Title

Variations in the Protein Hydration and Hydrogen-Bond Network of Water Molecules Induced by the Changes in the Secondary Structures of Proteins Studied through Near-Infrared Spectroscopy

Author(s)

Mika Ishigaki, Yoshiki Kato, Eri Chatani, and Yukihiro Ozaki

Journal

The Journal of Physical Chemistry B, 2023 127 (32), pp.7111-7122

Published

July 21, 2023

URL

https://doi.org/10.1021/acs.jpcb.3c01803

Variations in the protein hydration and hydrogen-bond network of water molecules induced by the changes in the secondary structures of proteins studied through near-infrared spectroscopy

Mika Ishigaki[1]*, Yoshiki Kato[1], Eri Chatani[2], Yukihiro Ozaki[3]

[1]*Institute of Agricultural and Life Sciences, Academic Assembly, Shimane University, 1060 Nishikawatsu, Matsue, Shimane, 690-8504, Japan*

[2] *Department of Chemistry, Graduate School of Science, Kobe University, Nada, Kobe, 657-8501, Japan*

[3]*School of Biological and Environmental Sciences, Kwansei Gakuin University, 1 Gakuen-Uegahara, Sanda, Hyogo 669-1330, Japan*

*Authors to whom correspondence should be sent.

*E-mail: ishigaki@life.shimane-u.ac.jp (M.I.)

Abstract

This study investigated how the secondary structural changes of proteins in aqueous solutions affect their hydration and the hydrogen-bond network of water molecules using near-infrared (NIR) spectroscopy. The aqueous solutions of three types of proteins, i.e., ovalbumin, β-lactoglobulin, and bovine serum albumin, were denatured by heating, and changes in the NIR bands of water reflecting the states of hydrogen bonds induced via protein secondary structural changes were investigated. On heating, the intermolecular hydrogen bonds between water molecules as well as between water and protein molecules were broken, and protein molecules were no longer strongly bound by the surrounding water molecules. Consequently, the denaturation was observed to proceed depending on the thermodynamic properties of the proteins. When the aqueous solutions of proteins were cooled after denaturation, the hydrogen-bond network was reformed. However, the state of the protein hydration was changed owing to the secondary structural changes of proteins, and the variation patterns were different depending on the protein species. These changes in the protein hydration may be derived from the differences in the surface charges of proteins. The elucidation of the mechanism of protein hydration and the formation of the hydrogen-bond network of water molecules will afford a comprehensive understanding of the protein functioning and dysfunctioning derived from the structural changes in proteins.

Introduction

Water is a crucial component in living organisms, and biochemical reactions occur within an aqueous environment. The water molecules form a complex network of hydrogen bonds, which render anomalous properties such as high boiling point in spite of having low molecular weight and the density of water get maximum at 4 °C [1, 2]. Although the water molecule is one of the simplest and the most common molecules, the reasons behind the anomalous properties have not been fully elucidated yet.

When biomolecules such as proteins and DNA are dissolved in water, the hydrogen-bond network of water molecules is locally disrupted due to hydration at the interface of biomolecules named “hydration shell”[3, 4], and it begins to show different properties of the hydrogen-bond network from those of bulk water, where nothing is dissolved in water [5-7]. Protein molecules can exist stably as they are supported by the interactions with water, and the water molecules at the interfaces of protein molecules control the molecular structures of proteins and their functions [3, 4, 8, 9]. The structural changes of proteins not only vary their original functions but also their hydration properties, leading to changes in the hydrogen-bond network of water. It is very important to comprehensively understand the interactions between protein and water molecules to clarify the protein functions at a molecular level in living organisms.

Analytical methods such as X-ray, neutron scattering, nuclear magnetic resonance (NMR) spectroscopy, vibrational spectroscopy, and molecular dynamics (MD) simulation are often used for investigating biomolecular

hydration [5, 10]. For example, NMR experiments reported a slowdown in the reorientation and translational motion of water in protein aqueous solutions than in bulk water [11, 12]. It is supposed to be derived from the disorder of hydrogen bonds between water and protein sites in hydration shells [12, 13]. The pattern of X-ray scattering for protein aqueous solutions proved the existence of hydration water [14-16]. As the mentioned techniques are also good at analyzing structural changes in proteins, in principle, they can be used to reveal the effects of protein structural changes on variations in hydration. However, earlier studies could explain only the static state of hydrogen-bond networks around protein molecules, while the detailed dynamics of protein hydration are still unknown. It is because proteins have too complicated structures of their own and interactions with other molecules to elucidate the whole system for protein solutions. MD simulations shows that hydrogen bonds are formed at a protein surface, and hydration water around protein molecules showed different dynamical properties from those of bulk water [17-20]. Such variation in the state of hydrogen bonds is described as "the presence of biomolecules induces a glass-like property in the surrounding solvent" [20].

Raman spectroscopy is one of the vibrational spectroscopies, and it has been used for investigating the hydrogen-bond network of water [21-23]. The main Raman bands of water are due to O–H stretching (~3450 cm^{-1}) and O–H bending (~1640 cm^{-1}) vibrational modes. Raman spectroscopy can also detect bands of proteins such as N–H stretching vibration (~3300 cm^{-1}), amide I (~1650 cm^{-1}), and amide III (~1300 cm^{-1}) bands. Especially, the

amide bands clearly reflect the secondary structure of proteins. Moreover, the N–H stretching band is also sensitive to the strength of the hydrogen bonding of the amide groups [24-27]. However, the bands corresponding to water and proteins overlap each other in the 3500–3200 and 1700–1600 cm^{-1} regions, and it is necessary to analyze Raman spectra by resolving a Raman band into several component bands of water and proteins affected by the changes in hydration, hydrogen-bond network, and protein structures. Thus, the results obtained by band resolution depend on the spectral analytical methods, and it is impossible to rule out the uncertainties of the results. Near-infrared (NIR) spectroscopy is also one of the vibrational spectroscopies based on the overtone and combination modes of molecular vibrations that uses the 12500–4000 cm^{-1} (800–2500 nm) region of the electromagnetic spectrum [28, 29]. These modes are derived from the anharmonicity of the potential energy of molecular vibrations. Because the anharmonicity becomes large for the molecules, including a hydrogen atom, NIR bands due to the overtones of O–H, C–H, and N–H stretching modes are observed. Furthermore, band shifts caused by hydrogen bonds and intermolecular interactions are much larger for overtone and combination modes than for fundamental modes. The fact that the absorbance of NIR light by water is much smaller than that of mid-IR light enables us to perform the spectroscopic characterization of the samples with thicknesses in the order of millimeters. Thus, NIR spectroscopy is suitable for analyzing the protein aqueous solutions because it can provide information about changes in the protein secondary structure, protein hydration, and hydrogen-bond network

of water simultaneously.

Two broad water bands in the vicinity of 6900 and 5200 cm^{-1} are observed in the NIR wavenumber region that can be assigned to the combination of the antisymmetric and symmetric O–H stretching modes ($\nu_1 + \nu_3$) and that of the antisymmetric O–H stretching and O–H bending modes ($\nu_2 + \nu_3$) of water, respectively [30, 31]. Here, the bands of ν_1 (~3657 cm^{-1}), ν_2 (~1595 cm^{-1}), and ν_3 (~3756 cm^{-1}) are symmetric O–H stretching, O–H bending, and asymmetric O–H stretching vibrations, respectively [32]. The combination band ($\nu_2 + \nu_3$) of water rarely overlaps with the other overtone and combination bands of proteins, while the combination band ($\nu_1 + \nu_3$) heavily overlaps with that of proteins. Thus, the NIR bands reflecting the secondary structural changes of proteins in the 4800–4500 cm^{-1} region can be independently and simultaneously investigated with those of water in the 5500–5000 cm^{-1} region affected by the changes of the protein hydration and hydrogen-bond network of water.

In this study, the variations in the protein hydration and hydrogen-bond network of water caused by the protein secondary structural changes were investigated through NIR spectroscopy. Heating was used to denature three types of proteins from α-helix to β-sheet, and variations of NIR water bands were analyzed in relation to the secondary structural changes. Three types of proteins, i.e., ovalbumin, β-lactoglobulin, and bovine serum albumin, were chose as representative native proteins with different composition ratios of the secondary structures of α-helix and β-sheet. Here, this study revealed not only the changes in the hydrogen-bond network due to hydration on the

protein surface but also the impact on water as a whole. The increase in protein concentration made the hydrogen-bond network stronger due to the hydration between protein and water molecules. The variation patterns of the hydrogen-bond network depended on the type of protein and slightly on the solvent species. Such differences may be caused by the differences in hydrophobic and hydrophilic properties on the surface of protein molecules depending on the protein species and their interaction with solvents.

Hydration water of proteins have a crucial influence of protein structure and function including protein folding, enzyme function, and molecular recognition[3, 4]. A deep understanding of the interactions between protein and water molecules would help elucidate the protein function and dysfunction within living organisms.

Materials and Methods

Sample preparation

Protein samples of ovalbumin (OVA; 012-09885, from a chicken egg, FUJUFILM Pure Chemical Co., Japan), β-lactoglobulin (BLG; L3908, from bovine milk, Sigma-Aldrich Co., USA), and bovine serum albumin (BSA; 013-15121, FUJUFILM Pure Chemical Co., Japan) were diluted with ultrapure water (214-01301, FUJIFILM Pure Chemical Co., Japan) or 0.05-M Tris-HCl buffer (pH = 7.6, FUJUFILM Pure Chemical Co., Japan) to 10 wt%.

NIR measurements

The NIR spectra were recorded using an IRTracer-100 NIR system (Shimadzu, Kyoto, Japan). An open quartz cell with an optical path length of 0.1 mm (AB20-IR-0.1, GL Sciences Inc., Japan) was used, and a NIR measurement was performed under a flow of dry air (Precision Nitrogen Trace 1000, PEAK Scientific, UK). The wavenumber region for recording the spectra was 10000–4000 cm^{-1} at a spectral resolution of 4 cm^{-1}, and the accumulation was 128 times. The cell was set into a mid-IR base (PN 111-40XX, PIKE, USA) using a digital temperature controller (PN 076-1510, Pike Technologies, USA). The NIR spectra were obtained every 5 min for 60 min while heating at 60 °C, 70 °C, 80 °C, or 90 °C after the cell was placed into the cell holder for 5 min. The heated cells were removed from the cell holder and cooled using ice for 5 min. The cooled cell was restored in the temperature controller reset at 25 °C, and the NIR spectra were recorded again after standing for 15 min. Each analysis was repeated three times for confirming reproducibility. The stability of the temperature controller was ±0.1 °C.

Spectral data analysis

For the spectral analysis, the chemometrics software, Unscrambler X 10.3 (Camo Analytics, Oslo, Norway) and OriginPro 6.1 (OriginLab Corporation, Massachusetts, USA), were used. The second derivative spectra were analyzed using the Savitzky–Golay (SG) method with a second-order polynomial set of the window size as 15 (total 31) after smoothing with the same SG conditions. Capturing the changes in the curvature of the NIR

spectra of protein aqueous solutions by calculating the second derivatives facilitates their resolution into some bands of the individual components overlapping with each other.

Confirmation of the changes in the secondary structure of the protein by Raman spectroscopy

The glass capillaries (outer diameter: 1.52 mm, inner diameter: 1.13 mm, Drummond Scientific Company, USA) were filled with three types of protein aqueous solutions, and both ends of the capillary were sealed by heating using a gas burner. The sealed glass capillaries were warmed in a water bath (DTU-1CN, TAITECH, Japan) at 60 °C, 70 °C, 80 °C, and 90 °C for 20 min. After heating, the samples were taken out from the water bath and cooled in ice water.

Raman spectra were recorded for the preprocessed protein samples at room temperature in the 1900–400 cm^{-1} region using a Raman microscope (XploRA INV, HORIBA Ltd., Kyoto, Japan), which consisted of a 532-nm diode laser, spectrometer, 1800-gr/mm grating, CCD camera, and microscope (Eclipse Ti2 Inverted Microscope, Nikon Co., Tokyo, Japan) with a 40× objective lens (Plan Apochromat, NA=1.20, Nikon). The exposure time was set as 15 s, and 40 scan accumulations led to a total exposure time of 600 s. Raman spectra were calibrated using the peak of silicon as standard, and they were preprocessed using background subtraction and fifth-order polynomial fitting for fluorescence background removal. All Raman spectra obtained from the samples were normalized

using a standard peak at 1003 cm^{-1} due to phenylalanine. Three samples were prepared for each heating time. Raman spectra were recorded at three points for each sample. Mean Raman spectra were calculated based on the nine preprocessed Raman spectra. The experimental results for Raman spectroscopy are given in Supporting Information (SI) 1.

Confirmation of changes in the protein secondary structure by the fluorescent pigment of thioflavin T

To confirm the changes in the secondary structure of the protein, the formation of aggregates was investigated using the fluorescent pigment of thioflavin T (ThT) [33]. The ThT powder was dissolved in distilled water and mixed with a glycine–sodium hydroxide buffer solution (pH = 8.5). The final concentrations of ThT and the buffer solution were 5 μM and 50 mM, respectively. The protein aqueous solutions heated at four temperatures (60 °C, 70 °C, 80 °C, and 90 °C) for 60 min were taken using capillaries and transferred to microtubes with ThT solutions. After stirring the mixture, the fluorescence of each of the three samples was measured at 60 °C, 70 °C, 80 °C, and 90 °C using a microplate reader (MTP-450, Corona, Japan). The excitation and detection wavelengths were 445 and 485 nm, respectively. All these experiments were repeated three times for confirming the reproducibility of the results. The detailed results are shown in SI 2.

Results and Discussion

Changes in the secondary structures of proteins

Figure 1A depicts an NIR absorbance spectrum in the 10000–4000 cm^{-1} region for a 10-wt% OVA aqueous solution at 25 °C. Two broad bands at around 6900 and 5200 cm^{-1} were assigned to the combination of the antisymmetric and symmetric O–H stretching modes and that of the antisymmetric O–H stretching and O–H bending modes of water, respectively [30, 31]. A second derivative spectrum was calculated to specify the NIR bands attributed to proteins. Figure 1B shows the second derivative spectrum in the 5300–4500 cm^{-1} region. The weak absorbance bands of proteins were resolved as downward peaks at 4863, 4610, and 4528 cm^{-1} in addition to the strong peak at 5250 cm^{-1} due to water, i.e., the structural changes in the protein and hydrogen-bond network of water can be simultaneously and independently analyzed by focusing on the second derivative spectrum in the 5500–4500 cm^{-1} region for protein aqueous solutions.

Figure 1C depicts the NIR second derivative spectra of 10-wt% aqueous solutions of the three types of proteins in the 4900–4200 cm^{-1} region. Some characteristic bands due to proteins were observed. The bands at ~4860 and ~4610 cm^{-1} are due to the combination of N–H stretching and amide II modes and that of N–H stretching and amide III modes, respectively [28, 30, 34, 35]. The bands near 4530 and 4410 cm^{-1} were ascribed to a peptide, especially with β-sheet structure, and the band at ~4360 cm^{-1} was assigned to a peptide with an α-helix structure [30, 35]. Furthermore, the C–H components in the proteins were also detected through the bands at ~4410 and ~4260 cm^{-1},

which were due to the combination of CH stretching and CH_2 deformation modes and the first overtone of the symmetric stretching of CH and CH_2 deformation modes, respectively [30, 36, 37]. The four bands at approximately 4860, 4610, 4530, and 4360 cm^{-1} showed similar spectral patterns in the band positions and intensities regardless of solvents. On the contrary, the intensity of the band at ~4410 cm^{-1} got stronger in the second derivative spectra in Tris-HCl buffer solution than that in ultrapure water because the bands of proteins and Tris-HCl buffer overlapped each other.

Izutsu et al. reported NIR second derivative spectra of seven types of protein aqueous solutions and their freeze-dried solids [37]. They concluded that NIR bands at 4860 and 4530 cm^{-1} strongly suggest the presence of the proteins with β-sheet structures while the intense ones at 4615 and 4365 cm^{-1} suggest the proteins with α-helix rich structures. The α-helix structure is predominant in BSA molecules with a minority of β-sheet structure. The band intensities at 4613 and 4370 cm^{-1} in the second derivatives in BSA were certainly stronger than those of the other proteins. Moreover, BLG contains a β-sheet structure higher than the ratio of 40%, and the 4861 and 4532 cm^{-1} bands were more clearly observed than those of OVA and BSA. OVA is composed of α-helix and β-sheet structures in almost equal proportions, and its second derivative spectra showed the intermediate properties of the spectra of the two proteins mentioned above. The spectra of the three types of protein solutions with the two types of mediums confirmed the presence of the spectral patterns reflecting the ratio of the protein secondary structures.

Figures 2A and 2B show the second derivative spectra of OVA in

ultrapure water in the 4880–4850 and 4650–4500 cm^{-1} regions collected every 5 min while heating at 70 ℃ for 60 min. As the heating time progressed, the bands at ~4860 and ~4610 cm^{-1} shifted to lower and higher wavenumbers, respectively. Additionally, it can be seen that the intensity of the peak at 4530 cm^{-1} in the second derivative spectrum got stronger with the progression of the heating time. Because the denaturation temperature of OVA is about 75 °C [38], these spectral variations indicate the changes in the protein secondary structure due to heating. Liu et al. indicated the correlation between the shift of the NIR band at 4800 cm^{-1} due to the combination of N–H stretching and amide II modes and the shift of the N–H stretching vibration in mid-IR wavenumber region (~3290 cm^{-1}) [26]. The present study reconfirmed that the lower frequency shift of the NIR band derived from the combination of N–H stretching and amide II modes reflected the changes in the hydrogen bonds of N–H groups based on the changes in the secondary structure of proteins. The shift of the band at ~4610 cm^{-1} to a higher frequency after the secondary structure of the protein changes into β-sheet can be explained by the higher frequency shift of the amide III band [24, 39]. Furthermore, the band at ~4530 cm^{-1} is assigned to the N–H and C=O groups in the peptide backbone referred to as the β-sheet structure [30, 35]. Thus, the increase in the peak intensity of the second derivative spectrum becomes an indicator to monitor the changes in the protein's secondary structure.

The changes in the positions of bands at ~4860 and ~4610 cm^{-1} and that of in the intensity of the second derivative band near 4530 cm^{-1} of OVA ultrapure water solutions at four heating temperatures are summarized in

Figures 2C, 2D, and 2E, respectively. Figures 2C and 2D revealed that the secondary structures of proteins did not change at 60 °C and that they changed over time at temperatures higher than 70 °C. It was found that the secondary structure already changed within 5 min after the beginning of heating at 80 °C and 90 °C. The intensity of the second derivative bands tended to increase with time as shown in Figure 2E. However, the band shifts are more sensitive and useful than the changes in their intensities to follow the changes in the secondary structure as the sample-to-sample variations in the intensities are large. Thus, in this study, the secondary structure changes of proteins were mainly tracked by the band shifts of ~4860 and ~4610 cm^{-1}.

The corresponding plots of the band positions for the other two proteins are shown in SI 3. It was confirmed that the secondary structural changes in BSA occurred at 60 °C based on the shifts of these two bands (~4860 and ~4610 cm^{-1}), and in the case of BLG, they occurred at temperatures higher than 70 °C, similar to those in OVA. These results corresponded to the thermal stability of the proteins; thermal transition midpoint of OVA and BLG are at 70 °C and that of BSA is at ~55 °C [40-42]. Thus, it was established that changes in the secondary structures of all the three proteins upon heating could be traced by the shifts of the two NIR bands at 4860 and 4610 cm^{-1}.

Hydrogen-bond network in ultrapure water

The states of hydrogen bonds of ultrapure water were analyzed by temperature-dependent NIR spectral variations. Figure 3A shows the mean NIR absorbance spectra in the 8000–4480 cm^{-1} region obtained over a

temperature range of 5–70 °C at a 5 °C interval; the spectra were normalized by the area in the wavenumber region. It was confirmed that increasing the temperature shifted the bands near 6900 and 5200 cm^{-1} to higher wavenumbers.

Principal component analysis (PCA) was conducted for the NIR data in the 5870–4480 cm^{-1} region, where the band due to the combination of antisymmetric O–H stretching and O–H bending modes of water is observed. The loading and score plots of principal component (PC) 1 are shown in Figures 3B and 3C, respectively. The contribution of PC 1 to the explained variables amounted to 99.6%. In the PC 1 loading, peaks appeared at 5250 and 5016 cm^{-1} in positive and negative directions, respectively. As the score values of PC 1 became large with increasing temperature, the positive and negative peaks in the PC 1 loading worked additive and suppressive to the average spectra at about 35 °C with which the score value is almost zero. The temperature-dependent shifts of the water bands could be interpreted based on the components of PCA.

The PCA analysis was also performed on the second derivative spectral data shown in Figure 3D (5500–5000 cm^{-1}), and the NIR spectral changes caused by the temperature variation were interpreted with the help of the second derivative spectra. It was confirmed that the absorbance band near 5200 cm^{-1} (Figure 3D) was composed of two bands at 5250 and 5180 cm^{-1}, and their intensities in the second derivative spectrum increased and decreased with a change in temperature.

In the loading plots of the PCA results for the dataset of the second

derivatives, the peaks at 5250 and 5160 cm^{-1} appeared in negative and positive directions, respectively (Figure 3E), and the score values monotonically became large with increasing temperature (Figure 3F), i.e., the PCA result indicated that the one component (PC 1, whose contribution to the explained variables was 98.9%) consisting of two elements that showed different temperature-dependent contributing patterns could explain the spectral variations of water. These results are in good agreement with the previous studies on water through NIR spectroscopy. Šašić et al. reported that more than 99% of the NIR spectra of water in the 1300–1600 nm (7700–6200 cm^{-1}) region in a temperature range of 6–76 °C could be explained with only two species of water [43]. Czarnik-Matusewicz et al. also concluded that the spectral variations in a temperature range of 25–80 °C of water were mainly due to the two spectral components that changed relative to the temperature [44]. These two components of water are referred to as two species of water depending on the two states of hydrogen bonds: strongly hydrogen bonded (SHB: ~5180 cm^{-1}) and weakly hydrogen bonded (WHB: ~5250 cm^{-1}) water species [43, 45].

Water is one of the most simple and common molecules; however, its physicochemical properties have not yet been fully unveiled [46]. The models of water structure constructed using hydrogen bonds take two major stances: (1) water is a mixture of components with different numbers of hydrogen bonds [43-45, 47, 48] and (2) water is a continuum in which the strength of hydrogen bonds vary [49, 50]. The former model is often adopted when vibrational spectroscopies such as IR, Raman, and NIR spectroscopies are

used. A water molecule has two donors in the form of hydrogen atoms and two acceptors in the form of an oxygen atom. Thus, the possible numbers of hydrogen bonds within a water molecule are 0, 1, 2, 3, and 4. The water without a hydrogen bond is called “free OH” [44], and the other four species of water with a maximum of four hydrogen bonds are grouped together into ones with asymmetrically bonded OH group [51]. When the water absorption band is theoretically resolved into these components, depending on the number of hydrogen bonds, the bands with a lower number of hydrogen bonds are assigned to those in order from higher wavenumber [43, 44, 51]. However, only two components are phenomenologically changed in NIR spectra depending on their temperature, and the explicit definition has not been established such that WHB includes the spectral components with the number of hydrogen bonds up to 2 and SHB has those with hydrogen bonds higher than 3. In the present study, the water absorbance band is discussed based on the experimental data within the framework of two spectral components as WHB (higher wavenumber: ~5250 cm^{-1}) and SHB (lower wavenumber: ~5180 cm^{-1}), and further discussion about their explicit definitions is not included herein because it is beyond the scope of this study.

Thus, the relative contribution of SHB water species to that of WHB water species can be seen to decrease with increasing temperature. The relative intensity of the bands in the second derivative is defined as I_{5180}/I_{5250}, and the values are plotted against the temperature of water (Figure 3G). The changes in the hydrogen-bond states are numerically discussed using the ratio. The other analytical method: two-dimensional

correlation spectroscopy, was also applied to the spectral data of NIR absorbance of ultrapure water and protein aqueous solutions. It was investigated whether the analytical method based on the second derivative spectra was enough to study the changes of the hydrogen bonds. The detailed discussion is shown in SI 4.

Hydrogen-bond network in protein aqueous solutions

The states of hydrogen bonds in protein aqueous solutions were investigated. Figure 4A shows the second derivative spectra in the 5500–5000 cm^{-1} region for the three types of protein solutions in ultrapure water with variation in concentration from 5 to 20 wt% having an interval of 5 wt% at 25 °C as well as over a temperature range of 5 °C to 80 °C with the 5 °C increments. The spectra of protein aqueous solutions heavily overlapped with each other. To make it easier to observe, the relative changes in the intensities of the bands due to SHB and WHB water species in the second derivative spectra with variations in temperature and protein concentration were normalized using the second derivative intensities where they had the smallest values (~5250 cm^{-1}) (Figure 4B). The relative intensities of the bands due to SHB water compared to those of WHB water increased with the increase in protein concentration regardless of the protein species, i.e., the spectra of protein solutions at 25 °C with higher concentration were similar to those of ultrapure water at temperatures lower than 25 °C. The results show similar tendency with those of the other research reporting that the contributions of the lower frequency water band were dominant in

condense environment of the solutes at the order of milli mole concentration; it is the same order of the solute concentration in this study [52].

PCA analysis was conducted on the data set including ultrapure water and protein aqueous solutions to compare the changes in the protein hydration and hydrogen-bond network of water caused by the increase in the protein concentration. Figures 4C and 4D show the plots of scores (PC 1 vs PC 2) and loadings (PC 1 and PC 2) of PCA, respectively. The score values of PC 1 for ultrapure water reduced with increasing temperature. When the loading plots in the 5300–5000 cm^{-1} region were focused, two bands at 5250 and 5160 cm^{-1} were clearly observed in the positive and negative directions, respectively. By combing the information of scores, the PCA results also indicated that the second derivative intensity at ~5250 cm^{-1} due to WHB water increased and at ~5160 cm^{-1} water owing to SHB decreased for ultrapure water with the increasing temperature. In the loading plot of PC 2, the bands at 5280 and 5216 cm^{-1} were observed in the positive and negative directions, respectively. Since the score values of PC 2 increased as the protein concentrations increased, each band was analyzed to get stronger toward the positive and negative directions with increasing protein concentration. Thus, it was interpreted that the NIR band at 5250 cm^{-1} entirely shifted to a lower frequency. For ultrapure water, however, the PC 2 scores did not monotonically change. The results indicated that the band possibly shifted to higher and lower frequencies because of the superposition of the bands due to SHB and WHB water whose relative contributions changed depending on the temperature. In practice, the wavenumber of the

band position at ~5250 cm^{-1} was confirmed to change similarly to the PCA results (Figure 4E). In the case of protein aqueous solution, the band shifted to lower frequency with increasing protein concentration, and in the case of ultrapure water, the band shifted to lower and higher frequencies at ~35 °C as a boundary.

In summary, the addition of proteins into water yields different types of hydrogen bonds between water and protein molecules; only the O–H groups are involved in the hydrogen bonds in pure water, while the N–H groups are newly added in aqueous protein solutions. However, the PCA results revealed that the hydrogen bonds in ultrapure water and hydrated water are not distinguishable via NIR spectroscopy. Namely, the hydrogen-bond network of water caused by the increment of protein concentration were analyzable based on the framework of two components that vary their relative contributions with variations in temperature. The hydrogen-bond network can be commonly investigated by the relative band intensities of the second derivatives of SHB and WHB water using the formula I_{5180}/I_{5250}. In the following discussion, the states of the hydrogen bonds between water molecules as well as between water and the protein molecules are uniformly analyzed in a word "hydrogen-bond network."

Variations in the hydrogen-bond network of the aqueous solutions of the protein with changes in the secondary structure of the protein

The states of hydrogen bonds of water along with the protein secondary structural changes were studied. Figure 5A shows the second derivative

spectra in the 5300–5050 cm^{-1} region normalized by their intensities at the minimum point at ~5250 cm^{-1} for OVA solutions in ultrapure water that were recorded for 60 min during heating, where, at each heating temperature, 12 spectra were superimposed. The states of hydrogen bonds of water in OVA solutions changed depending on the heating temperature as in the case of ultrapure water. However, the changes in the states of hydrogen bonds were not detected in NIR water bands along with the changes in the secondary structure of the protein during heating. The results indicated that NIR spectra of water only reflected the states of hydrogen bonds depending on the heating temperature but not the changes in the interactions with the protein molecules.

The protein solutions were cooled to 25 °C after 60 min of heating, and their NIR measurements were performed at 25 °C to investigate the changes in the states of hydrogen bonds caused by the interactions between proteins having native and denatured structures and water molecules. Under these conditions, the species of proteins, their concentrations, solvents, and temperatures were the common factors, whereas the only difference between samples was the secondary structures of the proteins.

For the three types of aqueous solutions of proteins that were cooled after heating for 60 min, the NIR bands in the second derivative spectra derived from the proteins were focused initially. Figure 5B shows the second derivative spectra at ~4860 and ~4630 cm^{-1} analyzed at 25 °C for the three types of protein solutions. As the lower and higher frequency shifts were detected even after cooling, the changes in the secondary structures of the

proteins were confirmed to be retained. Raman spectroscopy and ThT fluorescence techniques were used as supplementary tools to confirm that the changes in the secondary structures of the proteins were retained even after cooling, as shown in SI 1 and SI 2. Further, the NIR bands due to water were also examined. Figure 5C depicts the second derivative spectra in the 5300–5050 cm^{-1} region due to water normalized by the lowest intensity (~5250 cm^{-1}) measured at 25 °C. In the case of OVA, the intensity at 5180 cm^{-1} due to SHB water got weaker with increasing temperature. BLG, on the contrary, showed an opposite trend even though the variations in the intensities were small. The intensities did not practically change in the case of BSA although they temporarily became stronger at 60 °C. The ratios of the second derivative intensities between SHB and WHB, defined as I_{5180}/I_{5250} analyzed at 25 °C after cooling, were plotted against the heating temperatures for ultrapure water solutions and Tris-HCl buffers (Figure 5D). The aqueous solutions for the three types of proteins showed very similar ratio properties without strong dependence on their species. In the case of OVA, the characteristic ratio properties were also observed in the Tris-HCl buffer, where the proportion of SHB water species became high after the changes in the secondary structures of the proteins. As for BLG, the ratio contributed by the SHB water species tended to get lower in Tris-HCl buffer as compared to that in ultrapure water. The states of the hydrogen bonds of water associated with the changes in the secondary structures of the proteins were not detected regardless of the solvent species in BSA.

Thus, the variations of hydrogen bonds with the changes in the

secondary structures of the proteins showed different patterns depending on the protein species, and the differences depending on the solvent species were insignificant. However, the way of interactions between water and protein molecules sometimes differed slightly depending on the solvent species. Even though further discussions about these differences are not included here, these differences may be due to the degree of the surface charge embedded inside the aggregates and polymorphism beyond the secondary structures of proteins.

A model for the interactions between protein and water molecules

The variations in the hydrogen-bond states along with the changes in the secondary structures of the proteins were analyzed in the above subsection. Here, the interactions between proteins and water molecules through hydrogen bonds are discussed in depth. The changes in the NIR bands of water caused by increasing the protein concentrations were investigated. Figure 6A depicts the normalized second derivative spectra of OVA solutions in ultrapure water with the variation in protein concentration from 0 to 20 wt%. The normalization was based on the intensity at ~5250 cm^{-1}. As the protein concentration increased, the second derivative intensity at ~5180 cm^{-1} increased. The plots of the relative intensities of the bands of water (I_{5180}/I_{5250}) for the three types of aqueous solutions of the proteins are shown in Figure 6B. The relative contribution due to SHB water species compared to that of WHB water species increased for all the three types of proteins. It could be interpreted from the results that the contribution from

the SHB water species was from the hydrogen bonds newly built between protein and water molecules.

Based on the above results, the changes in the hydrogen-bond states along with the changes in the secondary structures of the proteins are summarized. First, a higher contribution of SHB water species with increased protein concentration was observed. This may be due to the fact that protein molecules are hydrated when they are in a dissolved state (Figures 7A and 7B). Then, the bands due to water in the second derivative spectra did not vary in accordance with the changes in the protein secondary structure while they were heated. The result indicated that many hydrogen bonds between water molecules and between water and protein molecules got disconnected by heating, and protein molecules were no longer being trapped by surrounding water molecules through hydrogen bonds (Figure 7C). Hydrophobic interactions are the driving force of protein folding [10, 53]. Such protein molecules tended to lose their three-dimensional structure, expose the hydrophobic amino acids buried inside them, and form intermolecular associations (Figure 7D). When it was cooled down to 25 °C after heating for 60 min, the network of hydrogen bonds was reconstructed by forming the hydrogen bonds between water molecules and water and protein molecules (Figure 7E). At this time, in the case of OVA, it was very likely that more charges on the molecular surface were buried inside, and it became more difficult to form hydrogen bonds with water molecules compared to the proteins with the native structure. In the cases of BLG and BSA, it seemed that the states of the hydrogen bonds did not significantly

change because the charges on the molecular surface might not have been drastically changed compared to those before heating. The differences in the states of hydrogen bonds caused by the solvent species for BLG may detect the variation in the surface charge due to the difference in protein conformations.

Thus, the dissolution process of proteins in water was explored by changes in the NIR absorption bands of water caused by the variations of hydrogen bonds. In aqueous solutions, protein molecules with native structures are trapped by water molecules through hydrogen bonds, and they can stably exist. When the hydrogen bonds are partly broken by heating, the proteins lose their stability and are capable of changing the structure according to their own thermodynamic properties. Because the distributions of the surface charges vary due to the secondary structural changes of proteins, the network of hydrogen bonds is thought to change.

Conclusions

In the present study, variations in the hydrogen-bond network caused by the secondary structural changes of proteins owing to denaturation were explored through NIR spectroscopy. The network was revealed to become stronger in protein aqueous solutions due to the formation of new hydrogen bonds caused by the protein hydration. Protein molecules independently existed in the surrounding water molecules when the hydrogen bonds were broken due to increasing temperature. Consequently, the protein conformations changed according to their thermodynamic property. When

they were cooled after denaturation, hydrogen bonds were formed again between water molecules and protein and water molecules while retaining the conformational changes in the protein. This reconstructed hydrogen-bond network showed different tendencies, such as getting stronger or remaining almost unchanged, than those before heating depending on the protein species. This was likely to be derived from the differences in the surface charges before and after heat denaturation according to the type of proteins.

Thus, NIR spectroscopic analysis succeeded in detecting the variations of hydrogen-bond network due to the secondary structural changes of proteins. The differences in the interactions between protein and water molecules are thought to be involved in possible pathogenesis caused by protein aggregation. From the viewpoint of protein hydration, the method is expected to be further applied for the elucidation of the mechanism of the protein structural changes as well as its aggregation.

Supporting Information

Confirmation of protein conformational changes by Raman spectroscopy, thioflavin T, and NIR band shifts, and results of two-dimensional correlation spectroscopy.

Acknowledgments

The study was supported by MEXT Leading Initiative for Excellent Young Researchers (M.I.).

References

1. Kell, G. S. Density, thermal expansivity, and compressibility of liquid water from 0. deg. to 150. deg.. Correlations and tables for atmospheric pressure and saturation reviewed and expressed on 1968 temperature scale. *J. Chem. Eng. Data* **1975**, *20*(1), 97-105.
2. Lang, E. W.; Lüdemann, H. D. Anomalies of liquid water. *Angew. Chem. Int. Ed.* **1982**, *21*(5), 315-329.
3. Fogarty, A. C.; Laage, D. Water dynamics in protein hydration shells: the molecular origins of the dynamical perturbation. *J. Phys. Chem. B* **2014**, *118*(28), 7715-7729.
4. Perticaroli, S.; Ehlers, G.; Stanley, C. B.; Mamontov, E.; O'Neill, H.; Zhang, Q.; Cheng, X.; Myles, D. A. A.; Katsaras, J.; Nickels, J. D. Description of hydration water in protein (green fluorescent protein) solution. *J. Am. Chem. Soc.* **2017**, *139*(3), 1098-1105.
5. Bagchi, B. Water dynamics in the hydration layer around proteins and micelles. *Chem. Rev.* **2005**, *105*(9), 3197-3219.
6. Pratt, L. R.; Pohorille, A. Hydrophobic effects and modeling of biophysical aqueous solution interfaces. *Chem. Rev.* **2002**, *102*(8), 2671-2692.
7. Nandi, N.; Bhattacharyya, K.; Bagchi, B. Dielectric relaxation and solvation dynamics of water in complex chemical and biological systems. *Chem. Rev.* **2000**, *100*(6), 2013-2046.
8. Robinson, G. W. *Water in biology, chemistry, and physics: experimental overviews and computational methodologies (Vol. 9)*. World Scientific:

Singapore, 1996.

9. Deal, A. M.; Rapf, R. J.; Vaida, V. Water–air interfaces as environments to address the water paradox in prebiotic chemistry: a physical chemistry perspective. *J. Phys. Chem. A* **2021**, *125*(23), 4929-4942.
10. Bellissent-Funel, M. C.; Hassanali, A.; Havenith, M.; Henchman, R.; Pohl, P.; Sterpone, F.; Van Der Spoel D.; Xu, Y.; Garcia, A. E. Water determines the structure and dynamics of proteins. *Chem. Rev.* **2016***, 116*(13), 7673-7697.
11. Halle, B. Protein hydration dynamics in solution: a critical survey. *Philos. Trans. R. Soc. Lond., **B**, Biol. Sci.* **2004**, *359*(1448), 1207-1224.
12. Grebenkov, D. S.; Goddard, Y. A.; Diakova, G.; Korb, J. P.; Bryant, R. G. Dimensionality of diffusive exploration at the protein interface in solution. *J. Phys. Chem. B* **2009**, *113*(40), 13347-13356.
13. Makarov, V. A.; Andrews, B. K.; Smith, P. E.; Pettitt, B. M. Residence times of water molecules in the hydration sites of myoglobin. *Biophy. J.* **2000**, *79*(6), 2966-2974.
14. Svergun, D. I.; Richard, S.; Koch, M. H. J.; Sayers, Z.; Kuprin, S.; Zaccai, G. Protein hydration in solution: experimental observation by x-ray and neutron scattering. *Proc. Natl. Acad. Sci.* **1998**, *95*(5), 2267-2272.
15. Fujisawa, T.; Uruga, T.; Yamaizumi, Z.; Inoko, Y.; Nishimura, S.; Ueki, T. The hydration of Ras p21 in solution during GTP hydrolysis based on solution X-ray scattering profile. *J. Biochem.* **1994**, *115*(5), 875–880.
16. Perkins, S. J.; Smith, K. F.; Kilpatrick, J. M.; Volanakis, J. E.; Sim, R. B. Modelling of the serine-proteinase fold by X-ray and neutron scattering

and sedimentation analyses: occurrence of the fold in factor D of the complement system. *Biochem. J.* **1993**, *295*(1), 87–99.

17. Iben, I. E. T.; Braunstein, D.; Doster, W.; Frauenfelder, H.; Hong, M. K.; Johnson, J. B.; Luck, S.; Ormos, P.; Schulte, A.; Steinbach, P. J. et al. Glassy behavior of a protein. *Phys. Rev. Lett.* **1989**, *62*, 1916-1919.
18. Bizzarri, A. R.; Wang, C. X.; Chen, W. Z.; Cannistraro, S. Hydrogen bond analysis by MD simulation of copper plastocyanin at different hydration levels. *Chem. Phys.* **1995**, *201*(2-3), 463-472.
19. Tirado-Rives, J.; Jorgensen, W. L. Molecular dynamics of proteins with the OPLS potential functions. Simulation of the third domain of silver pheasant ovomucoid in water. *J. Am. Chem. Soc.* **1990**, *112*(7), 2773-2781.
20. Bizzarri, A. R.; Cannistraro, S. Molecular dynamics of water at the protein− solvent interface. *J. Phys. Chem. B* **2002**, *106*(26), 6617-6633.
21. Kocherbitov, V.; Latynis, J.; Misiunas, A.; Barauskas, J.; Niaura, G. Hydration of lysozyme studied by Raman spectroscopy. *J. Phys. Chem. B* **2013**, *117*(17), 4981-4992.
22. Paolantoni, M.; Sassi, P.; Morresi, A.; Santini, S. Hydrogen bond dynamics and water structure in glucose-water solutions by depolarized Rayleigh scattering and low-frequency Raman spectroscopy. *J. Chem. Phys.* **2007**, *127*(2), 07B606.
23. Davis, J. G.; Rankin, B. M.; Gierszal, K. P.; Ben-Amotz, D. On the cooperative formation of non-hydrogen-bonded water at molecular hydrophobic interfaces. *Nat. Chem.* **2013**, *5*(9), 796-802.

24. Movasaghi, Z.; Rehman, S.; Rehman, I. U. Raman spectroscopy of biological tissues. *Appl. Spectrosc. Rev.* **2007**, *42*(5), 493-541.
25. Majzner, K.; Czamara, K.; Pacia, M. Z.; Dybas, J.; Bik, E.; Chrabaszcz, K.; Wiercigroch, E.; Dorosz, A.; Wislocka, A.; Marzec, K. M. et al. Vibrational imaging of proteins: changes in the tissues and cells in the lifestyle disease studies. In *Vibrational Spectroscopy in Protein Research*. Academic Press: London, 2020; pp 177-218.
26. Liu, Y.; Cho, R. K.; Sakuri, K.; Miura, T.; Ozaki, Y. Studies on spectra/structure correlations in near-infrared spectra of proteins and polypeptides. Part I: A marker band for hydrogen bonds. *Appl. Spectrosc.* **1994**, *48*(10), 1249-1254.
27. Ishigaki, M.; Morimoto, K.; Chatani, E.; Ozaki, Y. Exploration of insulin amyloid polymorphism using Raman spectroscopy and imaging. *Biophys. J.* **2020**, *118*(12), 2997-3007.
28. Ishigaki, M.; Ozaki, Y. Near-infrared spectroscopy and imaging in protein research. In *Vibrational Spectroscopy in Protein Research*. Academic Press: London, 2020; pp 143-176.
29. Czarnecki, M. A.; Morisawa, Y.; Futami, Y.; Ozaki, Y. Advances in molecular structure and interaction studies using near-infrared spectroscopy. *Chem. Rev.* **2015**, *115*(18), 9707-9744.
30. Workman, J.; Weyer, L. *Practical guide and spectral atlas for interpretive near-infrared spectroscopy*, CRC Press: Boca Raton, 2012.
31. Ozaki, Y.; Huck, C.; Tsuchikawa, S.; Engelsen, S. B. *Near-Infrared Spectroscopy: Theory, Spectral Analysis, Instrumentation, and*

Applications, Springer: Berlin/Heidelberg, 2021.

32. Siesler, H. W.; Ozaki, Y.; Kawata, S.; Heise, H. M. (Eds.). *Near-infrared spectroscopy: principles, instruments, applications*. John Wiley & Sons: Weinheim, 2002.
33. Nielsen L.; Frokjaer S.; Brange J.; Uversky V. N.; Fink A. L. Probing the mechanism of insulin fibril formation with insulin mutants. *Biochemistry* **2001**, *40*(28), 8397-8409
34. Robert, P.; Devaux, M. F.; Mouhous, N.; Dufour, E. Monitoring the secondary structure of proteins by near-infrared spectroscopy. *Appl. Spectrosc.* **1999**, *53*(2), 226-232.
35. Izutsu, K. I.; Fujimaki, Y.; Kuwabara, A.; Hiyama, Y.; Yomota, C.; Aoyagi, N. Near-infrared analysis of protein secondary structure in aqueous solutions and freeze-dried solids. *J. pharm. Sci.* **2006**, *95*(4), 781-789.
36. Hug, W.; Chalmers, J. M.; Griffith, P. R. *Handbook of vibrational spectroscopy.* John Wiley & Sons: New Yolk, 2002.
37. Wu, P.; Siesler, H. W. The assignment of overtone and combination bands in the near infrared spectrum of polyamide 11. *J. Near Infrared Spectrosc.* **1999**, *7*(2), 65-76.
38. Kitabatake, N.; Ishida, A.; Doi, E. Physicochemical and functional properties of hen ovalbumin dephosphorylated by acid phosphatase. *Agric. Biol. Chem.* **1988**, *52*(4), 967-973.
39. Lippert, J. L.; Tyminski, D.; Desmeules, P. J. Determination of the secondary structure of proteins by laser Raman spectroscopy. *J. Am.*

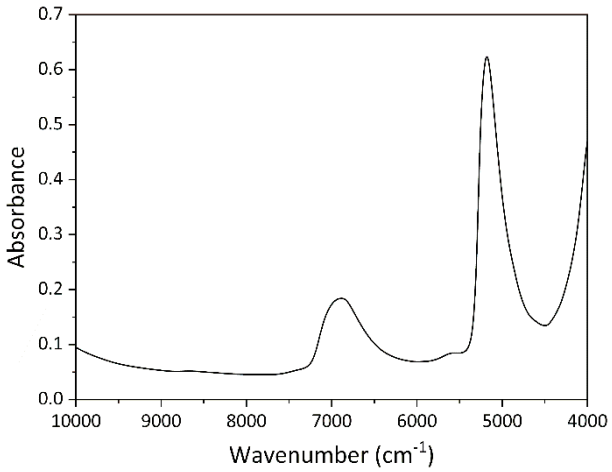
Chem. Soc. **1976**, *98*(22), 7075-7080.

40. Yamasaki, M.; Yano, H.; Aoki, K. Differential scanning calorimetric studies on bovine serum albumin: I. Effects of pH and ionic strength. *Int. J. Biol. Macromol.* **1990**, *12*(4), 263-268.
41. Photchanachai, S.; Mehta, A.; Kitabatake, N. Heating of an ovalbumin solution at neutral pH and high temperature. *Biosci. Biotechnol. Biochem.* **2002**, *66*(8), 1635-1640.
42. Lapanje, S.; Poklar, N. Calorimetric and circular dichroic studies of the thermal denaturation of β-lactoglobulin. *Biophys. Chem.* **1989**, *34*(2), 155-162.
43. Šašic, S.; Segtnan, V. H.; Ozaki, Y. Self-modeling curve resolution study of temperature-dependent near-infrared spectra of water and the investigation of water structure. *J. Phys. Chem. A* **2002**, *106*, 760-766.
44. Czarnik-Matusewicz, B.; Pilorz, S.; Hawranek, J. P. Temperature-dependent water structural transitions examined by near-IR and mid-IR spectra analyzed by multivariate curve resolution and two-dimensional correlation spectroscopy. *Anal. Chim. Acta* **2005**, 544 (1-2), 15-25.
45. Walrafen, G. E.; Fisher, M. R.; Hokmabadi, M. S.; Yang, W. H. Temperature dependence of the low-and high-frequency Raman scattering from liquid water. *J. Chem. Phys.* **1986**, *85*(12), 6970-6982.
46. Pettersson, L. G. M.; Henchman, R. H.; Nilsson, A. Water – The Most Anomalous Liquid. *Chem. Rev.* **2016**, *116*(13), 7459-7462.
47. Segtnan, V. H.; Šašić, Š.; Isaksson, T.; Ozaki, Y. Studies on the structure of water using two-dimensional near-infrared correlation spectroscopy

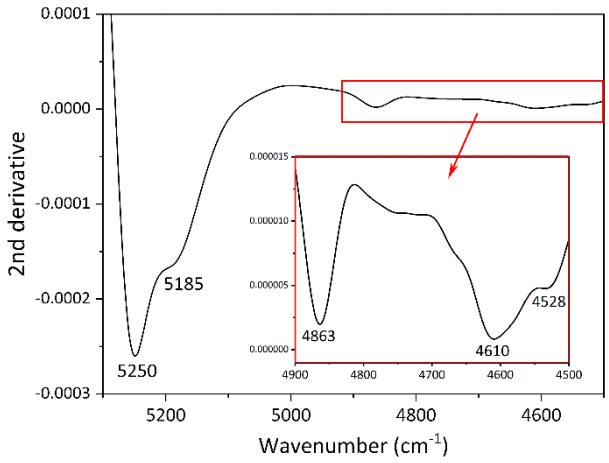
and principal component analysis. *Anal. Chem.* **2001**, 73(13), 3153-3161.

48. Maeda, H.; Ozaki, Y.; Tanaka, M.; Hayashi, N.; Kojima, T. J. Near infrared spectroscopy and chemometrics studies of temperature-dependent spectral variations of water: relationship between spectral changes and hydrogen bonds. *J. Near Infrared Spectrosc.* **1995**, *3*(4), 191-201.
49. Okajima, H.; Ando, M.; Hamaguchi, H.O. Formation of "nano-ice" and density maximum anomaly of water. *Bull. Chem. Soc. Jpn* **2018**, *91*(6), 991-997.
50. Luck, W. A. P. A model of hydrogen-bonded liquids. *Angew. Chem. Int. Ed.* **1980**, *19*(1), 28-41.
51. Tomlinson-Phillips, J.; Davis, J.; Ben-Amotz, D.; Spångberg, D.; Pejov, L.; Hermansson, K. Structure and dynamics of water dangling OH bonds in hydrophobic hydration shells. Comparison of simulation and experiment. *J. Phys. Chem. A* **2011**, *115*(23), 6177-6183.
52. Giuffrida, S.; Cupane, A.; Cottone, G. "Water Association" Band in Saccharide Amorphous Matrices: Role of Residual Water on Bioprotection. *Int. J. Mol. Sci.* **2021**, *22*(5), 2496.
53. Kauzmann, W. Some factors in the interpretation of protein denaturation. In *Advances in protein chemistry*. Academic Press: Massachusetts, 1959; pp 1-63.

(A)


Absorbance
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
10000
9000
8000
7000
6000
5000
4000
Wavenumber (cm-1)

(B)


2nd derivative
0.0001
0.0000
-0.0001
-0.0002
-0.0003
5185
5250
0.000015
0.000010
0.000005
0.000000
4863
4610
4528
4900
4800
4700
4600
4500
5200
5000
4800
4600
Wavenumber (cm-1)

(C)

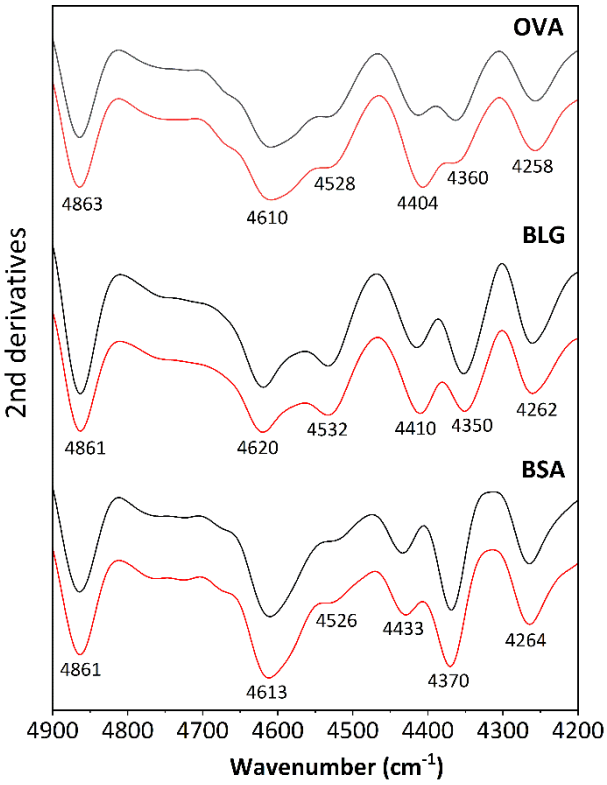
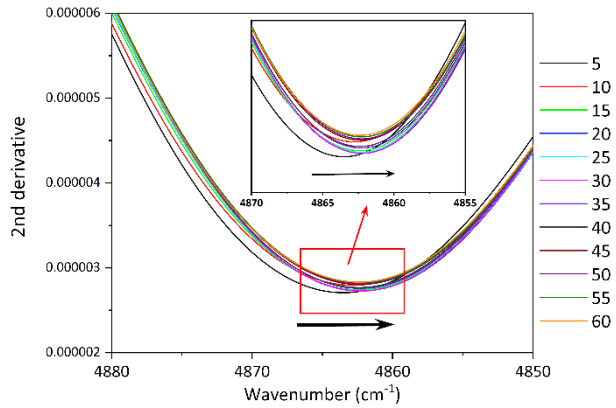
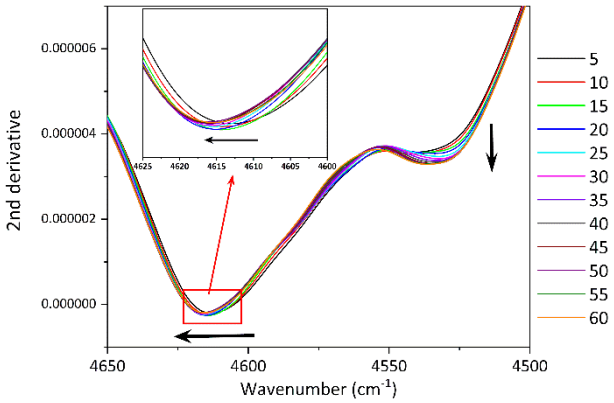


Figure 1: (A) NIR absorbance spectrum in the region of 10000–4000 cm^{-1} and (B) its second derivative spectrum in the 5300–4500 cm^{-1} region of a 10wt% OVA ultrapure water at 25 °C. Inset in (B), an enlargement of the 4900–4500 cm^{-1} region. (C) The second derivative spectra in the 4900–4200 cm^{-1} region for 10 wt% aqueous solutions of three types of proteins at 25 °C. The black and red lines correspond to ultrapure water and Tris-HCl buffer solution media, respectively. The scales of the vertical axis for the second derivatives were unified with respect to all the spectra.

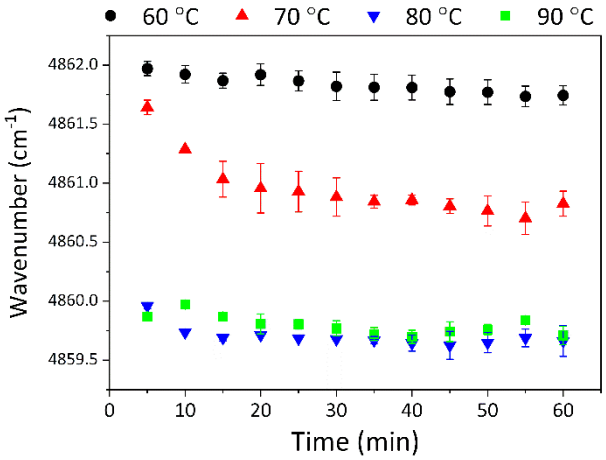
(A)


0.000006
0.000005
0.000004
0.000003
0.000002
2nd derivative
4880
4870
4860
4850
Wavenumber (cm-1)
4870
4865
4860
4855
5
10
15
20
25
30
35
40
45
50
55
60

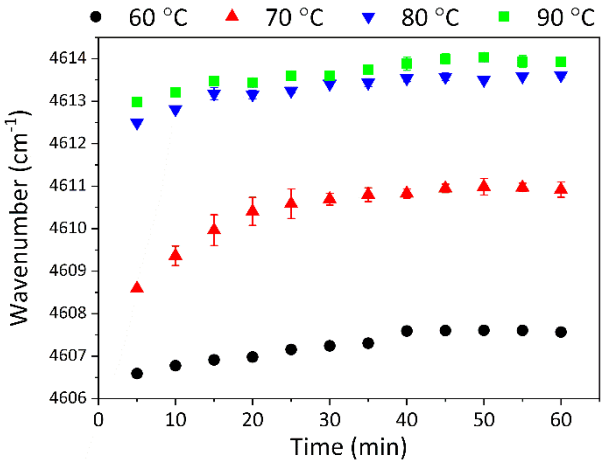
(B)


0.000006
0.000004
0.000002
0.000000
2nd derivative
4650
4600
4550
4500
Wavenumber (cm-1)
5
10
15
20
25
30
35
40
45
50
55
60

(C)


60 °C
70 °C
80 °C
90 °C
4862.0
4861.5
4861.0
4860.5
4860.0
4859.5
Wavenumber (cm-1)
0
10
20
30
40
50
60
Time (min)

(D)



(E)

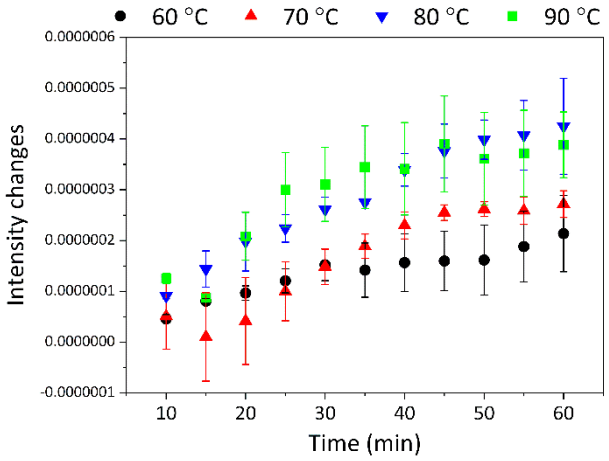
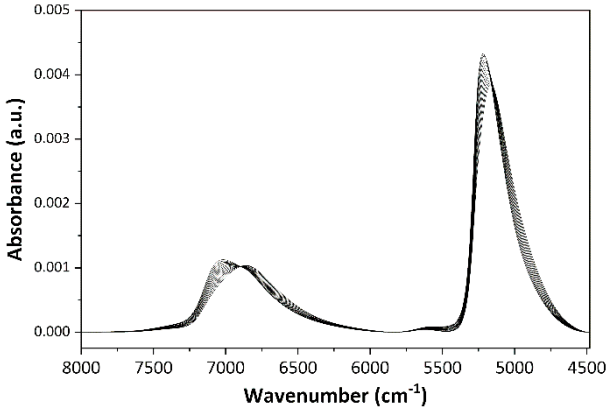
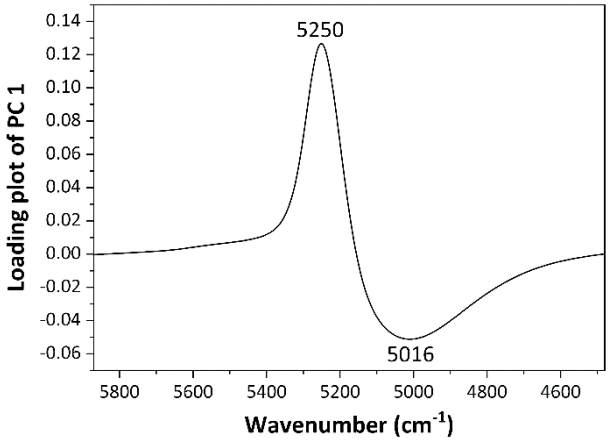


Figure 2: Second derivative spectra in the (A) 4880–4850 cm^{-1} and (B) 4650–4500 cm^{-1} regions that were recorded every 5 min during heating at 70 °C for 60 min of the OVA ultrapure solution. The arrows explicit the spectral variations with increasing temperature. (C) and (D) are the plots of the band positions with the standard error (SE) near 4860 and 4610 cm^{-1}, respectively, and (E) is the second derivative intensities at ~4510 cm^{-1} with the SE along with heating time at four different temperatures.

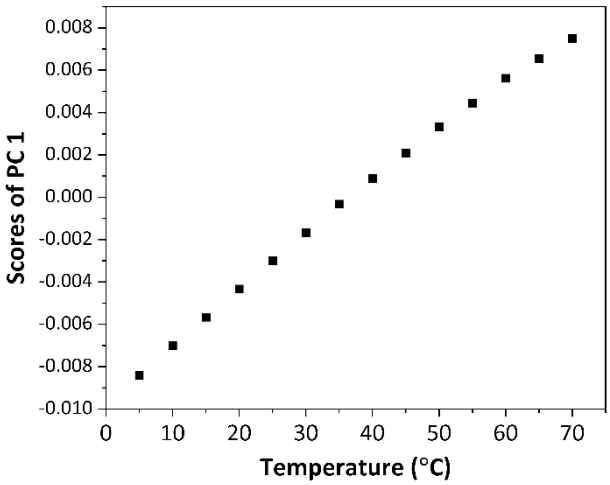
(A)


Absorbance (a.u.)
0.005
0.004
0.003
0.002
0.001
0.000
8000
7500
7000
6500
6000
5500
5000
4500
Wavenumber (cm-1)

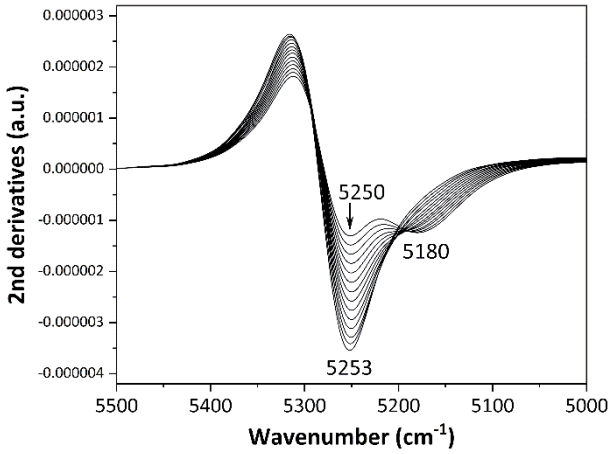
(B)


Loading plot of PC 1
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
-0.02
-0.04
-0.06
5250
5016
5800
5600
5400
5200
5000
4800
4600
Wavenumber (cm-1)

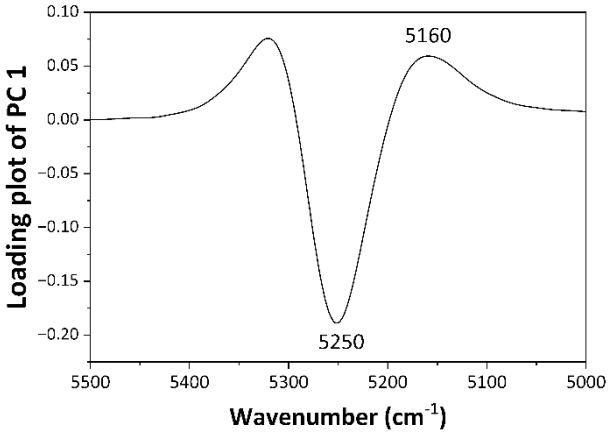
(C)


Scores of PC 1
0.008
0.006
0.004
0.002
0.000
-0.002
-0.004
-0.006
-0.008
-0.010
0
10
20
30
40
50
60
70
Temperature (°C)

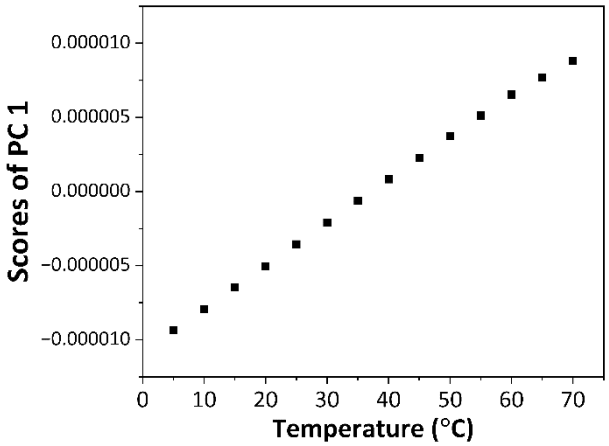
(D)


0.000003
0.000002
0.000001
0.000000
-0.000001
-0.000002
-0.000003
-0.000004
2nd derivatives (a.u.)
5250
5180
5253
5500
5400
5300
5200
5100
5000
Wavenumber (cm-1)

(E)


0.10
0.05
0.00
−0.05
−0.10
−0.15
−0.20
Loading plot of PC 1
5160
5250
5500
5400
5300
5200
5100
5000
Wavenumber (cm-1)

(F)


0.000010
0.000005
0.000000
−0.000005
−0.000010
Scores of PC 1
0
10
20
30
40
50
60
70
Temperature (°C)

(G)

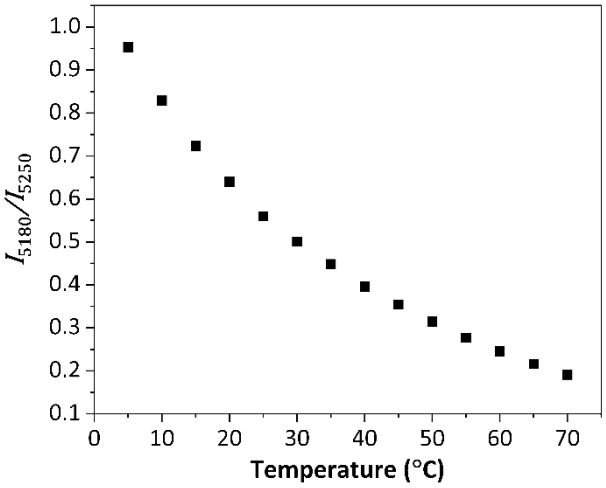
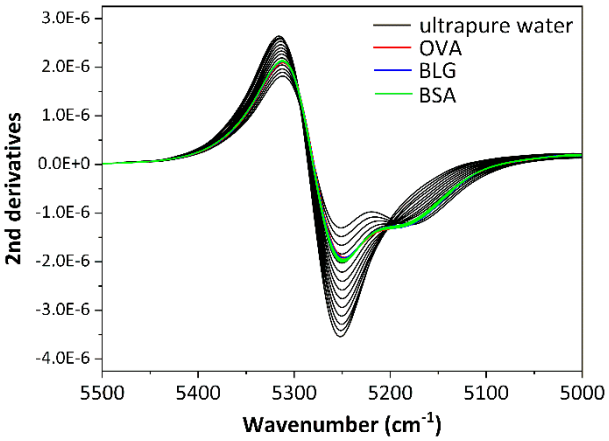
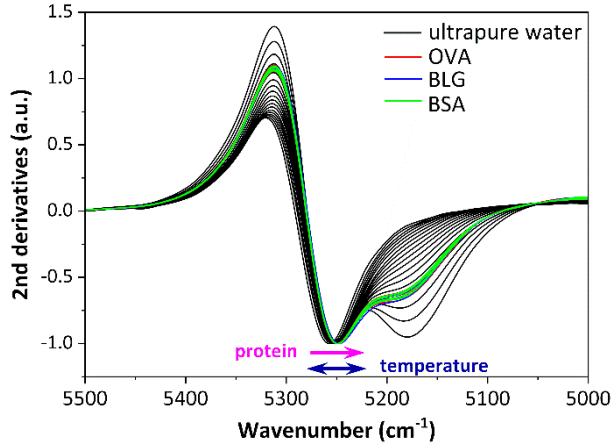


Figure 3: (A) NIR absorbance spectra of ultrapure water in the 8000–4480 cm^{-1} region measured in a temperature range of 5–70 °C at 5 °C intervals. (B) Loading and (C) score plots of PC1 for PCA conducted for the data set shown in (A). (D) Second derivative spectra in the 5500–5000 cm^{-1} region calculated for the data set (A). (E) Loading and (F) score plots of PC1 for PCA conducted for the data set shown in (D). (G) The plots of the ratio for the second derivative intensities, defined as I_{5180}/I_{5250}, with the SE depending on the temperature.

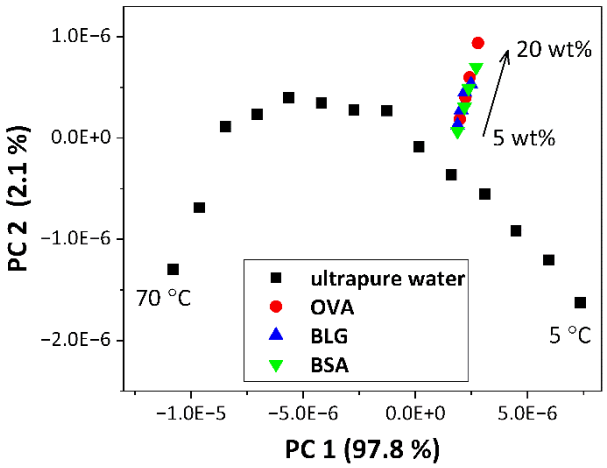
(A)


ultrapure water
OVA
BLG
BSA
2nd derivatives
3.0E-6
2.0E-6
1.0E-6
0.0E+0
-1.0E-6
-2.0E-6
-3.0E-6
-4.0E-6
5500
5400
5300
5200
5100
5000
Wavenumber (cm-1)

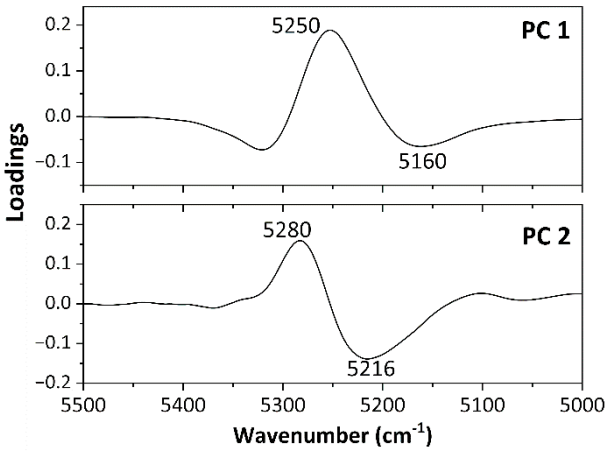
(B)


ultrapure water
OVA
BLG
BSA
2nd derivatives (a.u.)
1.5
1.0
0.5
0.0
-0.5
-1.0
protein
temperature
5500
5400
5300
5200
5100
5000
Wavenumber (cm-1)

(C)


1.0E-6
0.0E+0
-1.0E-6
-2.0E-6
PC 2 (2.1 %)
20 wt%
5 wt%
70 °C
5 °C
ultrapure water
OVA
BLG
BSA
-1.0E-5
-5.0E-6
0.0E+0
5.0E-6
PC 1 (97.8 %)

(D)



(E)

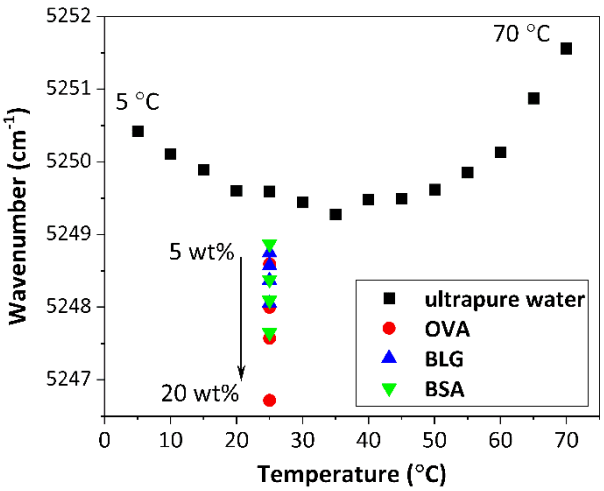
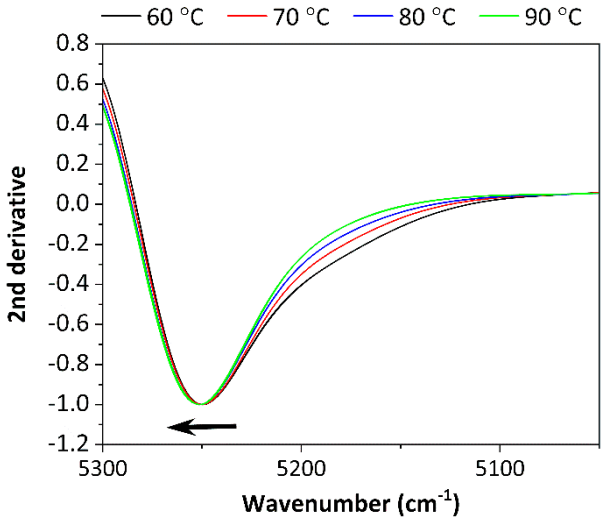


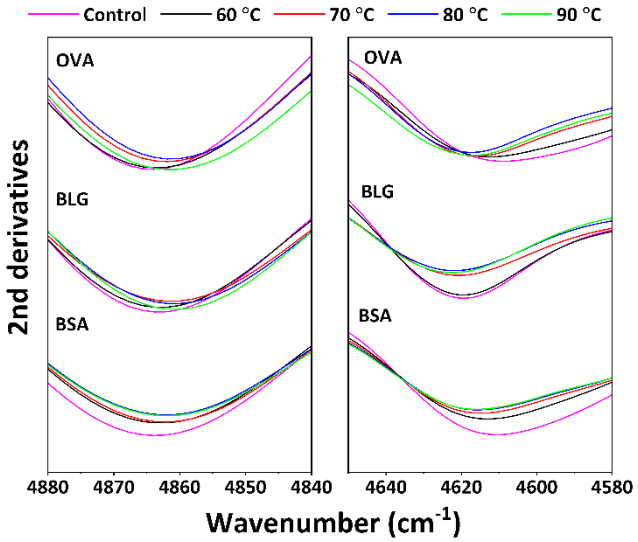
Figure 4: (A) Second derivative spectra in the 5500–5000 cm^{-1} region of ultrapure water solutions for three types of proteins with concentrations of 5–20 wt% with increments of 5 wt% (red: OVA, blue: BLG, green: BSA) and ultrapure water within a temperature range of from 5 °C to 80 °C with 5 °C increments (black). (B) Normalized second derivative spectra using the second derivative intensities having the smallest values at ~5250 cm^{-1}. The arrows explicit the band shifts according to the increasing of protein concentration and temperature. (C) Plots of the scores (PC 1 vs PC 2) and (D) loadings (PC 1 and PC 2) of PCA. (E) Wavenumbers of the band near

5250 cm^{-1} in the second derivative spectra due to WHB water against the heating temperature.

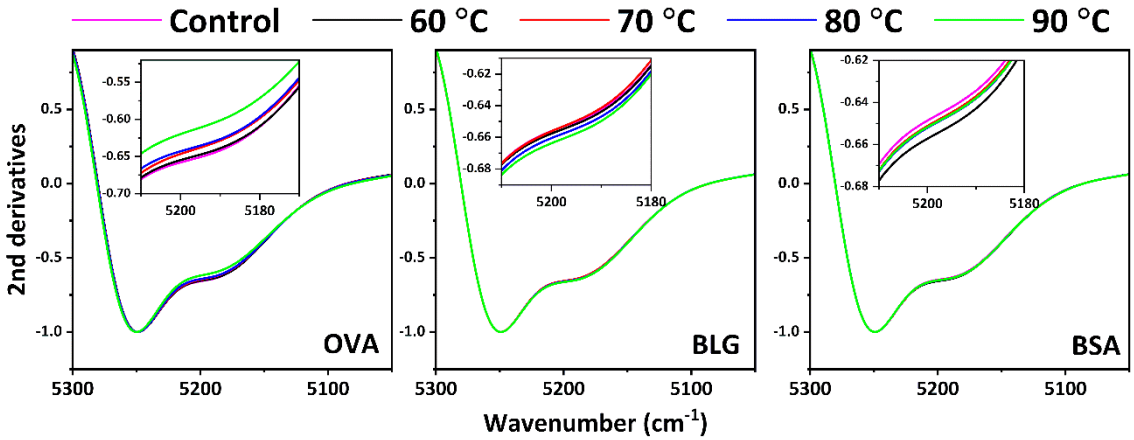
(A)



(B)



(C)



(D)

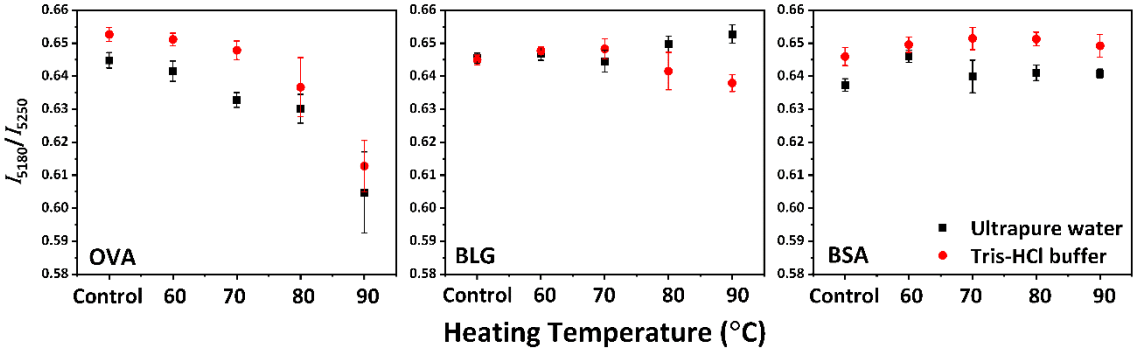
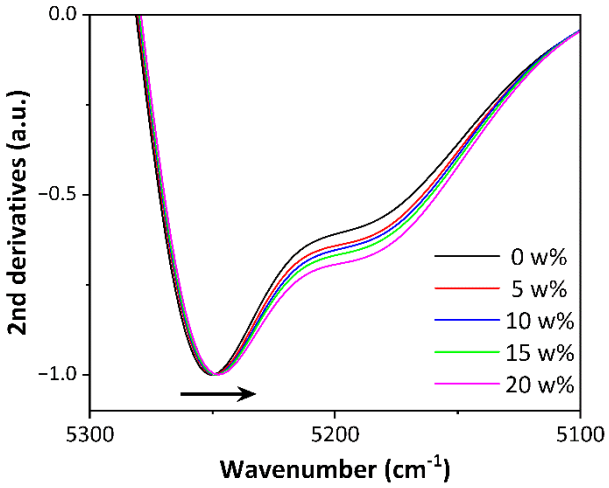


Figure 5: (A) Second derivative spectra in the 5300–5000 cm^{-1} region normalized at the second derivative intensities at ~5250 cm^{-1} for OVA ultrapure water solutions heated at four different temperatures. The bands showed higher frequency shifts with increasing temperature. Second derivative spectra at approximately (B) 4860 and 4630 cm^{-1}; (C) 5200 cm^{-1} after cooling to 25 °C. The sample that is not heated is assigned as a control. (D) Plots of the ratios of the second derivative intensities with the SE between SHB and WHB, defined as I_{5180}/I_{5250}, against the heating temperature.

(A)



(B)

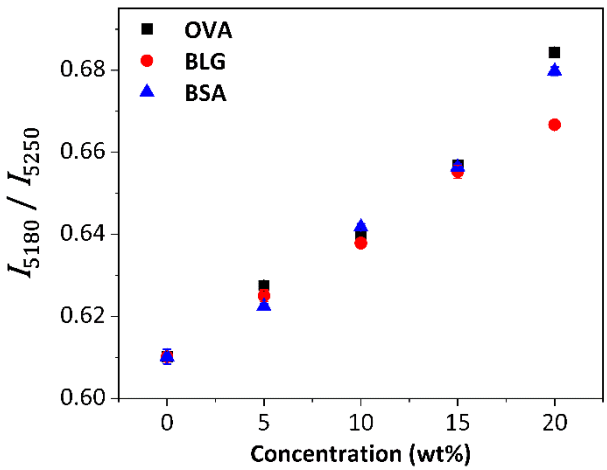


Figure 6: (A) Second derivative spectra in the 5300–5100 cm^{-1} region normalized by the second derivative intensities at ~5250 cm^{-1} for the OVA ultrapure water solutions with concentrations of 0–20 wt%. The arrows explicit the spectral variations with increasing protein concentration. (B) Plots of the ratios for the second derivative intensities with the SE between SHB and WHB, defined as I_{5180}/I_{5250}, against the concentration of proteins.

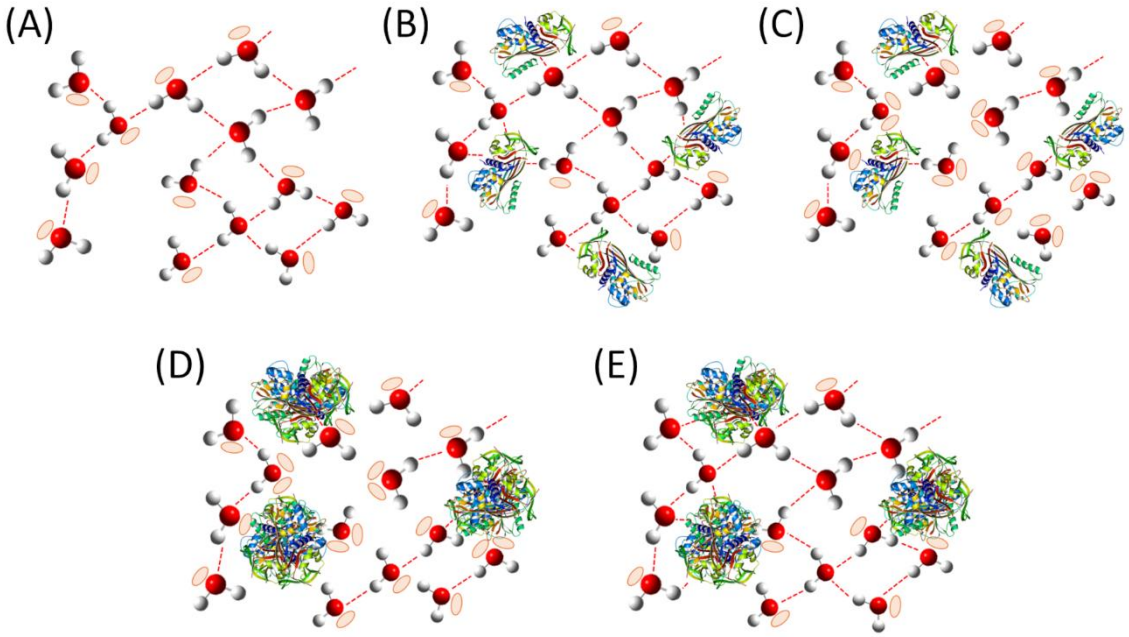


Figure 7: Schematic model of the hydrogen-bond network for (A) ultrapure water and (B)–(E) protein aqueous solutions. (A), (B), and (E) are the states assumed at 25 °C, and (C) and (D) are the states at a denaturation temperature. The red dotted lines depict the hydrogen bonds, and the orange circles show the electron pairs of oxygen atoms.

Supporting Information

Variations in the Protein Hydration and Hydrogen-Bond Network of Water Molecules Induced by the Changes in the Secondary Structures of Proteins Studied through Near-Infrared Spectroscopy

Mika Ishigaki[1]*, Yoshiki Kato[1], Eri Chatani[2], Yukihiro Ozaki[3]

[1]*Institute of Agricultural and Life Sciences, Academic Assembly, Shimane University, 1060 Nishikawatsu, Matsue, Shimane, 690-8504, Japan*
[2] Department of Chemistry, Graduate School of Science, Kobe University, Nada, Kobe, 657-8501, Japan
[3]School of Biological and Environmental Sciences, Kwansei Gakuin University, 1 Gakuen-Uegahara, Sanda, Hyogo 669-1330, Japan

*Authors to whom correspondence should be sent.
*E-mail: ishigaki@life.shimane-u.ac.jp (M.I.)

Table of Contents

SI 1: Confirmation of protein secondary structural changes via Raman spectroscopy

Figures S1A and S1B show the mean Raman spectra of three types of protein aqueous solutions for ultrapure water in the amide I (~1660 cm^{-1}) and amide III (~1250 cm^{-1}) band regions, respectively. In the amide I wavenumber region, the bands shifted to higher wavenumber at higher temperatures; however, in the amide III wavenumber region, the band intensities of the lower wavenumbers were observed to get more intense. The spectral variations with increasing temperature corresponded to that derived from the protein secondary structural changes from α-helix to β-sheet[18, 21]. The results confirmed that the secondary structural changes were retained even after cooling. For protein aqueous solutions in Tris-HCl buffer, similar spectral variations were also detected.

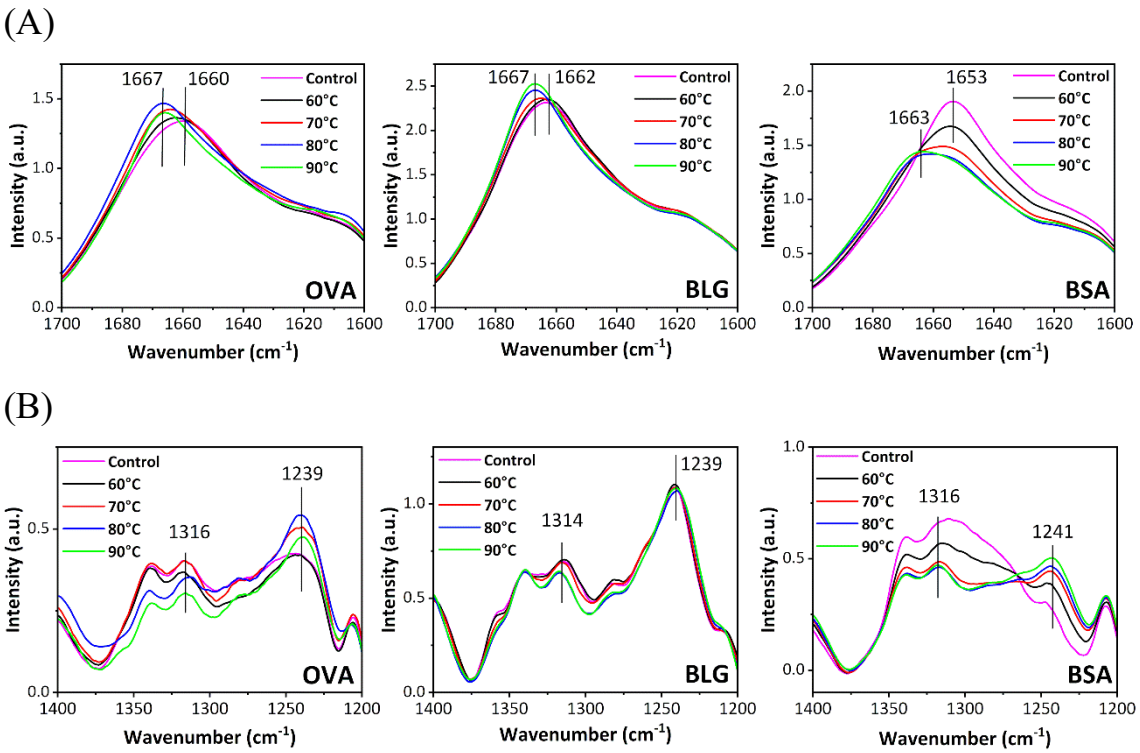


Figure S1: Raman spectra of three types of protein aqueous solutions in the wavenumber regions due to (A) amide I and (B) amide III.

SI 2: Confirmation of the changes in the secondary structure of protein by fluorescent pigment thioflavin T

Figure S2 depicts the ThT fluorescence intensity based on the heating temperature. The drastic changes in the fluorescence intensities were detected for the three types of proteins at temperatures higher than 70 °C as compared with those of the controls without heating. Thus, aggregation, such as an amyloid fibril, was confirmed to be formed along with the secondary structural changes of proteins.

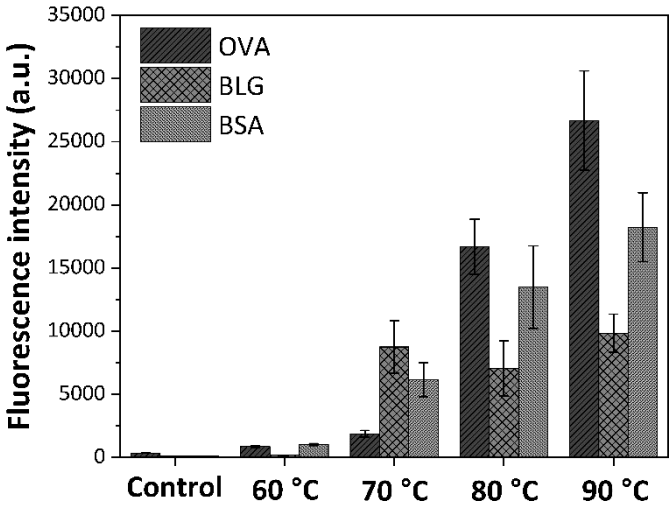


Figure S2: ThT fluorescence intensity, with the standard error (SE) depending on the heating temperature.

SI 3: Plots of the band positions of the combination bands of amide modes for BLG and BSA

The secondary structural changes of protein in the course of denaturation by heating were monitored by plotting the band positions of the combination bands of amide modes for BLG and BSA (Figure S3).

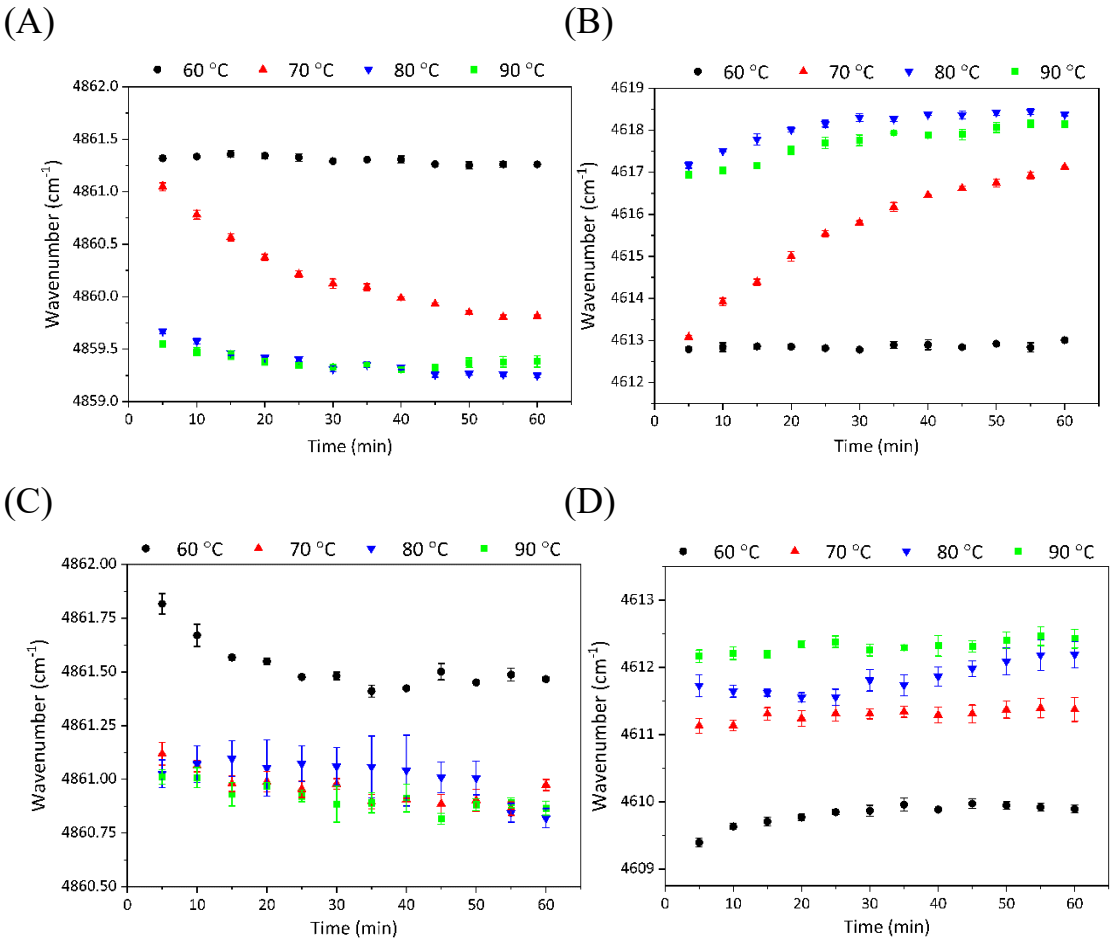


Figure S3: Plots of the band positions of the combination bands of amide modes at ~4860 and ~4610 cm^{-1} with the SE for (A, B) BLG and (C, D) BSA in ultrapure water.

SI 4: Application of the two-dimensional correlation spectroscopy to the spectral data of NIR absorbance of ultrapure water and protein aqueous solutions.

The two-dimensional correlation spectroscopy (2D-COS) analysis was applied to the spectral dataset of NIR absorbance of ultrapure water and protein aqueous solutions. The 2D-COS analysis was performed using the 2D-Shige software developed by Shigeaki Morita (Osaka Electro-Communication University) [S1]. The correlation between the bands appeared in one-dimensional spectrum varying with the perturbation given to the sample can be visually analyzed.

Temperature-dependent NIR absorbance spectra of ultrapure water

The 2D-COS analysis was executed to the dataset including NIR absorbance spectra of ultrapure water in the 8000-4800 cm^{-1} region obtained over a temperature range of 5–70 °C at a 5 °C interval. The spectra were normalized by the area in the wavenumber region. The synchronous and asynchronous spectra in the 5400-4900 and 7500-6300 cm^{-1} regions are shown in Figure S4. For simplicity, the 5400-4900 cm^{-1} region is focused to discuss.

In the 5400-4900 cm^{-1} region, two auto peaks appeared around 5250 and 5020 cm^{-1}, and two cross peaks approximately (5250, 5020) and (5020, 5250) in both side of the diagonal appeared in the synchronous spectra (Figure S3A). Since these two off diagonal peaks showed negative correlation, the peaks at 5250 and 5020 cm^{-1} increased and decreased with increasing temperature, respectively.

In the asynchronous spectra, two peaks around (5180, 5255) and (5255, 5180) were observed and the results suggested that the changes of two peaks are not completely synchronized. Since the signs of these peaks in synchronous and asynchronous were both positive, the peak at 5255 cm^{-1} was proved to increase after decreasing of the peak at 5180 cm^{-1}. That is, the hydrogen bonded water was dissociated into free form due to the increment of temperature. Furthermore, two band position corresponded to those of the peaks in the second derivative spectra shown in the main text that captured the changes in curvature of the absorbance spectra. Thus, spectral variations correlating with temperature variation can be discussed in the changes in

spectral curvature as the second derivative spectra at 5180 and 5255 cm^{-1}. This is consistent with our results using the second derivative spectra.

Protein aqueous solutions with concentration variation

Figure S5 shows the NIR absorbance spectra of protein aqueous solutions of OVA with its concentration variation. It was made clear that the intensities of the bands due from proteins around 4850 and 4600 cm^{-1} changed depending on the protein concentration. Under the influence of protein concentration, the variations of water absorbance in the normalized spectrum based on the area in the whole range (8000-4480 cm^{-1}) became large. Since the uncertainties of the intensity of water band caused by protein concentration sensitively reflect the 2D-COS results, NIR absorbance spectra of protein aqueous solutions with its concentration variations can't be correctly analyzed by 2D-COS analysis.

The NIR bands due to water (~5200 cm^{-1}) and protein (~4850, ~4600 cm^{-1}) overlap each other when the 5500-4480 cm^{-1} region is focused. In the second derivative spectra, on the other hand, the bands due to water and proteins can be separated, and the band due to water can be independently analyzed (Figure 1B in the main text).

Time-dependent NIR absorbance spectra of OVA solution heated at 70 °C

The NIR absorbance spectra with heating at 70°C for 60 minutes were analyzed using 2D-COS. At the temperature, the structural changes in proteins should be most clearly observed in time course during the heating (Figure 2C and 2D in the main text). Figures S6A and S6B depict the synchronous and asynchronous spectra in the 5400-4900 cm^{-1} regions. One auto peak around 5215 cm^{-1} appeared in the synchronous spectra, and two peaks around (5040, 5220) and (5220, 5040) were observed in the asynchronous spectra.

From the signs of the intensities in synchronous and asynchronous spectra at (5040, 5220), the variations of the band at 5040 cm^{-1} occurred before those at 5220 cm^{-1}. That is, hydrogen bonds seemed to dissociate with heating. However, the peak intensity at (5040, 5220) extracted in the heating system was tiny by three orders of the magnitude compared to those at (5180, 5260) in the asynchronous spectra derived from the temperature-dependent NIR spectra of ultrapure water. Thus, it is difficult to conclude that the

hydrogen bond network significantly changed during the heating at 70°C. The 2D-COS results were consistent with those extracted using the second derivative spectra; the second derivative spectra recorded every five minutes did not show the variations even though the heating time past and the protein conformation changed.

NIR absorbance spectra of OVA ultrapure water measured at 25°C after heating at 60, 70, 80, and 90 °C for 60 min

The auto peak appeared around 5180 cm^{-1}, and four cross peaks approximately (5180, 5300), (4900, 5180), (5300, 5180), and (5180, 4900) in both side of the diagonal appeared in the synchronous spectra (Figure S7A). In the asynchronous spectra, the peaks around (5120, 5200), (4990, 5180), (5180, 5460), (5200, 5120), (5180, 4990), and (5460, 5180) seemed to appear (Figure S7B). However, the power spectra revealed that the peaks except for the one at 5180 cm^{-1} were very small and the discussions for their changes are nonsense (Figure S7C). The intensity of the power spectra in this case was smaller than those of the power spectra calculated for the temperature-dependent absorbance spectra in the 5-70 °C by the second order. However, it was bigger than those of the power spectra extracted from the absorbance spectra recorded during the heating at 70 °C by one order. Thus, it was revealed to be reasonable for discussing the state of hydrogen bonds based on the ratio of the second derivative spectra defined as I_{5180}/I_{5250} in the main text.

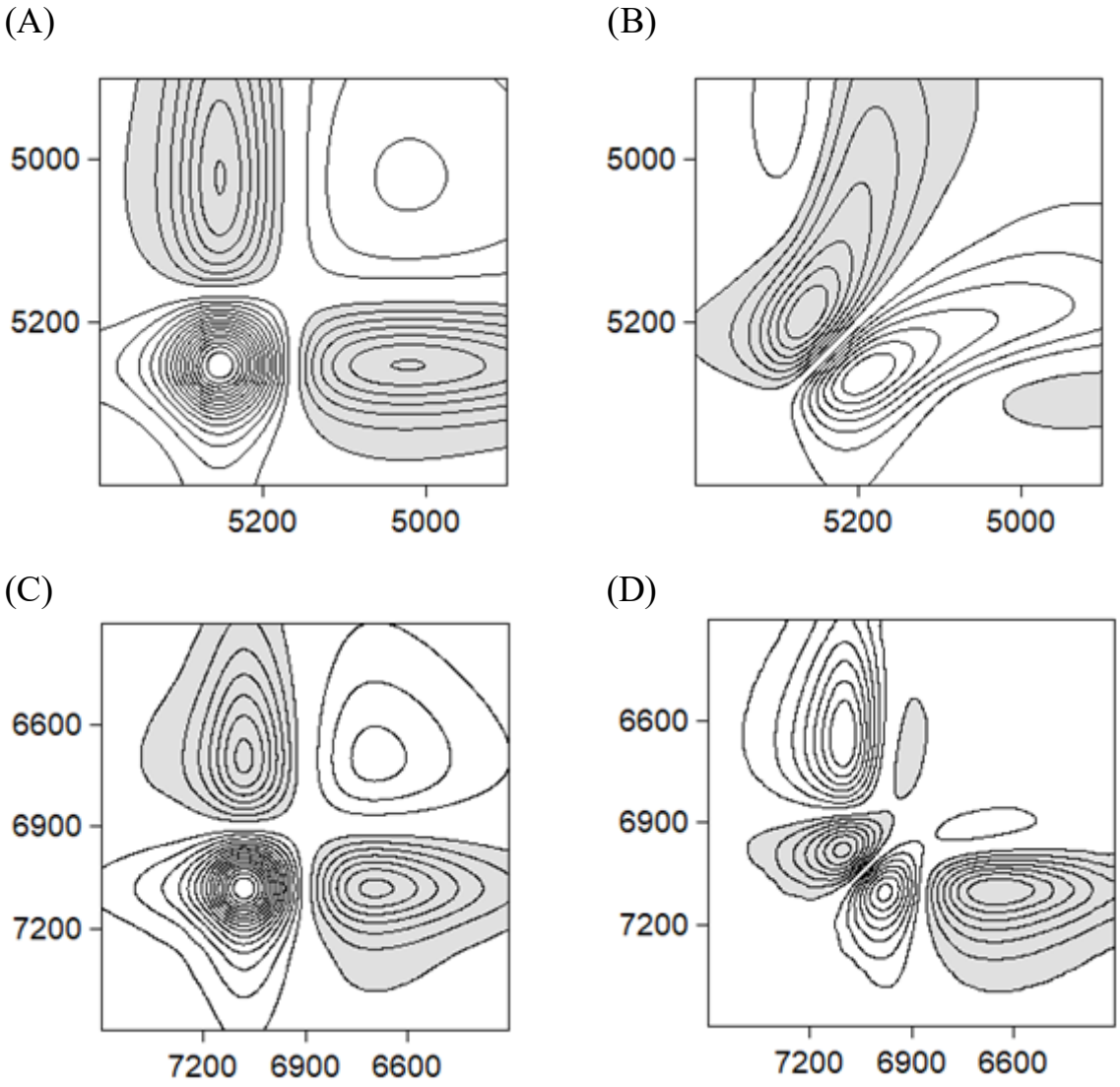


Figure S4: (A) (C) Synchronous and (B) (D) asynchronous correlation spectra calculated for NIR absorbance spectra of ultrapure water in the (A) (B) 5400-4900 and (C) (D) 7500-6300 cm^{-1} regions.

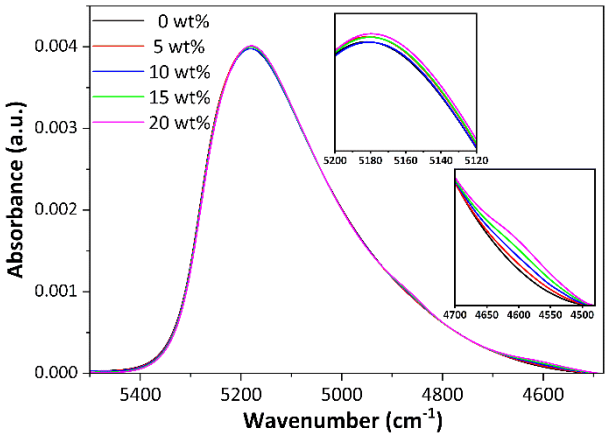


Figure S5: The normalized mean NIR absorbance spectra in the 5500-4480 cm^{-1}. The intensity variations of the bands due to proteins at around 4850 and 4580 cm^{-1} were observed.

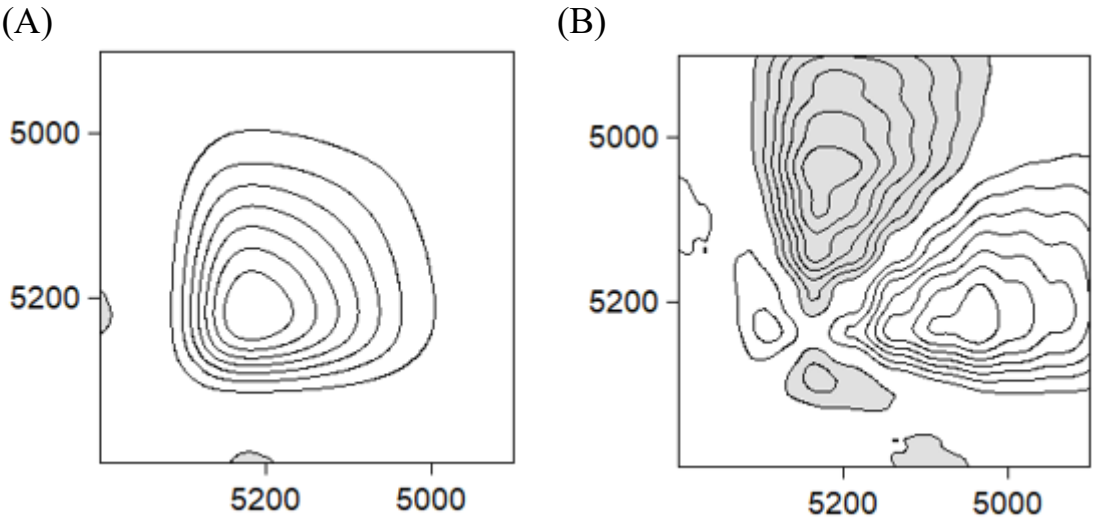


Figure S6: (A) Synchronous and (B) asynchronous correlation spectra in the 5400-4900 cm^{-1} regions calculated for the normalized absorbance spectra recorded every five minutes during the heating for 60 minutes at 70 °C.

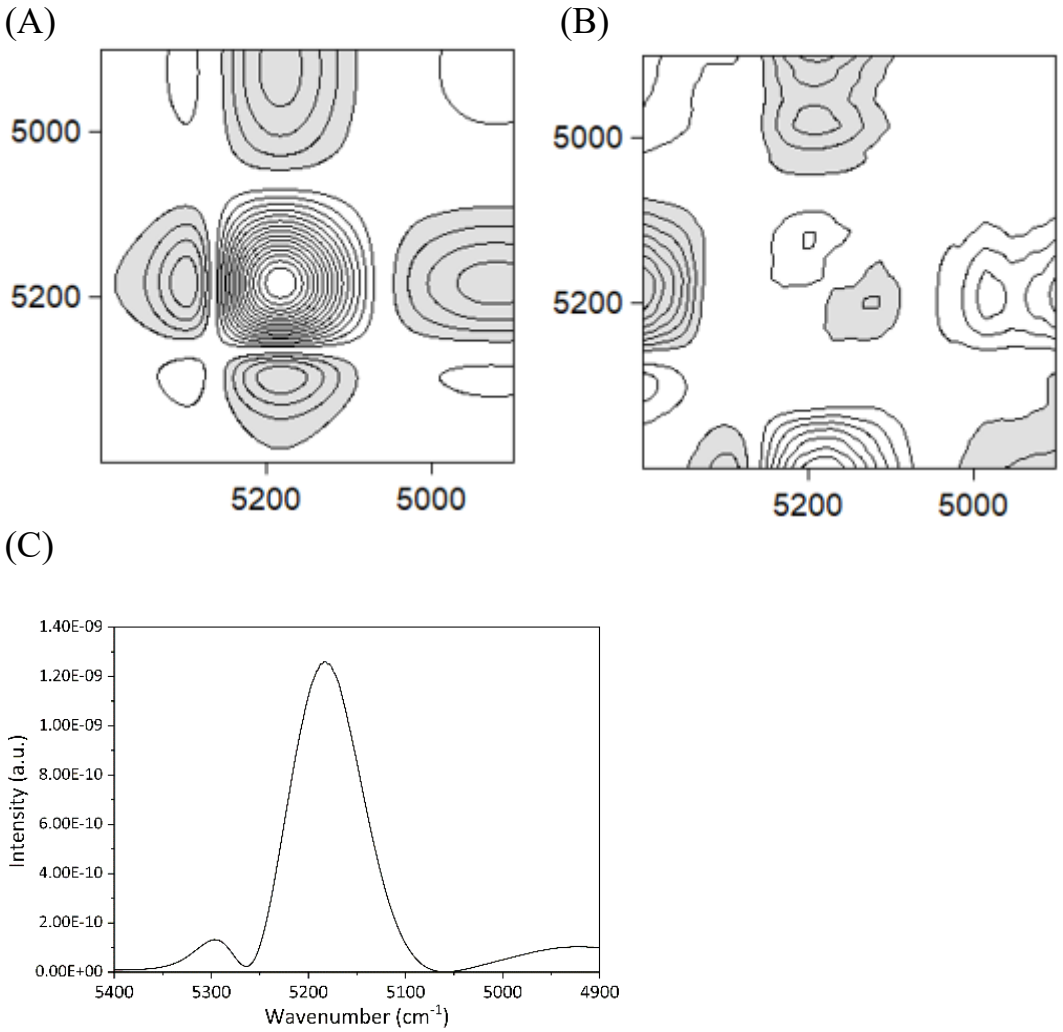


Figure S7: (A) Synchronous and (B) asynchronous correlation spectra in the 5400-4900 cm^{-1} regions calculated for the absorbance spectral data set recorded at 25 °C after heating at 60, 70, 80, and 90 °C for 60 minutes, and the corresponding (C) power spectrum.

Reference

S1. Park, Y., Jin, S., Noda, I., & Jung, Y. M. (2020). Emerging developments in two-dimensional correlation spectroscopy (2D-COS). *Journal of Molecular Structure*, *1217*, 128405.