Application of far-infrared spectroscopy to the structural identification of protein materials

Yanchen Han, Shengjie Ling, Zeming Qi, Zhengzhong Shao and Xin Chen

Although far-infrared (IR) spectroscopy has been shown to be a powerful tool to determine peptide structure and to detect structural transitions in peptides, it has been overlooked in the characterization of proteins. Herein, we used far-IR spectroscopy to monitor the structure of four abundant non-bioactive proteins, namely, soybean protein isolate (SPI), pea protein isolate (PPI) and two types of silk fibroins (SFs), domestic Bombyx mori and wild Antheraea pernyi. The two globular proteins SPI and PPI result in broad and weak far-IR bands (between 50 and 700 cm$^{-1}$), in agreement with those of some other bioactive globular proteins previously studied (lysozyme, myoglobin, hemoglobin, etc.) that generally only have random amino acid sequences. Interestingly, the two SFs, which are characterized by a structure composed of highly repetitive motifs, show several sharp far-IR characteristic absorption peaks. Moreover, some of these characteristic peaks (such as the peaks at 260 and 428 cm$^{-1}$ in B. mori, and the peaks at 245 and 448 cm$^{-1}$ in A. pernyi) are sensitive to conformational changes; hence, they can be directly used to monitor conformational transitions in SFs. Furthermore, since SF absorption bands clearly differ from those of globular proteins and different SFs even show distinct adsorption bands, far-IR spectroscopy can be applied to distinguish and determine the specific SF component within protein blends.

1. Introduction

Spectroscopic analysis at far-infrared (IR) frequencies (between 200 and 700 cm$^{-1}$) began to receive great attention since the characterization of polyamides and peptides in the 1960s and 1970s. Compared to mid-IR frequencies, far-IR is more sensitive to vibrational modes from peptide skeletons and hydrogen bonds. Thus, far-IR spectroscopy is more suitable for investigating highly-ordered structures, such as fibrillar formation. In addition, it can also provide information on protein dynamics that are complementary to nuclear magnetic resonance and fluorescence-based analyses. A series of spectroscopic characterizations of polyamides and peptides have been disclosed and some features were assigned to specific vibrational modes. For example, in N-methylacetamide, the peaks at around 100 cm$^{-1}$ and 201 cm$^{-1}$ have been assigned to the CO···HN hydrogen bonding vibration and C-N torsional vibration, respectively. The analysis of crystalline polyamides, including polyglycine, polyalanine and other polypeptides variants, further revealed that the far-IR absorption bands of the peptides are associated with their conformation. For instance, the peak at 380 cm$^{-1}$ has been assigned to the $\alpha$-helix, while the peak at 440 cm$^{-1}$ is indicative of $\beta$-sheets.

Unfortunately, peak assignment of far-IR spectra from polyamides does not lead to a realistic interpretation of the protein spectra. Most proteins, such as lysozyme, $\beta$-lactoglobulin, myoglobin, hemoglobin, serum albumin, ribonuclease, chymotrypsinogen, and subtilisin, only exhibit weak and broad absorption bands in the far-IR spectral region. For this reason, in order to disclose their far-IR band information, difficult and time-consuming theoretical calculations (e.g., density functional theory$^{11,12}$ and simulations (e.g., atomic molecular dynamics simulation$^{13}$) need to be applied to interpret experimental data. Moreover, these two methods are only practically applicable for small-size proteins due to time- and dimensional-limited computer processing power. As a result, far-IR has long been overlooked in the characterization of large-size proteins.

Nevertheless, for practical applications, large-size proteins have greater potential than small-size proteins due to their superior mechanical properties and tuneable biodegradability. For example, mulberry and non-mulberry silk fibroins (SFs) have shown promising biomedical, optical, and environmental applications.$^{14-16}$ Soy protein isolate (SPI) and pea protein isolate (PPI) are two of the most abundant plant proteins that have been extensively used to replace traditional plastics to produce life utensils and to fabricate biomaterials.$^{17-19}$ To tune

---

*State Key Laboratory of Molecular Engineering of Polymers, Department of Macromolecular Science, Laboratory of Advanced Materials, Fudan University, Shanghai, 200433, People’s Republic of China. E-mail: chenx@fudan.edu.cn

School of Physical Science and Technology, ShanghaiTech University, Shanghai, 201210, People’s Republic of China. E-mail: lingshj@shanghaitech.edu.cn

National Synchrotron Radiation Laboratory, University of Science and Technology of China, Hefei, 230029, People’s Republic of China

Received 3rd February 2018, Accepted 2nd April 2018
DOI: 10.1039/c8cp00802g
rsc.li/pccp

Cite this: Phys. Chem. Chem. Phys., 2018, 20, 11643
the physical properties to diverse applications, one type of protein usually needs to be blended with other proteins to generate protein composites.\textsuperscript{20} For instance, the addition of keratin or elastin to SF was shown to possibly accelerate the process of cell proliferation, differentiation, adhesion, and migration.\textsuperscript{14,20–22} The integration of SF and amyloid fibrils can lead to membranes with tunable mechanical properties and pore size by selecting etch amyloid components from the hybrid membranes.\textsuperscript{23}

Similar to other polymer blends, the mass ratio and the structure of each component in protein blends are critical to the mechanical properties and the function of the final materials. However, it is still a significant challenge to distinguish a single element from two (or several) protein components in protein blends. Mid-IR\textsuperscript{23,24} and Raman spectroscopies\textsuperscript{25,26} are two of the most used methods to analyze the structure of proteins; however, the spectra of most of the proteins exhibit very similar amide I (1700–1600 cm\textsuperscript{-1}, C=O stretching vibration), amide II (1600–1500 cm\textsuperscript{-1}, C-N stretching), and amide III (1300–1200 cm\textsuperscript{-1}, H\textsuperscript{3} bending and C-N stretching) bands.\textsuperscript{27} The reason for this is that all of these amide bands directly depend on the skeletal vibration of the protein backbone, which does not differ much among proteins.

Herein, we investigated the far-IR absorbance of different proteins and showed that in this spectral region, fibrous SF proteins can be clearly distinguished from globular SPI and PPI proteins. Far-IR spectroscopy can also be used to distinguish between the two different SFs. Moreover, in protein blends, we can estimate the ratio between the different protein components. Remarkably, some far-IR bands are even sensitive to the secondary structure of SFs; thus, they can be used to probe conformational transitions in SFs, which could possibly allow measurement and tuning of the properties of the resultant SF-based materials.

2. Experimental section

2.1 Preparation of regenerated Bombyx mori SF films

B. mori silk cocoons were purchased from Jiangsu Province, China. Cocoons were degummed twice in 0.5% (w/w) NaHCO\textsubscript{3} solution at 100 °C for 30 min and then washed with distilled water and air dried at room temperature. The dried degummed B. mori silk fibers were dissolved in 9.3 mol L\textsuperscript{-1} LiBr aqueous solution. After dialysis against deionized water for 3 days at room temperature, the B. mori SF aqueous solution had a concentration of approximately 4% (w/w). The as-prepared regenerated B. mori SF solution was then diluted to 1% (w/w) and cast onto a polyethylene plate to form a thin film with thickness of approximately 5 μm at 25 °C and 40% relative humidity.

2.2 Preparation of regenerated Antheraea pernyi SF films

A. pernyi silk cocoons were purchased from Shandong Province, China. Cocoons were degummed twice in 0.5% (w/w) Na\textsubscript{2}CO\textsubscript{3} solution at 100 °C for 30 min and then washed with distilled water and allowed to air dry at room temperature. The dried degummed A. pernyi silk fibers were dissolved in 7.5 mol L\textsuperscript{-1} Ca(NO\textsubscript{3})\textsubscript{2} aqueous solution at 90 °C. The resultant solution was dialyzed against deionized water for 3 days at 4 °C to yield approximately 2% (w/w) regenerated A. pernyi SF aqueous solution. The as-prepared regenerated A. pernyi SF solution was then diluted to 1% (w/w) and cast onto a polyethylene plate to form a thin film with thickness of approximately 5 μm at 25 °C and 40% relative humidity.

2.3 Preparation of SPI and PPI films

SPI (protein content >90%) was bought from Shenyuan Food Co., Ltd, China. The SPI powder was dissolved in 6 mol L\textsuperscript{-1} guanidine hydrochloride aqueous solution and stirred at room temperature for 3 h. Then, 25 mmol L\textsuperscript{-1} dithiothreitol was added into the solution in order to break disulfide bonds. The resultant solution was dialyzed against NaOH aqueous solution (pH = 10) for 2 days and then against deionized water for another day at room temperature. The as-prepared SPI solution was then diluted to 1% (w/w) and cast onto a polyethylene plate to form a thin film with thickness of approximately 5 μm at 25 °C and 40% relative humidity. The PPI powder was purchased from Staerkle & Nagler AG, Switzerland and films were prepared according to the same procedure as SPI films.

2.4 Preparation of A. pernyi SF/B. mori SF blend films

Different mass ratios of 1% (w/w) B. mori SF and A. pernyi SF solutions were mixed to form homogeneous solutions. Mixtures were then cast onto polyethylene plates and allowed to dry at approximately 25 °C and 40% relative humidity to form A. pernyi SF/B. mori blend films with an approximate thickness of 5 μm.

2.5 Preparation of A. pernyi SF/SPI blend films

Different mass ratios of 1% (w/w) SPI and A. pernyi SF solutions were mixed to form homogeneous solutions. The mixture was then cast onto polyethylene plates and allowed to dry at approximately 25 °C and 40% relative humidity to form A. pernyi SF/SPI blend films with an approximate thickness of 5 μm.

2.6 Far-IR spectroscopy of protein films

Far-IR spectra were acquired at the National Synchrotron Radiation Laboratory (NSRL, Hefei, China) in the range of 50 and 700 cm\textsuperscript{-1} using a Bruker 66v/s Fourier transform (FTIR) spectrometer. In order to eliminate the absorption contribution from atmospheric water vapour (water molecules absorption occurs between approximately 80 and 200 cm\textsuperscript{-1} due to the bending and stretching vibration of hydrogen bonds\textsuperscript{3}), the entire light path of the instrument was continuously evacuated with a Nidec rotary vacuum pump. For each measurement, 256 interferograms were co-added and transformed by employing a Genzel–Happ apodization function to yield spectra with a nominal resolution of 4 cm\textsuperscript{-1}. All the far-IR spectra shown in this article were original spectra without any post-treatment.

3. Results and discussion

3.1 Far-IR spectra of four pure protein films

Fig. 1 shows the far-IR spectra of SPI and PPI films, which present almost the same far-IR fingerprint characterized by
four broad and weak bands at 660, 622, 330 and 150 cm\(^{-1}\). The far-IR spectra are also similar to those of other globular proteins, such as myoglobin, hemoglobin, and lysozyme.\(^6\) The similarity between SPI and PPI studied herein with myoglobin, haemoglobin and lysozyme reported in the literature is that they all lack highly repetitive segments in their amino acid sequences. For example, glycamin (11S component), the main component of SPI consists of highly disordered hydrophobic (Ala, Val, Ile, Leu and Phe) and hydrophilic amino acid residues (Lys, His, Agr, Asp and Glu).\(^28,29\) Therefore, we ascribe the broad bands in the far-IR spectra observed in a multitude of globular proteins to the population of overlapping discrete vibrational modes from random amino acid sequences.

Contrary to SPI and PPI, \(B.\) mori SF is well-known to possess highly repetitive core sequences with alternating hydrophilic and hydrophobic segments and flanked by highly conserved shorter terminal domains (N- and C-termini).\(^30\) The hydrophobic motifs, like well-ordered B-sheet domains, are made up of (GAGAGS)\(_n\) (\(n = 1–11\)), while the hydrophilic motifs are formed by tyrosine-rich sequences.\(^30\) Fig. 2 shows the far-IR spectra of \(B.\) mori SF films, which are very different from SPI and PPI spectra shown in Fig. 1. In particular, the as-cast \(B.\) mori SF film shows two sharp peaks at 553 and 335 cm\(^{-1}\) and two weak peaks at 387 and 371 cm\(^{-1}\) (curve a). In order to identify the bands that are sensitive to conformation changes, we used 70% ethanol aqueous solution to induce a conformation transition of SF from helix/random coil to B-sheet.\(^31,32\) Compared with the as-cast film, after 70% ethanol treatment (curve b), two additional peaks are clearly observed at 260 and 428 cm\(^{-1}\), while the two peaks at 371 and 387 cm\(^{-1}\) disappear. These results prove that it is possible to distinguish the detailed conformation of \(B.\) mori SF from the far-IR spectrum. Additionally, because of the high content of the repeated GAGAGS motif\(^33,34\) in SF sequences, we can assign the far-IR absorption of \(B.\) mori SF according to the analysis of poly(\(-\)alanine) and poly(\(-\)alanylglycine).\(^3\) Our results confirm that this band can be assigned to helical conformation (silk I) in \(B.\) mori SF. The detailed band assignment in \(B.\) mori SF between 700 and 50 cm\(^{-1}\) is summarized in Table 1.

![Fig. 1 Far-IR spectra of SPI (a) and PPI (b) films.](image1)

![Fig. 2 Far-IR spectra of \(B.\) mori SF films: (a) \(B.\) mori SF film as cast; (b) \(B.\) mori SF film treated with 70% ethanol aqueous solution at room temperature for 24 h.](image2)

Table 1. The far-IR band assignment of \(B.\) mori SF\(^a\)

<table>
<thead>
<tr>
<th>Wavenumber (cm(^{-1}))</th>
<th>Assignment</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>553</td>
<td>C=O opb, C(^\text{CN}) def</td>
<td>1</td>
</tr>
<tr>
<td>428</td>
<td>C(^{\text{b}}) bend 1, NC(^{\text{C}})C def, C(^{\text{b}}) bend 2, NH opb; (\beta)-sheet</td>
<td>1</td>
</tr>
<tr>
<td>387</td>
<td>CO ipb, NH opb, C(^{\text{b}})C def, NC(^{\text{C}})C def</td>
<td>1</td>
</tr>
<tr>
<td>371</td>
<td>C(^{\text{b}}) bend 2, NC(^{\text{C}})C def, C(^{\text{b}})CN def; helix (silk I)</td>
<td>35</td>
</tr>
<tr>
<td>335</td>
<td>C(^{\text{b}}) bend 2</td>
<td>1</td>
</tr>
<tr>
<td>260</td>
<td>C(^{\text{b}})C(^{\text{b}}) tor, CNC(^{\text{b}}) def, H(-)O str; (\beta)-sheet</td>
<td>1</td>
</tr>
<tr>
<td>125</td>
<td>H(-)O str</td>
<td>1</td>
</tr>
</tbody>
</table>

\(^a\) Abbreviations: opb = out-of-plane angle bend, def = deformation, ipb = in-plane angle bend, tor = torsion, str = stretch.
as-cast A. pernyi SF films have also been previously discussed based on solid-state NMR, XRD, Raman and FTIR spectroscopies.\textsuperscript{23–26,38,39} Therefore, we are confident with the \( \alpha \)-helix assignment of these peaks (658, 526, and 373 cm\(^{-1}\)). After the 70% ethanol aqueous solution treatment, two additional peaks appear at 448 and 245 cm\(^{-1}\) (Fig. 3, curve b) and the peak at 620 cm\(^{-1}\) becomes increasingly sharp. Since these peaks have also been previously observed in the far-IR spectrum of poly-(\( l \)-alanine) and assigned to the \( \beta \)-sheet conformation, we can also confidently assign them to the \( \beta \)-sheet in A. pernyi SF. The detailed band assignment of A. pernyi SF in the 700–50 cm\(^{-1}\) spectral range is summarized in Table 2.

### 3.2 Determination of specific components in SF-based protein blends by far-IR spectroscopy

Our results clearly show that with respect to SPI and PPI, B. mori and A. pernyi SFs possess characteristic absorption peaks in far-IR region, although they display very similar absorbance in the mid-IR spectral region. Furthermore, within SFs, B. mori and A. pernyi can be clearly distinguished by their far-IR absorption spectra. Thus, we applied far-IR spectroscopy to characterize different SF-based protein blends. First, we investigated the protein blends from one globular protein (SPI) and one fibrous protein (A. pernyi SF). As shown in Fig. 1, SPI has broad absorption only in the far-IR region, so the characteristic peaks of A. pernyi SF are highly distinguishable in the far-IR spectra of the blend films (Fig. 4A and B). On increasing the A. pernyi SF content in the blend film, the normalized absorbance (normalized to the film thickness) of these characteristic peaks (namely, at 373 and 526 cm\(^{-1}\)) increases accordingly (Fig. 4C and D). Therefore, we may use the absorbance values of the characteristic peaks to determine the A. pernyi SF content within the blend film. There is a linear relationship between the absorbance at both the 373 and 526 cm\(^{-1}\) bands and the A. pernyi SF content in the blend film (Fig. 5), suggesting that it is possible to directly use the absorbance value of a specific characteristic peak of A. pernyi SF to determine its content in the protein blends. After analyzing the relatively simple A. pernyi SF/SPI blend film, we further analyzed two fibrous protein blend films, namely, A. pernyi SF and B. mori SF blend films. As shown in Fig. 6A, both A. pernyi SF and B. mori SF have characteristic peaks in the far-IR region and they are all

### Table 2  The far-IR band assignment in A. pernyi SF

<table>
<thead>
<tr>
<th>Wavenumber (cm(^{-1}))</th>
<th>Assignment</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>658</td>
<td>CN tor, NH opb, C=O opb</td>
<td>35</td>
</tr>
<tr>
<td>620</td>
<td>C=O ipb, C( ^\circ ) str; ( \beta )-sheet</td>
<td>3</td>
</tr>
<tr>
<td>526</td>
<td>C( ^\beta ) bend 1, C=O ipb, C( ^\circ ) str; ( \alpha )-helix</td>
<td>35</td>
</tr>
<tr>
<td>448</td>
<td>C( ^\beta ) bend 2, NC( ^\beta ) def; ( \beta )-sheet</td>
<td>3</td>
</tr>
<tr>
<td>373</td>
<td>C( ^\beta ) bend 2, NC( ^\beta ) def, C( ^\circ )CN def; ( \alpha )-helix</td>
<td>35</td>
</tr>
<tr>
<td>330</td>
<td>NC( ^\circ )C def</td>
<td>3</td>
</tr>
<tr>
<td>245</td>
<td>C( ^\circ )C( ^\circ ) tor, NC( ^\circ )C def; ( \beta )-sheet</td>
<td>3</td>
</tr>
<tr>
<td>118</td>
<td>NC( ^\circ )C def, NCC( ^\circ )C def, NCC( ^\circ ) def</td>
<td>3</td>
</tr>
</tbody>
</table>

\( ^\circ \) Abbreviations: tor = torsion, opb = out-of-plane angle bend, ipb = in-plane angle bend, str = stretch, def = deformation.
visible in the blend film spectra (Fig. 6B). Similar to the A. pernyi SF/SPI blend films, the absorbance of the characteristic peaks from A. pernyi SF at 373, 526, and 620 cm\(^{-1}\) increased gradually with A. pernyi SF content, while the absorbance at 553 cm\(^{-1}\) from B. mori SF decreased drastically (Fig. 6C and D) with the increase in A. pernyi SF content.

Fig. 7 shows the relationship between the absorbance of characteristic peaks from A. pernyi SF at 373, 526, and 620 cm\(^{-1}\) and the mass fraction of A. pernyi SF in the A. pernyi SF/B. mori SF blend films. Generally, they all result in a linear relationship, which is the same as that in the A. pernyi SF/SPI blend films. Moreover, the absorbance of the characteristic peaks in B. mori SF at 553 cm\(^{-1}\) also shows a linear relationship with the A. pernyi SF mass fraction. This implies that the absorbance of the characteristic peak in both A. pernyi and B. mori SF can be used as an indicator to determine both their conformation as well as their content in the blend films. However, we noticed when A. pernyi SF content in the blend film was low (mass fraction <0.2), the absorbance of A. pernyi SF characteristic peak at 620 cm\(^{-1}\) cannot fit into a linear relationship with the A. pernyi SF mass ratio (Fig. 7D). The same phenomenon was also found in the absorbance of the B. mori SF with the characteristic peak at 553 cm\(^{-1}\) when the B. mori SF content in the film was low (mass fraction <0.2, Fig. 7C). By carefully examining the four characteristic peaks shown in Fig. 7, we found that the peaks at 553 and 620 cm\(^{-1}\) were not as sharp as those at 373 and 526 cm\(^{-1}\). Therefore, when the content of A. pernyi or B. mori SF was low, the measurement of these weak and broad peaks may have relatively large errors. That is to say, if it is possible, choosing a sharp characteristic peak rather than a broad one in the Far-IR spectrum as a probe is the best choice when determining the content of a specific component in protein blends.

4. Conclusions

Herein, we used far-IR spectroscopy to analyze the structure of four abundant non-bioactive proteins, namely, SPI, PPI, B. mori SF, and A. pernyi SF. The two non-bioactive globular proteins (SPI and PPI) resulted in very similar far-IR spectra characterized by multiple weak and broad absorption peaks. The far-IR spectra agree well with those reported for other bioactive globular proteins, including myoglobin, hemoglobin, and lysozyme. In contrast, fibrous non-bioactive SF proteins (B. mori and A. pernyi), which feature highly repetitive amino acid sequences, exhibit distinct and sharp absorption peaks in the far-IR region. Moreover, we also found that some of the characteristic peaks in SF are sensitive to protein secondary structure, which can be used to investigate the conformational transition process of SFs. Based on the significant difference between the far-IR spectra of SFs and globular proteins and even between different types of SFs, we are able to quantitatively evaluate the specific SF ratio within protein blends. Characteristic absorbance peaks of SFs resulted in a linear relationship with the SF content within the protein blend. Therefore, far-IR spectroscopy has great potential for studying SFs and SF-based protein blends not only to directly assign the conformation of the SFs, but also to determine the ratio of the specific SF component in the protein blend.

Conflicts of interest

There are no conflicts of interest to declare.

Acknowledgements

This study was supported by the National Natural Science Foundation of China (No. 21574023 and 21574023). We thank Prof. Jinrong Yao and Dr Yuhong Yang at Fudan University for
their valuable suggestions and discussions. Prof. Shengjie Ling acknowledges the starting grant of ShanghaiTech University.

Notes and references