as the amide modes. With developments in FTIR instrumentation it is now possible to obtain high quality spectra from dilute protein solutions in H_2O [1–4]. The overlapping H_2O absorption can be digitally subtracted from the spectrum of the protein solution. In addition, the broad infrared bands in the spectra of proteins can be analysed in detail using second-derivative and deconvolution procedures [1–4]. These procedures can be utilised to reveal the overlapping components within the broad amide bands. Difference spectroscopy has the advantage in providing highly detailed information on conformational changes in proteins [5].

The most important advantage of FTIR spectroscopy for biological studies is that spectra of almost any biological material can be obtained in a wide variety of environments. Spectra of a protein can be obtained in single crystals, in aqueous solution, organic solvents, detergents micelles, lipid membranes, etc. The chemical environment in which a peptide or protein exists influences its structure and stability. This has important implications for the formulation, storage and delivery mechanisms for protein therapeutics. There is increasing evidence indicating that the environment can be important in determining the secondary structure formed by an amino acid sequence [6,7]. Other advantages of the technique include the following: the amount of protein required is relatively small (10 μg); the size of the protein is not important; there is no light scattering or fluorescent effects; kinetic and time-resolved studies are possible; and inexpensive compared to the cost of X-ray diffraction, NMR, ESR and CD spectroscopic equipments.

2. Materials

2.1. Instrumentation

Early studies of proteins using infrared spectroscopy was hindered by the low sensitivity of infrared spectrometers and absorption of liquid water over much of the infrared spectrum. With developments in FTIR instrumentation these problems have been solved. There are several companies that manufacture FTIR spectrometers and provide softwares for spectral analysis. A broad range of FTIR accessories are also available to overcome the problems of studying biological samples. For example, coupling of opti-

cal microscopes to an FTIR spectrometer enables the study of very small quantities of biological samples. Spectra of small crystals of peptides and proteins can be analysed using this approach. FTIR spectra of biological samples are most commonly recorded via the transmission mode, while the Attenuated Total Reflection (ATR) method is a useful method [8–10] for characterising the orientation of components of thin films, such as lipids and peptides in biomembranes.

2.2. Sample handling

Since H_2O is a strong absorber of infrared radiation, it is necessary to use short path length cells to record spectra of protein in H_2O. Consequently, a relatively high protein concentration ($\sim$ 10 mg/ml) is needed for such studies. However, ^2H_2O and many organic solvents do not have strong absorption bands in the amide I region (1600–1700) and in these cases much longer path length cells can be used, and hence much lower protein concentration ($\sim$ 1 mg/ml) is required to obtain high quality spectra.

For recording FTIR spectra, aqueous samples (between 10 μl to 50 μl) are placed in a thermostated cell fitted with usually CaF_2 windows (with a 6 μm tin spacer for measurements in H_2O or a 50 μm Teflon spacer for samples in ^2H_2O). Other window materials which are not soluble in water can also be used. Spectra of biological materials can be recorded at the desired temperature by linking a temperature controller unit to the FTIR spectrometer. Spectra are usually recorded at a resolution of 4 cm^{-1}. The number of scans accumulated depends on the concentration of the biological material used. For a proteins in H_2O, at a concentration of 10 mg/ml, approximately 200–400 scans are sufficient to obtain a high signal-to-noise ratio spectrum. In ^2H_2O one can use fewer number of scans, often 100–150 scans are sufficient.

Acquisition of high signal-to-noise ratio spectra of peptides and proteins requires the elimination of water vapour from the sample compartment of the spectrometer. This is because the narrow water vapour bands overlap with the conformation sensitive amide I band. Water vapour contributions can be reduced by purging the instrument with dry air or nitrogen, and by employing a sample shuttle which allows the background to be signal-averaged concurrently with the sample. Also a pre-recorded water vapour absorption spectrum can be subtracted from the protein absorption spectrum to reduce the vapour bands still further. However, subtraction of water vapour bands can be complicated

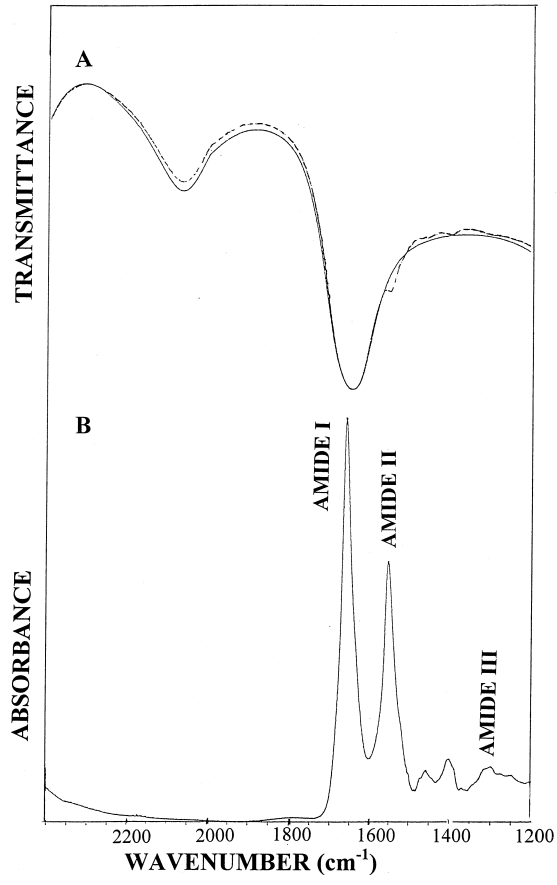

Fig. 1. (A) Transmittance spectrum of H_2O overlaid on top of the spectrum of a water-soluble protein in H_2O. (B) The absorbance spectrum of the protein obtained after the digital subtraction of the H_2O spectrum from the spectrum of the protein in H_2O. The absorbance spectrum shows bands corresponding to the amide I, amide II and amide III regions. The spectra were recorded using CaF_2 windows fitted with 6 μm tin spacers. The spectrum was recorded at 20°C.

and over-subtraction or under-subtraction can introduce artifacts.

Fig. 1A shows the transmittance spectrum of a water soluble protein in H_2O. The spectrum of H_2O itself is overlaid on top of the protein spectrum. Due to the strong absorbance of H_2O, the bands from the protein are masked out. In order to visualise the protein bands, it is necessary to digitally subtract the solvent absorption from the spectrum of the protein sample. Fig. 1B shows the absorbance spectrum of the protein in H_2O after subtraction of the overlapping water absorbance. Fig. 2 shows the absorbance spectra of a protein in H_2O and ^2H_2O. For accurate solvent subtraction, spectra of the solvent need to be recorded under identical conditions as the sample spectrum [3,11,12]. Thus, cell pathlength, temperature, number of scans, resolution, etc., need to be kept as identical as possible. Buffer subtraction are carried out digitally to give a straight baseline between 2000–1800 cm^{-1}. More details regarding digital subtraction of H_2O and ^2H_2O from spectra of protein solutions can be found in our earlier publications [3].

When carrying out FTIR spectroscopic measurements it is important to check the purity of the peptide and protein samples. This is because there are possibilities of trace contaminants remaining in the sample from the various steps involved in the purification and synthesis processes. For example, trifluoroacetic acid (TFA) is commonly used for the cleavage of peptides from the resin after synthesis, and is also used for HPLC purification procedures. This molecule can often remain strongly associated with the peptide. The COO^- stretching vibration of the trifluoroacetate counterion gives rise to an intense band at 1673 cm^{-1}, overlapping with the amide I band. The need to be cautious in assigning bands at this frequency in the spectra of peptides which have been in contact with TFA has been previously reported [2,13].

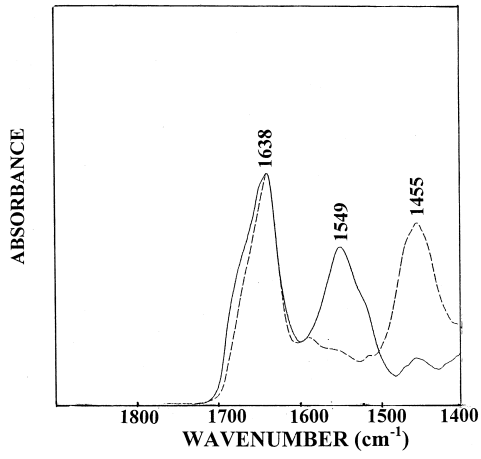

Fig. 2. Absorbance spectra of a predominantly β-sheet protein, α-chymotrypsin, in H_2O (continuous line) and in ^2H_2O after extensive hydrogen–deuterium exchange (broken line). Both spectra were recorded at 20°C. The amide I band maximum is centred at 1638 cm^{-1} and the amide II band at 1549 cm^{-1}. In ^2H_2O the amide II band is reduced in intensity as it shifts to $\sim$1455 cm^{-1}.

3. Methods

3.1. Data processing techniques

The amide I band of proteins consists of many overlapping component bands that represent different structural elements such as α-helices, β-sheets, turns and non-ordered or irregular structures. The width of the contributing component bands is usually greater than the separation between the maxima of adjacent peaks. As a consequence, the individual component bands cannot be resolved in the experimental spectra. The Fourier deconvolution procedure, sometimes referred to as 'resolution enhancement' involves narrowing the widths of infrared bands, allowing increased separation of the overlapping components present within the broad band envelope [14]. Increased separation of the overlapping bands can also be achieved by calculating the second derivative (rate of change of slope) of the absorption spectrum, [15]. Fig. 3 shows the absorbance, deconvolved and second derivative FTIR spectra of a predominantly α-helical protein. Bands revealed using these band-narrowing techniques enables one to identify the different structures present in a protein and also to detect conformational changes by monitoring alterations in the fre-

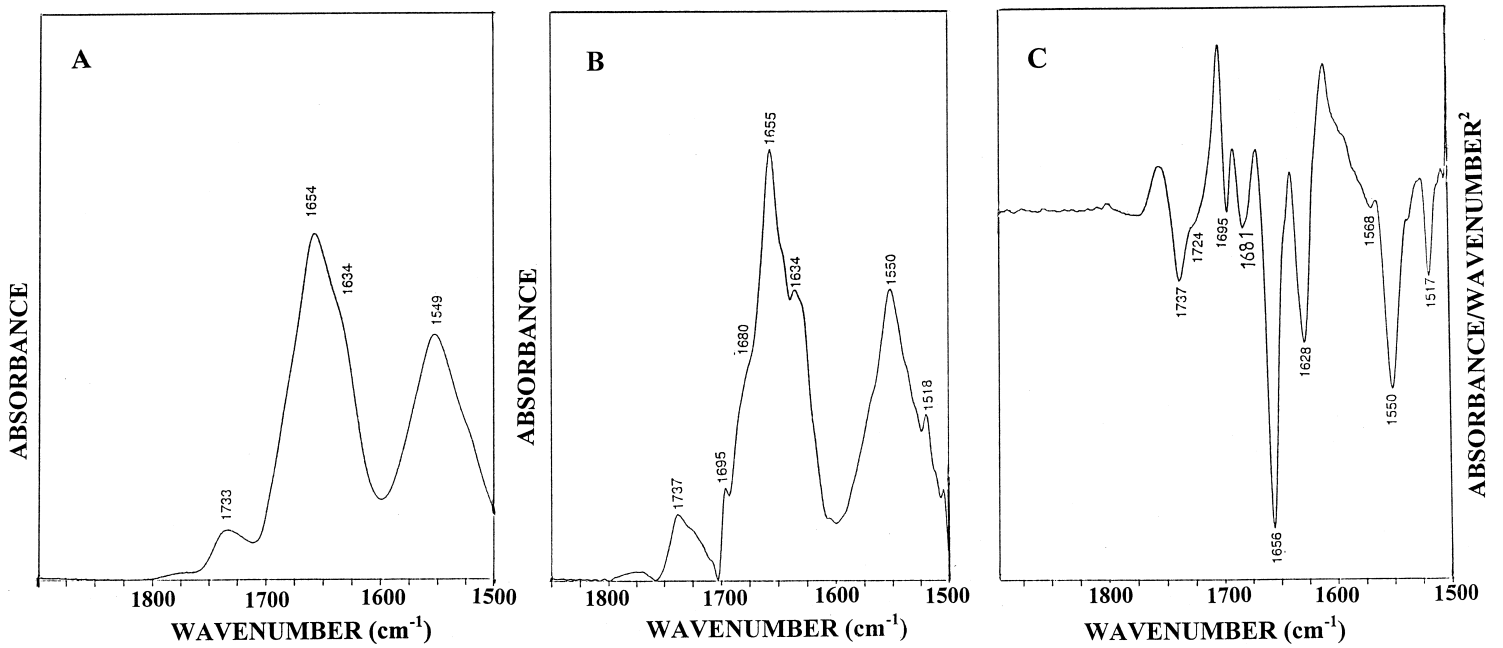

Fig. 3. Absorbance (A), deconvolved (B) and second-derivative (C) spectra of a predominantly α-helical membrane protein, the reaction centre from *R. sphaeroides*, in H_2O phosphate buffer. The spectrum was recorded at 20°C.

quency and intensity of these bands [1–4,17,18]. Both second-derivative and deconvolution procedures have been successfully applied in the qualitative study of a large number of proteins. In addition to providing valuable information about their secondary structure, the method has been shown to be useful for detecting conformational changes arising as a result of ligand binding, pH, temperature, organic solvents, detergents, etc. In many cases results obtained using this approach has been later supported by studies using other techniques such as X-ray diffraction and NMR. However, both derivative and deconvolution techniques should be applied with care since they amplify the noise significantly. Another problem that needs to be considered is that atmospheric water vapor display very narrow absorption bands which overlap with the amide I band and are also amplified in second-derivative and deconvolved spectra. As mentioned earlier a good purge of the spectrometer with dry air or nitrogen is necessary to overcome this problem.

Difference spectroscopy is another approach that is very useful for investigating subtle differences in protein structure. The principle of difference spectroscopy involves the subtraction of a protein absorbance spectrum in state A from that of the protein in state B. This results in the simplification of the spectrum since absorption from those groups within the protein that do not participate in the conformational change are subtracted. The resultant difference spectrum only shows peaks that are associated with those groups involved in the conformational change. Much of these types of spectroscopic studies have been carried out on large membrane bound proteins, that contain an internal trigger, such as the photosynthetic reaction centres and bacteriorhodopsin [5,19].

3.2. Analysis and assignment of protein infrared spectra

The amide bands that arise from the vibration of the peptide groups provide information on the secondary structure of polypeptides and proteins. Analysis of the peptide group vibration in model compounds and in polypeptide systems allows assignment of these characteristic bands. Changes in the hydrogen bonding involved in the peptide linkages results in changes in the vibrational frequency of the different amide modes. The modes most widely used in protein structural studies are amide I, amide II and amide III (see Fig. 1). The amide I band arises principally from the C=O stretching vibration of the peptide group. The amide II band is primarily N–H bending with a contribution from C–N stretching vibrations. The amide III absorption is normally very weak in the infrared, arising primarily from N–H bending and C–N stretching vibrations.

As seen from Figs. 1–3, the amide I absorption occurs in the region 1600–1700 cm^{-1}. Hydrogen bonding and the coupling between transition dipoles are amongst the most important factors governing conformational sensitivity of the amide bands. Amide I bands in the spectral range 1620–1640 cm^{-1} with proteins can be attributed to β-sheet structure [1–4, 18,20]. Fig. 3 shows a predominantly β-sheet containing-protein (α-chymotrypsin) in H_2O and in ^2H_2O. As can be seen from the figure, there is a large reduction in the intensity of the amide II band in ^2H_2O. This is due to the hydrogen–deuterium exchange of the amide bonds resulting in the shift of the amide II band towards lower frequency (approximately 1455 cm^{-1}) [3]. It is interesting to note from Fig. 2 that there is very little change in the position of the amide I band in ^2H_2O compared to its position in H_2O. However, in some proteins such as in ribonuclease A the amide I band shifts to lower frequency by about 10 cm^{-1}.

It has been reported that 'β-bands' can occur even below 1620 cm^{-1} for some proteins [17]. For the homopolypeptide poly-L-lysine, the β-sheet band frequency is located at 1610 cm^{-1}. Variation in hydrogen-bonding strength as well as differences in transition dipole coupling in different β-strands can strongly influence the

position of this band. The assignment of bands in the region 1620–1640 cm^{-1} to β-sheet structure is consistent with both theoretical calculations and experimental studies on a large number of peptides and proteins [1–4,17,18,20]. However, bands in this region can sometimes be observed for proteins which are known to contain very little or no β-sheet structure, for example, myoglobin [21]. It has been suggested that bands in this region in the spectrum of myoglobin can be assigned to short extended segments that do not form β-sheet structures. More recently, model calculations by Torii and Tasumi [22] for the protein myoglobin indicate that bands below 1640 cm^{-1} may also arise from vibrational motions of α-helical structures. Band assignment in this region is further complicated by the possibility of β-turn structure also absorbing in this region [23]. Anti-parallel β-sheet structure can be identified by the presence of another band near 1670–1695 cm^{-1}. This component is normally weak and its precise assignment is often made difficult by the overlap of absorption from β-turn and unordered structures. For the homopolypeptide poly-L-lysine, a weaker band associated with high-frequency vibration of anti-parallel β-sheet structure is seen at 1680 cm^{-1}.

Experimental studies on proteins of known structure show that α-helical conformations gives rise to infrared absorption in the range 1650–1658 cm^{-1}. Fig. 3 shows the absorbance, second-derivative and deconvolved spectra of a membrane protein, photosynthetic reaction centre from *Rhodobacter sphaeroides*. The crystal structure of this protein has been determined [24] and has been shown to contain a predominantly α-helical structure. The amide I maximum at 1654–1656 cm^{-1} can be assigned to α-helical structure with some overlap from random coil structure. For samples in H_2O the amide I band arising from α-helical absorption can overlap with that arising from random coil structures. These two types of structures can normally be distinguished by carrying out experiments with ^2H_2O. In this media rapid hydrogen–deuterium exchange of the peptide N–H groups in the random coil structure results in a large shift of the amide I band to lower frequency. In ^2H_2O, proteins with a random coil arrangement show absorption near 1644 cm^{-1}. It should be noted that the absorption of α-helical structures can often be sensitive to the surrounding environment. For example, in membrane proteins the α-helical absorption occurs in the range 1656–1658 cm^{-1} [25], whereas in water-soluble proteins it normally occurs near 1650–1655 cm^{-1}. Furthermore, with highly solvent exposed helices [26] in ^2H_2O, the amide I band can shift to a low frequency of 1644 cm^{-1}. In some cases the amide I maximum for α-helical structure can occur at even lower frequencies [27,28]. For example, the amide I maximum for α-helical structure in the homopolypeptide poly-L-lysine occurs at 1638 cm^{-1} [28]. It is possible for the band frequency for an α-helical structure in a short, solvent exposed peptide to be significantly different from the absorption of helical structure which may be buried within a solvent inaccessible region of a highly folded globular protein. These factors should be taken into consideration when assigning bands in the spectra of peptides and proteins.

3.3. Protein secondary structure prediction

While second-derivative and deconvolution procedures are useful for qualitative analysis of proteins, it is necessary to develop methods for quantitative analysis of infrared spectra of proteins which provide information on fraction of different secondary structural elements. Although several methods have been developed for the quantitative analysis of infrared spectra of proteins the most popular methods are based on curve fitting and pattern recognition.

3.3.1. Curve fitting analysis

Byler and Susi [21] introduced the curve-fitting analysis of deconvolved amide I band contours as a linear combination of individual com-

ponent bands by iterative adjustment of heights, widths, and position of these bands. The fractional areas of the individual bands gives the fraction of secondary structure elements. This procedure provided a very good estimate of protein secondary structure and has been adopted by others with some modifications [9,17]. There are, however, a number of problems associated with the curve-fitting procedure. These include the need to assign all the bands within the amide I region. Another assumption in this method is that the molar absorptivities of the bands associated with different secondary structural elements are taken to be identical. Studies on poly-L-lysine have shown that the molar absorptivities of different secondary structures can be substantially different [27,28]. There is also a significant element of subjectivity associated with curve-fitting, such as the initial choice of input parameters. The results of fitting with a large number of adjustable parameters is not necessarily unique.

3.3.2. Approaches based on pattern recognition

This method avoids the need to deconvolve the amide I band and the assignment of the bands to different structures. The method employed is similar to that in use with CD and Raman spectroscopy. It involves the use of a calibration matrix of the infrared spectra of proteins of known X-ray structure. Mathematical tools such as factor analysis [29] and partial least squares analysis [30] are used in the analysis of the infrared spectra. The former method used the program CIRCOM [29]. Factor analysis of the infrared spectra of 18 proteins whose crystal structures are known from X-ray studies were used. A good correlation is observed between the infrared estimates and those calculated from X-ray data. This method gave standard errors of prediction of 3.9% for α-helices, 8.3% for β-sheets and 6.6% for turns. This compares well with values obtained using other spectroscopic techniques such as CD spectroscopy. Best results are obtained when only amide I bands of normalized infrared spectra of known proteins are used to construct the calibration matrix. Nevertheless, the approach has some drawbacks, such as the influence of overlap of absorbance from amino acid side chains in the amide I region. Dosseau and Pezolet [30] obtained the best result when both amide I and amide II regions were used to generate the calibration set. A number of other methods for quantitative estimation of protein secondary structure using FTIR has also been reported [31–34].

There are a number of problems still associated with all the methods that are in current use. These include normalization procedures which assumes equal absorptivity of amide bands irrespective of secondary structure variation. Furthermore, the technique encounters difficulties in cases where the spectral properties of the unknown protein lies outside the properties of the spectra within the calibration set. In these situations an incorrect estimation of the secondary structure may be possible. An additional common problem for both the curve-fitting method and the pattern recognition method is the overlap of amino acid side chains in the amide I region [35,36]. The different methods developed for the quantitative analysis infrared spectra are based upon soluble proteins. There are no methods available specifically catering for the analysis of peptides or membrane proteins. The pattern recognition method, for example, relies on calibration spectra derived from soluble proteins with well-characterised structures from X-ray crystallography.

4. FTIR spectroscopy of proteins in aqueous solutions

4.1. Investigation of protein structure and conformational changes

FTIR spectroscopy has been successfully used to investigate conformational changes in many different soluble proteins in aqueous solution (both H_2O and ^2H_2O). Here, we present FTIR

spectra of three proteins: Fig. 2 shows the FTIR spectra of a predominantly β-sheet containing water soluble protein in H_2O and ^2H_2O; FTIR spectrum of a predominantly α-helical membrane protein in H_2O is presented in Fig. 3; and finally Fig. 4 shows the spectra of a predominantly α-helical water-soluble protein in ^2H_2O.

Infrared spectroscopy is extremely sensitive to changes in hydrogen bonding and therefore potentially useful for obtaining subtle changes in protein structure. In order to detect very small changes in hydrogen bonding in proteins using FTIR spectroscopy, it is important to employ derivative, deconvolution and difference spectroscopy. These techniques have been utilised for investigation of conformational changes in a large number of water-soluble proteins. For example, conformational differences between the native and reactive-centre cleaved forms of α_1-antitrypsin [37] have been shown. Differences were observed in the spectra indicative of greater hydrogen bonding occurring with the reactive-centre cleaved form of the protein. The hydrogen–deuterium exchange within the polypeptide backbone is also found to be less for the cleaved form. This difference in the hydrogen–deuterium exchange was further supported by NMR spectroscopy [37].

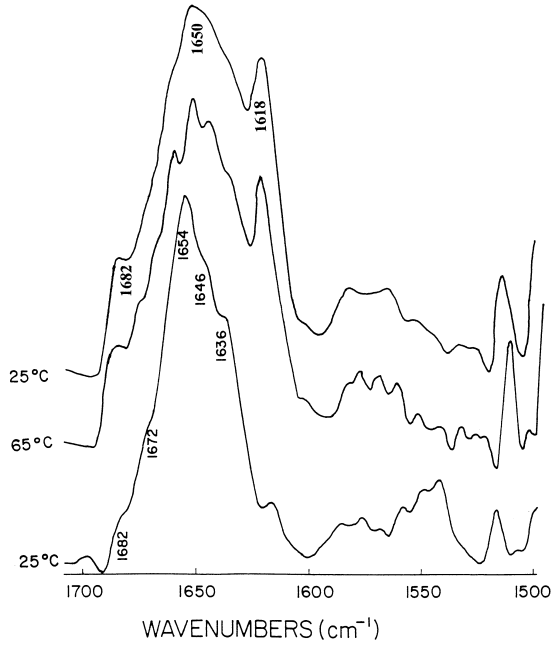

Fig. 4. Deconvolved spectra of pig citrate synthase recorded at 25°C, 65°C and after cooling back to 25°C. Thermal unfolding is characterised by the appearance of new bands near 1618 and 1682 cm^{-1}. The spectra were recorded for the protein in ^2H_2O.

A large number of metalloproteins such as calmodulin [26] and concanavalin A [38] have also been studied. In these studies changes in protein structure upon interaction of the proteins with metal ions were detected. The structure of iron binding proteins [39], human serum transferrin, rabbit serum transferrin and human lactoferrin have been characterised. The transferrins are involved in the regulation of iron homeostasis by the controlled release and binding of iron. Hydrogen–deuterium exchange measurements using the Amide II band indicate that the metal free proteins exchange to a greater extent than the corresponding metal-bound proteins. This difference was attributed to changes in flexibility/mobility in the proteins, since little or no change in secondary structure content is detected. The pathway of iron uptake into the vast majority of cell types involve the interaction of transferrin with its receptor. FTIR spectroscopy was used to investigate the solution structure of the water-soluble extracellular fragment of human transferrin receptor at extracellular and endosomal pH [40]. At extracellular pH, the FTIR quantitative analysis indicates the protein to consist of 56% α-helix, 19% β-sheet and 14% turns. Upon acidification to endosomal pH the α-helical content of the protein is reduced and the β-sheet content increased by nearly 10%.

FTIR spectroscopy has been applied to gain insights into the structure of a number of different complement proteins. Many of these proteins are extremely large in size and very little is known about their structure. Proteins that we have studied include factor H [16], properdin [41] and factor B [42]. Factor H is a regulatory component of the complement cascade of immune defense and functions as a cofactor for the enzyme factor I in the breakdown of C3b of complement to form iC3b. It is composed entirely of 20 repetitive sequences of polypeptide

known as short consensus repeats (SCRs), where each sequence is 61 residues long. FTIR spectroscopy was used to investigate the secondary structure of human complement factor H in H_2O and ^2H_2O buffers [16]. The amide I band components are consistent with the existence of an extensive antiparallel β-strand secondary structure. The hydrogen–deuterium exchange properties of factor H as measured in ^2H_2O are rapid and lead to an estimate of N–H proton non-exchange that is comparable with those for small globular proteins. The application of FTIR spectroscopy to factor H was complemented by comparisons with secondary structure predictions of the main-chain conformations in the SCR. The predictive method was applied simultaneously to compute the average prediction for all 101 SCR sequences that are currently available in order to improve the precision of the calculations. A structural model based on antiparallel β-strands was proposed [16]. The large size of factor H prevents its structural characterisation using NMR spectroscopy. However, one of the SCR domains of factor H was expressed in a yeast secretion system and its solution structure determined using NMR spectroscopy [43]. The results of this study were in good agreement with our conclusions [16]. By linking FTIR spectroscopy with structure prediction methods, one can obtain a more reliable picture of the structural organisation of a protein molecule.

4.2. Investigation of protein thermostability

Several factors are involved in the destabilization of proteins. These are mainly pH, extreme salt conditions, ionic strength, denaturant such as urea and guanidine, and finally temperature. Here only the effect of temperature on protein stability will be discussed.

Understanding the molecular basis of thermostability in proteins remains an important challenge in biochemistry. Once the mechanism of thermostability is understood, the industrially important mesophilic enzymes will be engineered to form thermophilic enzymes which have biotechnological applications. FTIR spectroscopy is a particularly useful technique for monitoring temperature induced changes in protein structure. We have been using this technique to compare the thermostability of different proteins in order to identify the factors that are responsible for conferring greater stability to thermophilic proteins compared to their mesophilic counterparts. Citrate synthase is one protein that we have been investigating. This enzyme can be obtained from both mesophilic and thermophilic organisms. Fig. 4 shows the deconvolved FTIR spectra of a pig citrate synthase (PCS) recorded in ^2H_2O. Spectra of the protein recorded at 25°C (native), 65°C (denatured) and after coiling back to 25°C are presented. As can be seen from the figure, thermal denaturation leads to a loss of native secondary structure and the appearance of new bands near 1618 cm^{-1}. The latter band is close to the frequency associated with β-sheet structure. However, it occurs at a significantly lower frequency compared with β-sheet structure in native proteins. This structure has been attributed to aggregated β-sheet structures. When polypeptide chains unfold at high temperatures, they interact with each other forming strongly hydrogen-bonded aggregated structures that give rise to distinct bands in the region 1620–1615 cm^{-1}. These bands are good diagnostic tools for monitoring thermal unfolding of proteins. The irreversible nature of the denaturation process is indicated by the fact that these bands remain after cooling the sample back to 25°C. The irreversible change in the spectrum indicates that the PCS protein undergoes irreversible thermal denaturation. Precise information regarding the thermal unfolding temperature, and about the unfolding behaviour of specific secondary structural elements can also be determined using FTIR spectroscopy. For example, the temperature dependent variation in the intensities of the α-helix and ‘aggregation β-band’ is presented in Fig. 5. As seen from this figure, with progressive increase in temperature, the intensity of

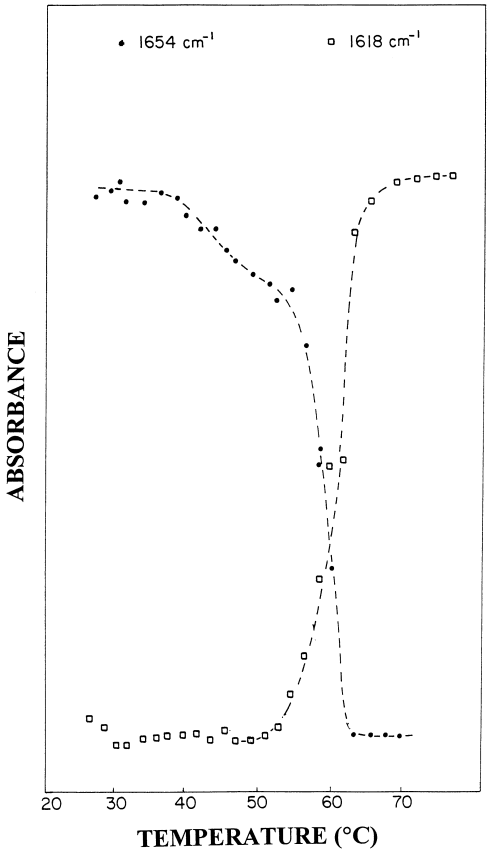

Fig. 5. A plot of the temperature induced changes in the intensity of Amide I bands at 1654 cm^{-1} (associated with α-helical structure) and 1618 cm^{-1} (aggregated β-sheet structure) for pig citrate synthase in ^2H_2O. The loss in intensity of the 1654 cm^{-1} coincides with the increase in intensity of the 1618 cm^{-1} band.

the α-helix decreases. Simultaneously, the intensity of the bands due to the aggregated β-sheet structure (1618 cm^{-1}), which appear upon thermal denaturation, increases. According to the mid-point transition of the α-helix and 'aggregated β-band', thermal unfolding temperature of PCS occurs at 59°C.

We have also used FTIR spectroscopy to investigate temperature-induced structural changes in a large number of other proteins such as albumin, immunoglobulin G, fibrinogen, lysozyme, α-lactalbumin, and ribonuclease S. In order to analyze the data, a new method was developed in which the data were analyzed globally with the aid of a spectral model [44]. Seven or eight bands were sufficient to fit the full data set of spectra ranging from 1420 to 1760 cm^{-1} with a root mean square error of 1–2% of the maximum. Subsequently, the estimated band amplitude curves which showed a sigmoidal progression with increasing temperature were fitted with a two-state thermodynamic model. In this way, information on structural changes as well as on the thermal stability of the proteins was obtained. We observed that an enhanced hydrogen–deuterium exchange occurred at temperatures well below the unfolding of the secondary structure. This was explained as being due to a change in tertiary structure leading to enhanced solvent accessibility. In all the proteins investigated, with the exception of ribonuclease S, an intermolecular β-sheet band indicative of aggregation appeared concomitant with the denaturation of the proteins.

FTIR spectroscopy was used to investigate the thermal stability of the water-soluble extracellular fragment of human transferrin receptor at extracellular and endosomal pH [40]. At the extracellular pH (7.4), the mid-point temperature of thermal denaturation for the loss of secondary and tertiary structure, and the formation of aggregated structures (1617 and 1683 cm^{-1}), is 71°C. At endosomal pH (5.6) this temperature is reduced by approximately 15°C.

Thermal denaturation of proteins such as calmodulin, transferrin and lactoferrin were also studied in the presence and absence of metal ions. The removal of metal ions leads to thermal destabilization of these proteins. FTIR analysis shows that a distinct loss of protein secondary structure occurs at the transition temperature shown by differential scanning calorimetry [39]. According to our studies thermostability of metalloproteins are significantly enhanced when they are bound to metal ions.

4.3. Structure and stability of recombinant proteins

Advances in recombinant DNA technology is rapidly generating a large number of proteins for medical and industrial applications. How-

ever, the process which lead to the correct folding of a denatured or newly-synthesised protein is poorly understood. Insight into the underlying principles and conditions which promote protein folding will help in the production and purification of proteins especially those produced in bacteria and yeast as inactive, misfolded, or aggregated proteins. Also little information is available as to how other factors such as variation in glycosylation influences the structure and stability of such proteins. At present most recombinant proteins intended for human therapy are glycoproteins. In order to successfully use recombinant proteins for medical and biotechnological applications, it is necessary to characterise their structure and stability under various conditions. FTIR spectroscopy is ideally suited for such studies as it can be used for recording spectra of proteins in virtually all types of environments. Furthermore, the ability of FTIR spectroscopy to detect small changes in protein structure makes it useful for rapid comparison of the structure and stability of wild-type proteins with their site-directed mutant form.

We have been using FTIR spectroscopy to characterise the structure of a number of recombinant proteins. One such protein is the Streptococcal antigen I/II (SA I/II). This is a cell surface adhesin of *Streptococcus* mutants which mediates attachment of the bacterium to teeth [45]. The deduced amino acid sequence of SA I/II (1561 residues) includes a leader sequence (residues 1–38), four tandem repeats of an Ala-rich sequence (residues 121–447), three tandem repeats of a Pro-rich sequence (residues 839–955) and a consensus recognition sequence for attachment to the bacterial cell wall. The conformation of the native and recombinant form of this adhesin has been studied by FTIR spectroscopy [45]. The studies shows that there are minor variations in thermal stability of the two forms of the protein. The reason for the variation in thermal stability is not clear but it may be due to some misfolding of the recombinant protein. These types of studies indicate that mere testing of functional activity of a recombinant protein is not sufficient to indicate that it has the same properties as the native version of the protein.

Recently, we have used a combination of secondary structure prediction methods and FTIR spectroscopy to investigate the structure of HPV-16 E7 protein. The E7 and E6 proteins are the main oncoproteins of human papillomavirus types 16 and 18 (HPV-16 and HPV-18). The structure of these proteins are unknown. Secondary structure prediction was carried out using the GOR I/III, Chou-Fasman, SAPIENS and PHD methods [46]. The results obtained using these methods were compared with data obtained for recombinant HPV-16 E7 in H_2O and ^2H_2O buffers using FTIR spectroscopy. Quantitative analysis of the FTIR spectra of the recombinant E7 protein, using the factor analysis method [29], revealed that it contained similar amounts of α-helix and β-sheet structures [46]. This is in good agreement with the results of the theoretical prediction methods [46] of α-helix and β-sheet structures in E7. The results of the spectroscopic and prediction methods were used to propose a topology model for the structure of the zinc-binding motif of HPV-16 E7 [46].

5. FTIR spectroscopy of proteins in organic solvents and membranes

Often structural studies are carried out on proteins far removed from their natural environment. This can be for a deliberate purpose (for example, to monitor how the structure of a protein is influenced by its environment, and if this influences its function) or because one has no choice (since no suitable technique can be used for the structural analysis in the biologically relevant media). As an example of the latter situation, structure of membrane proteins and peptides are often characterised in detergents or in organic solvents rather than phospholipid membranes. These systems are used to

mimic biological membranes as it is not possible to investigate the structure of the protein in phospholipid membranes using most techniques. The biotechnology industry has great interest in characterising the structure of proteins in non-conventional media such as in organic solvents. This is because some enzymes have been reported to display higher activity in organic solvents compared to aqueous media. FTIR spectroscopy is one of the techniques that can be applied to study peptides and proteins in all these environments. This is illustrated in Fig. 6 which shows the FTIR spectra of a synthetic peptide corresponding to the voltage-sensor of the *Shaker* potassium channel. Spectra of the peptide were recorded in an organic solvent (trifluoroethanol), phospholipid membranes and aqueous solution. The amide I maxima for this peptide is different for each of these media indicating that secondary structure is highly dependent on properties of the surrounding environment. Thus in aqueous solution (^2H_2O) it has a band position near 1641 cm^{-1} indicative of random coil structure. An increase in helicity is seen in phospholipid membranes where the band position is centred at 1651 cm^{-1}. Even a higher helical content is seen in TFE where the band position is located at 1654 cm^{-1}. In recent years FTIR spectroscopy has been successfully used in the structural characterisation of a large range of membrane associated peptides and proteins (for recent reviews see Refs. [4,20,47]).

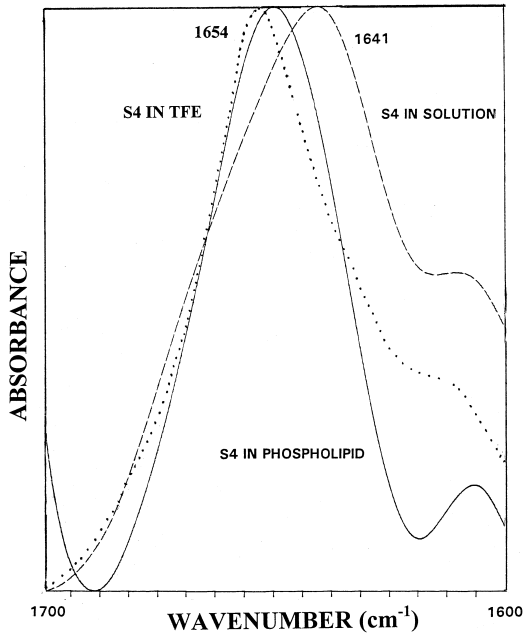

Fig. 6. Absorbance spectra of a synthetic peptide in ^2H_2O (---), phospholipid membranes (—) and in trifluoroethanol (TFE) solvent (…). Spectra were recorded at 25°C. Variations in the frequency of the amide I band illustrates environment dependent changes in the structure of the peptide.

6. Structural characterisation of immobilised proteins

Immobilization of proteins and peptides using different procedures is becoming increasingly important for a number of medical and biotechnological applications. However, there is a distinct lack of techniques that can be used to characterise the conformation of such immobilised proteins. FTIR spectroscopy is one of the few techniques that can provide information in this area. The secondary structure of a chemically cross-linked protein can be readily characterised using FTIR spectroscopy as long as there is no major overlap of absorbance, in the Amide I region, from the non-protein component. It may be possible to reduce spectral overlap by isotopically labelling either the protein or the cross-linking agent. Such isotope-edited FTIR spectroscopy has been successfully used in the investigation of protein–protein interactions [48].

FTIR spectroscopy has been used to characterise the secondary structure and orientation of immobilised peptide monolayers [49]. The peptides were immobilised on solid supports such as gold. FTIR spectroscopy has also been used to characterise the structure of proteins that have been chemically immobilised. For example, Jiang et al. [50] chemically immobilised a mutant myoglobin molecule on a silane derivatised substrate. The conformation of the immobilised protein was characterised using FTIR spectroscopy. The results revealed that the

structure of the immobilised protein was similar to that of the native myoglobin molecule.

Studies have also been performed on proteins that have been lyophilized in the presence of multifunctional ligands [51]. The structure of such proteins were characterised in anhydrous media after the removal of the ligands. Proteins exposed to these treatments show significant changes in their β-sheet and α-helical content compared to their native counterparts [51]. In another study, FTIR spectroscopy was used to quantitatively examine the secondary structure of some dozen different proteins in lyophilised forms [52]. This revealed that lyophilisation leads to a marked increase in the percentages of β-sheets and a parallel decrease in the content of α-helical and disordered structures.

7. Conclusion and future prospects

FTIR spectroscopy is a versatile tool that can be used to obtain information on different aspects of protein structure in a wide range of environments. Methods for quantitative estimation of protein secondary structure using this technique are progressively improving. Combining FTIR spectroscopy with secondary structure prediction methods provides a more reliable structure determination than is possible by using these methods individually. Second-derivative, deconvolution and difference spectroscopy are extremely useful for detection of subtle changes in protein structure. Information on the stability of different secondary elements in a protein can be monitored as a function of temperature and the precise unfolding temperature determined. We have also shown how the conformation of a peptide changes drastically when it is transferred from H_2O to phospholipid membranes and trifluoroethanol. The ability to characterise structure of proteins in a large range of environments makes FTIR spectroscopy a very important analytical tool. This is especially valuable for the biotechnology and pharmaceutical industry where there is a need for monitoring structure and stability of proteins in diverse, and often non-conventional, environments. With advances in both FTIR technology and molecular biology, the usefulness of FTIR spectroscopy for protein structure characterisation will further improve. As an example, the ability to produce isotopically labelled proteins has led to the development of a new method for investigating protein–protein interaction based on isotope-edited FTIR spectroscopy [48]. Similarly, advances in FTIR instrumentation has paved the way for time-resolved FTIR spectroscopy of protein structure and stability [53,54].

Acknowledgements

We would like to thank British Council-Academic Link Programme.

References

[1] P.I. Haris, D. Chapman, Trends Biochem. Sci. 17 (1992) 328.
[2] W.K. Surewicz, H.H. Mantsch, D. Chapman, Biochemistry 32 (1993) 299.
[3] P.I. Haris, D.C. Lee, D. Chapman, Biochim. Biophys. Acta 874 (1986) 255.
[4] J.L.R. Arrondo, F.M. Goni, in A. Watts (Ed.), Protein–Lipid Interactions, Elsevier, Amsterdam, 1993, p. 321.
[5] M.S. Braiman, K.J. Rothschild, Annu. Rev. Biophys. Biophys. Chem. 17 (1988) 541.
[6] D.V. Waterhous, W.C. Johnson, Biochemistry 33 (1994) 2121.
[7] M. Jackson, H.H. Mantsch, Biochim. Biophys. Acta. 1118 (1992) 139.
[8] U.P. Fringeli, H.H. Gunthard, in: E. Grell (Ed.), Membrane Spectroscopy, Vol. 31, Springer-Verlag, Berlin, 1981, p. 27.
[9] E. Goormaghtigh, R. Brasseur, P. Huart, J.M. Ruysschaert, Biochemistry 26 (1987) 1789.
[10] N.J. Harrick, Internal Reflection Spectroscopy, Wiley-Interscience, New York, 1967.
[11] J.R. Powell, F.M. Wasacz, R.J. Jakobsen, Appl. Spectrosc. 40 (1986) 339.
[12] F. Dosseau, M. Therrien, M. Pezolet, Appl. Spectrosc. 43 (1989) 538.
[13] H.H. Mantsch, W.K. Surewicz, A. Muga, D.J. Moffatt, H.H. Casal, in: D.G. Cameron (Ed.), Proc. of the 7th Int. Conf. on Fourier Transform Spectroscopy, SPIE, Vol. 1145, 1989, p. 580.
[14] J.K. Kauppinen, D.J. Moffatt, H.H. Mantsch, D.G. Cameron, Appl. Spectrosc. 35 (1981) 271.
[15] D.G. Cameron, D.J. Moffatt, Appl. Spectrosc. 41 (1987) 539.

[16] S.J. Perkins, P.I. Haris, R.B. Sim, D. Chapman, Biochemistry 27 (1988) 4004.
[17] W.K. Surewicz, H.H. Mantsch, Biochem. Biophys. Acta 952 (1988) 115.
[18] H. Susi, D.M. Byler, Methods. Enzymol. 150 (1986) 290.
[19] W. Mantele, Trends in Biochem. Sci. 18 (1993) 197.
[20] L.K. Tamm, S.A. Tatulian, Q. Rev. Biophys. 30 (1997) 365.
[21] D.M. Byler, H. Susi, Biopolymers 25 (1986) 469.
[22] H. Torii, M. Tasumi, J. Chem. Phys. 96 (1992) 3379.
[23] H.H. Mantsch, A. Perczel, M. Hollosi, G.D. Fasman, Biopolymers 33 (1993) 201.
[24] J.P. Allen, G. Feher, T.O. Yates, H. Komiya, D.C. Rees, Proc. Natl. Acad. Sci. 84 (1987) 6162.
[25] P.I. Haris, D. Chapman, Biochem. Soc. Trans. 17 (1989) 161.
[26] M. Jackson, P.I. Haris, D. Chapman, Biochemistry 30 (1991) 9681.
[27] Y.N. Chirgadze, B.V. Shestopalov, S.Y. Venyaminov, Biopolymers 12 (1973) 1337.
[28] M. Jackson, P.I. Haris, D. Chapman, Biochim. Biophys. Acta 998 (1989) 75.
[29] D.C. Lee, P.I. Haris, D. Chapman, R.C. Mitchell, Biochemistry 29 (1990) 9185.
[30] F. Dosseau, M. Pezolet, Biochemistry 29 (1990) 8771.
[31] V. Baumruk, P. Pancoska, T.A. Keiderling, J. Mol. Biol. 259 (1996) 774.
[32] K. Rahmelow, W. Hubner, Anal. Biochem. 241 (1996) 5.
[33] A. Dong, P. Huang, W.S. Caughey, Biochemistry 29 (1990) 3303.
[34] T.A. Keiderling, B.L. Wang, M. Urbana, P. Pancoska, R.K. Dukor, Faraday Discussions 99 (1994) 263.
[35] Y.N. Chirgadze, O.V. Federov, N.P. Trushina, Biopolymers 14 (1975) 679.
[36] S.Y. Venyaminov, N.N. Kalnin, Biopolymers 30 (1991) 1243.
[37] P.I. Haris, D. Chapman, R.A. Harrison, K.F. Smith, J. Perkins, Biochemistry 29 (1990) 1377.
[38] J. Alvarez, P.I. Haris, D.C. Lee, D. Chapman, Biochim. Biophys. Acta. 916 (1987) 5.
[39] J.M. Hadden, M. Bloemendal, P.I. Haris, S.K.S. Srai, D. Chapman, Biochim. Biophys. Acta. 1205 (1994) 59.
[40] J.M. Hadden, M. Bloemendal, P.I. Haris, I.H.M. Stokkum, D. Chapman, S.K.S. Srai, FEBS Lett. 350 (1994) 235.
[41] S.J. Perkins, A.S. Nealis, P.I. Haris, D. Chapman, D. Goundis, K.B.M. Reid, Biochemistry 28 (1989) 7176.
[42] S.J. Perkins, K.F. Smith, A.S. Nealis, P.I. Haris, D. Chapman, J. Mol. Biol. 228 (1992) 1235.
[43] P.N. Barlow, M. Baron, D.G. Norman, A.J. Day, A.C. Willis, R.B. Sim, I.D. Campbell, Biochemistry 30 (1991) 997.
[44] I.H.M. van Stokkum, H. Linsdell, J.M. Hadden, P.I. Haris, D. Chapman, M. Bloemendal, Biochemistry 34 (1995) 10508–10518.
[45] C.G. Kelly, J.M. Hadden, S. Todryk, P.I. Haris, P. Morgan, T.J. Mitchell, Immunology 89 (1996) FF369.
[46] C. G, P. Ullman, I. Haris, D.A. Galloway, V.C Emery, S.J. Perkins, Biochem. J. 319 (1996) 229.
[47] P.I., Haris, D. Chapman, in: P.I. Haris, D. Chapman (Eds.), Biomembrane Structures, IOS Press, Amsterdam, 1998, pp. 134.
[48] P.I. Haris, G.T. Robillard, A.A. Van Dijk, D. Chapman, Biochemistry 31 (1992) 6279.
[49] M. Boncheva, H. Vogel, Biophys. J. 73 (1997) 1056–1072.
[50] M. Jiang, B. Nolting, P.S. Stayton, S.G. Sligar, Langmuir 12 (1996) 1278–1283.
[51] P. Mishra, K. Griebeno, A.M. Klibanov, Biotechnol. Bioeng. 52 (1996) 609–614.
[52] K. Griebenow, A.M. Klivanov, Proc. Natl. Acad. Sci. 92 (1995) 10969–10976.
[53] D. Reinstadler, H. Fabian, J. Backmann, D. Naumann, Biochemistry 35 (1996) 15822.
[54] R. Gilmanshin, S. Williams, R.H. Callender, W.H. Woodruff, R.B. Dyer, Biochemistry 36 (1997) 15006.