Invited Paper

The Hydrogen Bonding Signature of Peptides and Proteins in the Far Infrared

Youssef El Khoury, Aurélien Trivella and Petra Hellwig *
Laboratoire de Spectroscopie Vibrationnelle et Electrochimie des Biomolécules Institut de Chimie, UMR 7177, Université de Strasbourg 1 rue Blaise Pascal, 67008 Strasbourg Cedex (France)
* E-mail: hellwig@unistra.fr

（Received November 22, 2010）

Abstract: Spectroscopic techniques that use the low frequency region are strongly emerging for the study of biological molecules. Far infrared and far Raman spectroscopies, THz time domain approaches and inelastic neutron scattering reveal the presence of vibrational modes involving inter- and intramolecular hydrogen bonding. Due to their collective nature, such modes are highly sensitive to the conformational state of the molecule. Here the influence of the secondary structure on these vibrational features in the far infrared for model compounds and proteins of well known structure are described. Since temperature has a large effect on hydrogen bonding, the development of the signature of poly-L-lysine between 14 and 294 *K* is presented. The data does not only reveal the increase of the number of hydrogen bonds with temperature, but also the reorganizations within the structures.

Keywords : Far infrared, Hydrogen bonding, Proteins, Polyamino acids, Temperature dependence

doi: 10.11906/TST.183-191.2010.12.18

1. Introduction

The biological function of a protein critically depends on its conformation and flexibility and on this basis on the inter- and intramolecular hydrogen bonding. Hydrogen bonds play an important role in both, charge shifting and the stabilization of molecules and they contribute to the affinity and selectivity of molecular recognition from unique systems like water to multi-component complexes of proteins, nucleic acids, and membranes. Importantly, hydrogen bonds play a crucial role in the three dimensional folding of proteins. Structural elements like α-helices and β-sheet folds are defined and stabilized by their unique pattern of hydrogen bonds.

Vibrational modes seen at low frequencies are highly sensitive to the collective intra- and intermolecular structure and provide a unique fingerprint of the molecule(s) arrangement. Several spectroscopic techniques give access to these vibrational modes including low frequency Raman, far infrared (far IR) or incoherent inelastic neutron scattering. In the far infrared, a continuum is found at about 400 to 0 cm^{-1}, the exact position and broadness depending on the polarizability of the hydrogen bonding features [1-3]. It includes the contribution of the inter– and intra–molecular hydrogen bonding H…O and H…N stretching vibrations within a protein or with water molecules. As a consequence far infrared spectroscopy and the THz spectroscopy on biological molecules experienced a flurry of activities. Research activities in this spectral domain started with the prediction of the collective structural vibrational modes first presented almost 40 years ago for DNA and for other large biological molecules [4-6]. Hydrogen bonds from water molecules were found to contribute to this same spectral region and within these lines the reorganization of the water solvation shell in the picosecond time range in the THz was recently

described [3, 7, 8]. Since bands sensitive to hydrogen bonds can be identified by perturbing or breaking the hydrogen bond, experiments varying pressure, temperature, pH and ionic strength have been presented [9-11]. The study of acetone: chloroform mixtures, a model system where formation of a hydrogen bond can be analyzed, lead to comparable spectroscopic data [12] and are important for the understanding of the data obtained for biological molecules.

The temperature dependent phase transition from phospholipids in the far IR domain and the specific behavior of the intrinsic spectral feature of the head groups within the phospholipid bilayer were recently described [13]. Similarly hydrogen bonding within proteins can be studies as reported for the complex I from the respiratory chain [14] and in a temperature dependent manner, for the Rieske protein [15], supporting the hypothesis, that structure mediated changes in the environment of the active centers play a critical role in regulating enzymatic catalysis.

Nevertheless the understanding of these spectral features for the study of proteins is still evolving. Due to the high polarizability of the signal, the influence of components like salt or detergent needs to be analyzed. The nature of the intra- and intermolecular interaction and importantly the influence of the secondary structure on this spectral feature are not clear. In this study we follow the well known structural changes of polyamino acids in the far infrared and we show on the one hand that the secondary structure and the hydrogen bonding feature is not correlated, however, that the signal is specific for each studied molecule.

2. Materials and methods

2.1 Sample preparation

N-Methylacetamide ≥ 99%, Poly (L-lysine) 45 kDa and Poly-L-glutamic acid 2-15 KDa were purchased from Sigma-Aldrich. The protonation and deprotonation effects were studied in the presence of 0.1 M HCl and 0.1 M NaOH, respectively. Spinach Ferredoxin was purchased from Sigma-Aldrich as a lyophilized powder (15% Ferredoxin). The protein was dissolved in 20 *mM* Tris-HCl buffer pH 7.4. Lyophilized Equine Myoglobin was purchased from Sigma-Aldrich and dissolved in pure water (pH 6). The Rieske protein from Thermus thermophilus was prepared as previously reported [16] and dissolved in 20 *mM* phosphate buffer pH 7, to a final concentration of ~2 *mM*.

2.2 Data Collection

The ATR-FTIR measurements in the mid IR and in the far IR range were performed on the attenuated total reflection unit (ATR Harrick crystal, Diamond Prism). 2 *μL* of the respective sample were used to obtain a film on the crystal.

For the temperature dependent measurements 20 *μL* of PLK were dried on a polyethylene window, placed in a copper sample holder, and mounted on the cold finger of a closed cycle He-cryostat (model DE-202 AE, Advanced Research Scientific, Allentown, PA, USA). The temperature was measured with a silicon diode (Scientific Instruments calibration, precision of ±0.5 *K* in the range 2–100 *K*). The temperature was regulated in the 294–14 *K* range with a

heating resistor monitored by a digital temperature controller (model 9700–1–1, Scientific Instruments, West Palm Beach, FL, USA).

Data in the mid IR and far IR domains were measured with a Vertex 70 instrument from Bruker, (Karlsruhe, Germany) the optics and the parameters where adapted for the different spectral ranges as follows: In the MIR domain a detector mercury cadmium telluride (MCT) was used and the scan velocity was 40 *kHz*. Ten spectra with a resolution of 4 cm^{-1} (256 scans) were averaged for each sample. In the far IR domain a deuterated triglycine sulfate (dTGS) detector was used and the scan velocity was 2.5 *kHz*. For one sample five spectra with a resolution of 4 cm^{-1} (128 scans) were averaged. The instrument was purged with dry air in order to avoid contributions from humidity in the spectra. No baseline correction or smoothing procedure was applied to the absorption spectra shown here.

3. Results and discussion

3.1 Describing a model for the hydrogen bonding in peptides - NMA

Fig. 1 shows the spectra of the neutral and the deprotonated forms of NMA in the mid IR and the far IR domain. In this simple model for a peptide bond and its intermolecular interactions, the deformation mode of the CN group can be found at 196 cm^{-1} and the intrinsic mode of the hydrogen bond at 100 cm^{-1}. Upon addition of NaOH (to a final pH of 11), this signal disappears and the CN^- vibrational mode appears at 1430 cm^{-1} due to the deprotonation of the NMA. The data is in line with previous studies with IR and Raman spectroscopies [17-19]. The C=O…H-N hydrogen bonding found in the protonated form corresponds to the structure found in proteins. It can thus be expected that it is also found for proteins and similar biological polymers or macromolecular structures. In a more general way, the presence of a large continuum-like spectral feature, involving hydrogen bonding modes was studied in detail by Zundel and colleagues on several models for hydrogen bonding [20-23].

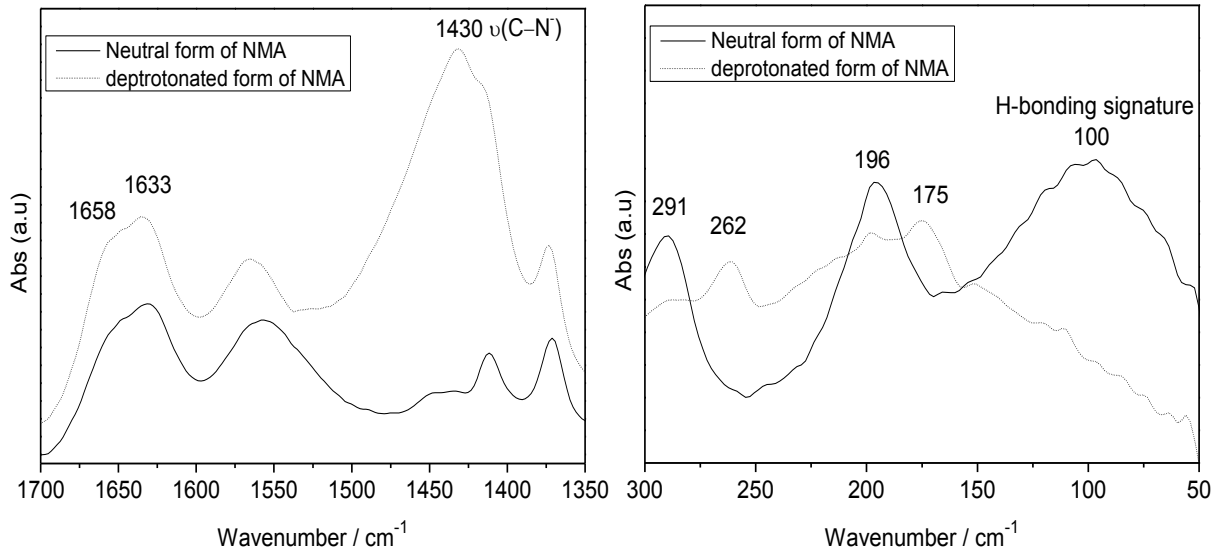

Fig. 1 ATR-FTIR spectra of protonated and deprotonated NMA in the 1700-1350 cm^{-1} (A) and in the 300-50 cm^{-1} range (B) at pH 6 and 11, respectively.

Since the hydrogen bonding strength should influence this spectral feature and the different secondary structural elements typically observed for proteins, like in the most common cases, alpha helix, random coil or beta sheet, show different hydrogen bonding strengths, polymers have been studied that may change structure in function of external factors. This is the case for polyaminoacids that change their organization in function of pH and temperature.

3.2 On the role of the secondary structure of the hydrogen bonding spectral feature

Fig. 2 shows the spectra of poly-L-glutamate (PLG) at pH 5.0 and pH 7.0 for the mid IR (A) and the far IR (B). This pH range involves the transition random coil conformation at neutral pH into alpha-helical conformation at acidic pH [24]. The transition, however, is not always complete and a mixture of both states may coexist. In the mid infrared the signal can be found at 1641 cm^{-1} for pH 5 (Fig. 2Aa) and at 1643 cm^{-1} for pH 7 (Fig. 2Ab). According to the extensive literature available on the amide I range, random coil contributions are expected around 1648 ± 4 cm^{-1} and alpha helical signals around 1653 ± 4 cm^{-1} [25, 26]. There is thus only a partial transistion. In the far IR a large signal can be seen, that evolves at 300 and is still evident below 50 cm^{-1}. In function of pH the maximum shifts from 138 cm^{-1} at pH 7 (Fig. 2Bb) to 164 cm^{-1} at pH 5 (Fig. 2 Ba). A shoulder at 125 cm^{-1} can be depicted, confirming the presence of residual random coil conformation. When comparing beta-sheet and random coil type-structures, a higher degree of organization and stronger hydrogen bonding is expected for the polyamino acids organized in beta-sheet structure. The observed frequency shifts would thus be in line with this suggestion.

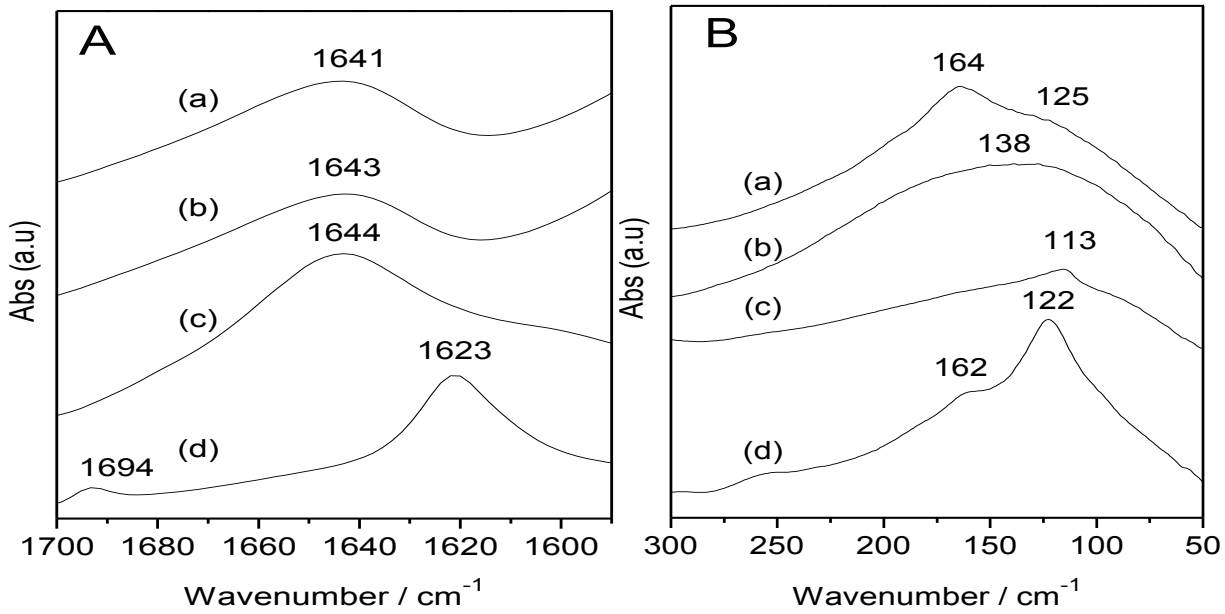

Fig. 2 ATR-FTIR spectra of PLG at pH 5 (a), pH 7 (b) and PLK at pH 7 (c) and pH 10.5 (d) in the amide I range (A) and in the 300-50 cm^{-1} range (B).

We thus probed further polyamino acids to see if there is any correlation between the secondary structure and the hydrogen bonding feature.

The data for poly-L-lysine is shown in Fig. 2. This well studied poly amino acid shows a mixture of beta-sheet and random coil with majority of random coil conformation at pH 7 [27-29]. At pH 10.5, the conformation predominantly involves beta-sheet contributions. It is noted that the exact structure and the phase transition depends on the polymer length, the temperature and other external factors [27]. The signal seen at 113 cm^{-1} in Fig. 2Bc for the data at pH 7, corresponds to the predominantly random coil type structure. Upon organization into beta sheet an up shift of the signal is seen at 122 and 162 cm^{-1} (Fig. 2Bd).

There is no specific signal in the far infrared that may be correlated with one of the specific secondary structures, but clearly the organization of the biopolymer into a higher organized structure like a helix or a beta sheet, lead to a shift to higher frequencies. We note that this statement is only valid for structural changes within one biopolymer. As soon the structures of different proteins are compared, the correlation seems more complex. This was previously described by inelastic neutron scattering experiments [30].

Fig. 3 shows the mid IR and far IR data for three different proteins: myoglobin, Rieske and Ferredoxin. Myoglobin has an overall alpha helical conformation [31] while the conformation of the Rieske protein is mainly beta sheet [32]. The crystal structure of the Ferredoxin from spinach shows that the protein has 47% random coils, besides 34% beta sheet and 19% alpha helices [33]. Whereas the far IR signal is localized at 145 and 142 cm^{-1} for Myoglobin and the Rieske protein respectively, the signal for Ferredoxin is found at 108 cm^{-1}.

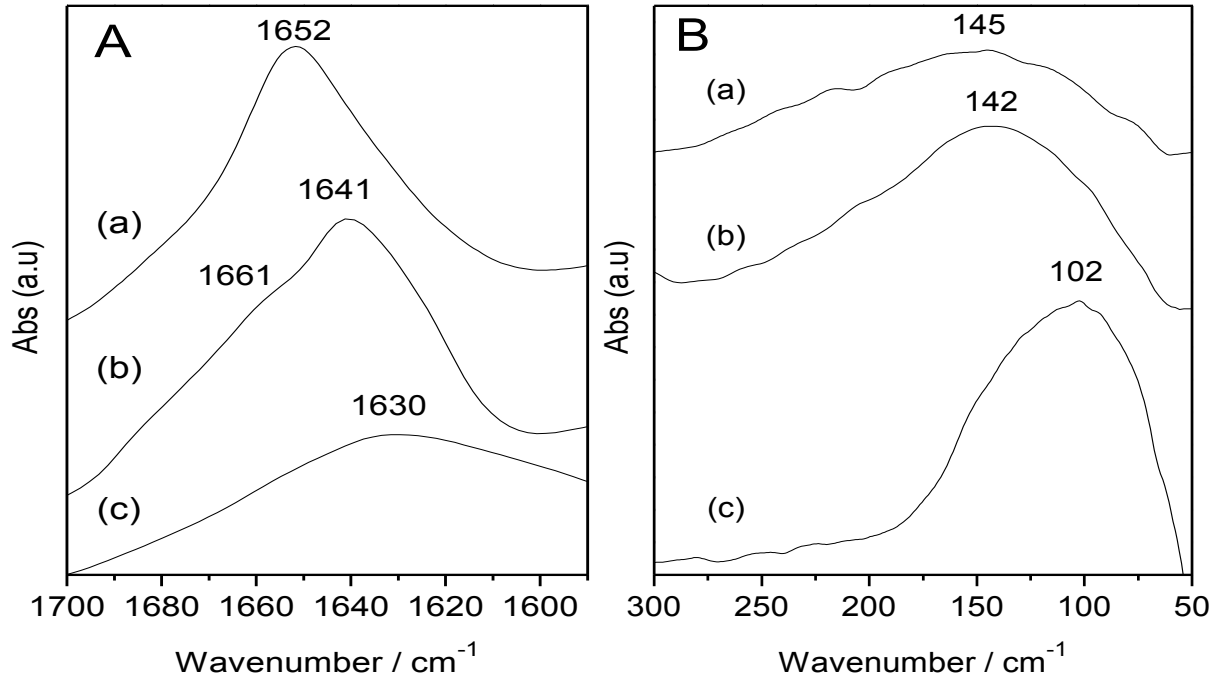

Fig. 3 ATR-FTIR spectra of Myoglobin, the Rieske protein and spinach Ferredoxin in the amide I range (A) and in the 300-50 cm^{-1} range (B) at neutral pH.

No clear correlation between different proteins and their secondary structure is possible. Still, the hydrogen bonding continuum seen below 300 cm^{-1} is specific for each of the studied protein and importantly, it presents a fingerprint for the protein in a specific conformation. Recently the change of the overall conformation of the respiratory enzyme, called NADH:ubiquinone

reductase in the presence and absence of its substrate was demonstrated in the far IR [14].

3.3 Temperature dependence of the hydrogen bonding collective mode

In order to analyze the information within the hydrogen bonding signature in the far infrared the perturbation within the studied system is necessary. While deuteration is widely used in the mid IR to perform assignments, in the far IR, deuteration is less efficient, since the shift of the hydrogen bond is about 1-5 cm^{-1} [3, 34]. Studies of the temperature dependence of IR spectra however, make it possible to compare the spectra in various phase states and to clarify the role of intermolecular forces. Temperature dependent studies aim to answer questions on the crucial role of hydrogen bonds, since temperature has a large effect on H-bonding, which is both, directional and cooperative. For many molecules, hydrogen bonds are what often determine structure, and when extreme temperature disrupts the structure, the IR spectra change [35].

Fig. 4 shows the temperature dependent development of the signal from poly-L-lysine in pure water. The secondary structure, as monitored in the amide I region, corresponds to the pH 7 data. We note that the relative intensities and the position of the signals in the far infrared are changed when working with transmission on polyethylene. As previously reported, the exact position of individual peaks are found shifted for 1–2 cm^{-1} in some cases and for up to 10 cm^{-1} in function of the window material used for the measurements [15]. In a possible explanation, the different charge of the each surface type may induce a different orientation and the arrangement of the protein. Importantly, the data obtained with each material is highly reproducible, and the comparison of data at different pH or temperatures values is then valid. It should also be noted that the optics of the ATR unit induces some polarization effects.

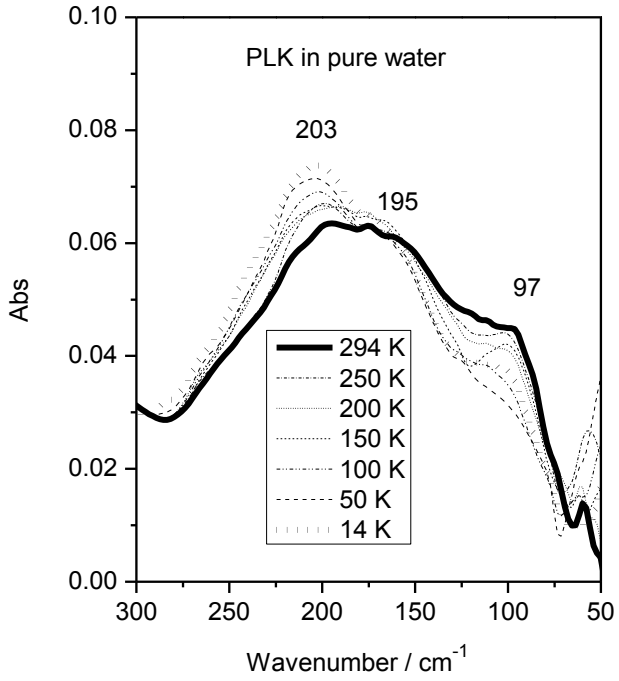

Fig. 4 Temperature dependence of PLK measured from 14 to 294 *K* in pure water on a polyethylene window. For details see materials and methods.

Importantly, there is a wavelength dependent penetration depth in ATR measurements and the data at lower frequencies corresponds to the contribution of a significantly larger amount of sample. For the diamond crystal the penetration depth in the mid infrared is typically around 1-2 μm, at 500 cm^{-1} it is around 4 and at 150 cm^{-1} it is close to 13 μm. The exact value depends on the diffraction index of the sample.

In the temperature dependent experiment shown in Fig. 4, an increae of the signal intensity and a shift from 195 to 203 cm^{-1} can be seen when cooling the sample from 294 to 14 *K*. The increase of the signal intensity can be correlated with the increase of the number of hydrogen bondings upon cooling, as for example also described for water [3]. Proteins have been reported to show a liquid-glass transition around 150-180 *K*, the signals increase essentially being linear [30, 36-39]. Extensive studies with Raman spectroscopies and inelastic neutron scattering have been performed correlating the structural flexibility above the transition temperature to the active form of the protein. Furthermore, the hydration shell of the biomolecule was found to be crucial [40, 41]. The internal perturbation of the hydrogen bonding structure by protonation reactions of specific residues may also change the temperature dependence of the signature in the far infrared as demonstrated for the Rieske protein from Thermus thermophilus, an iron sulfur protein whose activity crucially depends on the pH [15].

In the temperature dependence seen in Fig. 4, a signal at 97 cm^{-1} interestingly decreases upon cooling. A straightforward explation seems difficult. A similar phenomenon was however recently described for the temperature dependent phase transition of phospholipid bilayers as monitored in the far infrared. An alternation of the interaction between the phospholipids and water molecules was found to take place in function of the composition of the phospholipids and thus the headgroups involved in the bilayer [13]. The signal decrease seen here may reflect the reorganisation of internal water molecules.

4. Conclusions

In conclusion the far infrared spectral range involves the intermolecular hydrogen bonding interactions in large molecular assemblies in a very specific way. The use of this technique is still evolving and together with other low frequency techniques it is thus a powerful tool for investigations on structure, function and dynamics of biomolecules.

Abbreviations: ATR: attenuated total reflection, NMA: N-methylacetalide, PLK: poly-L-lysine, PLG: poly-L-glutamate

References

[1] G. Zundel, B. Brzezinski, J. Olejnik, *J. Mol. Struct.* 300, 573-592, (1993).

[2] F. Garczarek, K. Gerwert, *Nature* 439, 109-112, (2006).

[3] J. B. Brubach, A. Mermet, A. Filabozzi, A. Gerschel, P. Roy, *J. Chem. Phys.* 122, 184509, (2005).

[4] J. M. Eyster, E. W. Prohofsky, *Phys. Rev. Lett.* 38, 371-373, (1977).

[5] J. M. Eyster, E. W. Prohofsky, *Biopolymers* 13, 2527-2543, (1974).

[6] L. Young, V. V. Prabhu, E. W. Prohofsky, *Phys. Rev. A* 39, 3173-3180, (1989).

[7] U. Heugen, G. Schwaab, E. Brundermann, M. Heyden, X. Yu, D. M. Leitner, M. Havenith, *Proc. Natl. Acad. Sci. U. S. A.* 103, 12301-12306, (2006).

[8] S. Ebbinghaus, S. J. Kim, M. Heyden, X. Yu, U. Heugen, M. Gruebele, D. M. Leitner, M. Havenith, *Proc. Natl. Acad. Sci. U. S. A.* 104, 20749-20752, (2007).

[9] V. W. Srichatrapimuk, S. L. Cooper, *J. Macromol. Sci-Phys.* B15, 267-311, (1978).

[10] J. M. Vanderkooi, J. L. Dashnau, B. Zelent, *Biochim. Biophys. Acta* 1749, 214-233, (2005).

[11] D. D. Klug, M. Z. Zgierski, J. S. Tse, Z. Liu, J. R. Kincaid, K. Czarnecki, R. J. Hemley, *Proc. Natl. Acad. Sci. U. S. A.* 99, 12526-12530, (2002).

[12] T. S. Perova, D. H. Christensen, O. F. Nielsen, *Vib. Spectrosc.* 15, 61-67, (1997).

[13] R. Hielscher, P. Hellwig, *Chemphyschem* 11, 435-441, (2010).

[14] R. Hielscher, T. Friedrich, P. Hellwig, *Chemphyschem*, (2010).

[15] Y. El Khoury, A. Trivella, J. Gross, P. Hellwig, *Chemphyschem* 11, 3313-3319, (2010).

[16] T. Schroter, O. M. Hatzfeld, S. Gemeinhardt, M. Korn, T. Friedrich, B. Ludwig, T. A. Link, *Eur. J. Biochem.* 255, 100-106, (1998).

[17] K. Itoh, T. Shimanouchi, *Biopolymers* 5, 921-930, (1967).

[18] L. I. Maklakov, S. V. Aksakova, *Uspekhi Khimii* 66, 419-433, (1997).

[19] F. Fillaux, J. Tomkinson, *Chem. Phys.* 26, 295-300, (1977).

[20] J. Muehlinghaus, G. Zundel, *Biopolymers* 10, 711-719, (1971).

[21] G. Zundel, *J. Mol. Struct.* 381, 23-37, (1996).

[22] T. Keil, B. Brzezinski, G. Zundel, *J. Phys. Chem.* 96, 4421-4426, (1992).

[23] B. Brzezinski, B. Brycki, G. Zundel, T. Keil, *J. Phys. Chem.* 95, 8598-8600, (1991).

[24] S. Y. Venyaminov, N. N. Kalnin, *Biopolymers* 30, 1259-1271, (1990).

[25] J. Kong, S. Yu, Acta Biochim. *Biophys. Sin.* 39, 549-559, (2007).

[26] E. Goormaghtigh, V. Cabiaux, J. M. Ruysschaert, *Subcell. Biochem.* 23, 405-450, (1994).

[27] W. Dzwolak, T. Muraki, M. Kato, Y. Taniguchi, *Biopolymers* 73, 463-469, (2004).

[28] W. Dzwolak, V. Smirnovas, *Biophys. Chem.* 115, 49-54, (2005).

[29] N. N. Kalnin, I. A. Baikalov, S. Y. Venyaminov, *Biopolymers* 30, 1273-1280, (1990).

[30] S. Cusack, W. Doster, *Biophys. J.* 58, 243-251, (1990).

[31] J. C. Kendrew, G. Bodo, H. M. Dintzis, R. G. Parrish, H. Wyckoff, D. C. Phillips, *Nature* 181, 662-666, (1958).

[32] L. M. Hunsicker-Wang, A. Heine, Y. Chen, E. P. Luna, T. Todaro, Y. M. Zhang, P. A. Williams, D. E. McRee, J. Hirst, C. D. Stout, J. A. Fee, *Biochemistry* 42, 7303-7317, (2003).

[33] C. Binda, A. Coda, A. Aliverti, G. Zanetti, A. Mattevi, *Acta Crystallogr. D.* 54, 1353-1358, (1998).

[34] J. R. Powell, F. M. Wasacz, R. J. Jakobsen, *Appl. Spectrosc.* 40, 339-344, (1986).

[35] F. Demmel, W. Doster, W. Petry, A. Schulte, *Eur. Biophys. J.* 26, 327-335, (1997).

[36] V. Kurkal-Siebert, J. C. Smith, *J. Am. Chem. Soc.* 128, 2356-2364, (2006).

[37] S. Ciliberti, P. De Los Rios, F. Piazza, *Phys. Rev. Lett.* 96, 198103, (2006).

[38] A. L. Tournier, J. Xu, J. C. Smith, *Biophys. J.* 85, 1871-1875, (2003).

[39] W. Doster, S. Busch, A. M. Gaspar, M. S. Appavou, J. Wuttke, H. Scheer, *Phys. Rev. Lett.* 104, 098101, (2010).

[40] W. Doster, *Biochim. Biophys. Acta* 1804, 3-14, (2010).

[41] K. L. Ngai, S. Capaccioli, N. Shinyashiki, *J. Phys. Chem. B* 112, 3826-3832, (2008).