Biophysics and Physicobiology

Regular Article

Cancellation between auto- and mutual correlation contributions of protein/water dynamics in terahertz time-domain spectra

Yasumasa Joti[1,2] and Akio Kitao[3]

[1]*Japan Synchrotron Radiation Research Institute, Sayo-gun, Hyogo 679-5198, Japan*
[2]*RIKEN SPring-8 Center, Sayo-gun, Hyogo 679-5148, Japan*
[3]*School of Life Sciences and Technology, Tokyo Institute of Technology, Meguro-ku, Tokyo 152-8550, Japan*
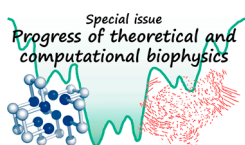
Terahertz time-domain spectra (THz-TDS) were investigated using the results of molecular dynamics (MD) simulations of Staphylococcal nuclease at two hydration states in the temperature range between 100 and 300 K. The temperature dependence of THz-TDS was found to differ significantly from that of the incoherent neutron scattering spectra (INSS) calculated from the same MD simulation results. We further examined contributions of the mutual and auto-correlations of the atomic fluctuations to THz-TDS and found that the negative value of the former contribution nearly canceled out the positive value of the latter, resulting in a monotonic increase of the reduced absorption cross section. Because of this cancellation, no distinct broad peak was observed in the absorption lineshape function of THz-TDS, whereas the protein boson peak was observed in INSS. The contribution of water molecules to THz-TDS was extremely large for the hydrated protein at temperatures above 200 K, in which large-amplitude motions of water were excited. The combination of THz-TDS, INSS and MD simulations

has the potential to extract function-relevant protein dynamics occurring on the picosecond to nanosecond timescale.

Conformational dynamics plays an important role in protein function. Collective motions in proteins slower than pico-seconds in timescale are of particular interest. Incoherent neutron scattering experiments are among the successful methods to study protein dynamics on the picosecond to nanosecond timescale. The protein boson peak is a broad peak found in the low-frequency region (16–32 cm^{-1}, corresponding to 1–2 ps) of the incoherent neutron scattering spectra (INSS) [1–4]. Frequencies of the protein boson peak shift higher upon hydration, indicating that solvent molecules are associated with the origin of the peak [2,4]. As the temperature increases, the boson peak for a hydrated protein shifts to lower frequency and becomes buried in the quasi-elastic contribution [1]. Another temperature-dependent phenomenon with regard to protein dynamics is a dynamical

Corresponding author: Akio Kitao, School of Life Science and Technology, Tokyo Institute of Technology, 2-12-1 Ookayama, Meguro-ku, Tokyo 152-8550, Japan.
e-mail: akitao@bio.titech.ac.jp

◀ *Significance* ▶

The temperature dependence of terahertz time-domain spectra (THz-TDS) is shown to differ significantly from that of incoherent neutron scattering spectra (INSS) using molecular dynamics simulations of Staphylococcal nuclease at two hydration states in the temperature range between 100 and 300 K. The mutual correlation between distinct atomic fluctuations and water dynamics to THz-TDS, which are negligible in INSS, contributes to this difference. This work provides the first theoretical and computational analysis of the dynamical components that contribute to THz-TDS.

transition characterized as an abrupt increase in atomic fluctuations at a temperature above ~200 K [1,5–7]. This increase has been interpreted as a result of a transition from harmonic to diffusive anharmonic motions of the protein. Such diffusive anharmonic motions are crucial to protein function [1,7]. Moreover, the transition is strongly suppressed in dry proteins, indicating that solvent molecules are implicated in the activation of the anharmonic motions [1,4].

INSS can be simulated using the results of molecular dynamics (MD) simulations [8–12]. Our previous studies have revealed that the protein boson peak is observed when protein motions distributed over the entire protein are trapped within narrow local energy minima with higher frequencies [11,12]. This trapping causes a frequency shift in the protein motions from the frequency range lower than the boson peak to higher, resulting in lowering of the spectral density in this range and enhancing the boson peak. Our simulations have also shown that most water molecules in the hydrated state (hydration ratio in weight, $h=0.49$) behave like bulk water above the transition temperature and act as a lubricant for protein dynamics, which excites diffusive anharmonic motion of the protein [13]. In contrast, water molecules in the dry state ($h=0.09$) tend to form more hydrogen bonds with the protein even at temperatures above ~200 K, which restricts the fluctuation of these water molecules to the fluctuation level of the protein. The dynamical transition in the hydrated state is significant in loop and terminal regions exposed to solvent, whereas these regions are restricted to protein-protein contacts in the dry state [13]. Consequently, restricted protein fluctuations in the dry state cause emergence of the protein boson peak even at 300 K [12].

As a result of technical advances in generating and detecting pulsed radiation in the terahertz region, terahertz time-domain spectroscopy has been widely conducted on bimolecular systems [14–21]. Terahertz time-domain spectra (THz-TDS) studies of proteins in solution have shown that water perturbs the protein dynamics via its hydrogen bond network by both direct short-range interactions in the first hydration shell and indirect long-range interactions up to several hydration shells [17–19]. The temperature and hydration dependence of biomolecular dynamics on the picosecond timescale has also been investigated with THz-TDS of biomolecules in the powder states [14–16]. In these studies, the relation between THz-TDS and the vibrational density of states (VDOS) is discussed. The auto-correlation of the atomic fluctuations is involved in VDOS and INSS, while the auto-correlation of the dipole moment of the sample reflects THz-TDS. In this work, we show that if charge distribution is expressed by partial charges assigned on atoms, both mutual correlation between distinct atomic fluctuations and auto-correlation of atomic fluctuations affect the fluctuation of the dipole moment, indicating that protein collective motions in the sample significantly reflect THz-TDS. The contribution of mutual correlation on THz-TDS has not been examined experimentally because its estimation

with only the THz-TDS experiment is difficult. However, contribution analysis can be conducted theoretically based on the results of MD simulations.

Here we show that mutual correlation between atomic fluctuations, as well as auto-correlation, significantly contribute to THz-TDS, whereas the mutual correlation does not contribute to INSS. To calculate THz-TDS, we assume that the dipole moment of the simulation system can be roughly approximated by a classical treatment using the results of all-atom MD simulations of the hydrated and dry proteins at different temperatures ranging from 100 to 300 K. In order to compare the physical characteristics of THz-TDS with those of INSS, the same MD trajectories as in our previous study [12,13] were used. We also show that water dynamics significantly affect the dipole moment of the system in this paper, whereas the contribution of water dynamics is negligible in INSS of proteins in deuterated water [12]. Although quantum treatment is shown to be important for quantitative estimation of THz-TDS in water dynamics [22], we would expect that the significance of the contributions of the mutual correlations hold true even without quantum treatment.

Methods

Classical and atomistic description of terahertz time-domain spectrum

The absorption scattering cross section function, $\alpha(\omega)$, measured by THz-TDS or infrared spectroscopy is defined [23] as:

$$\alpha(\omega) = \frac{4\pi^2}{3\hbar c n(\omega)} \omega(1-e^{-\beta\hbar\omega})I(\omega), \qquad (1)$$

where $\hbar$, c, β, $n(\omega)$ are the reduced Plank constant, the velocity of light, the thermodynamic beta and the refractive index, respectively. The absorption lineshape function, $I(\omega)$, is defined as the Fourier transform of the time-correlation function of the total dipole moment of the sample, $\mathbf{M}(t)$, as:

$$I(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \langle \mathbf{M}(0)\cdot\mathbf{M}(t)\rangle \exp(-i\omega t)dt$$

$$= \lim_{\tau\to\infty} \frac{2\pi}{\tau} |\tilde{\mathbf{M}}(\omega)|^2, \qquad (2)$$

where $\tilde{\mathbf{M}}(\omega)$ is defined as:

$$\tilde{\mathbf{M}}(\omega) = \frac{1}{2\pi} \int_0^{\tau} \mathbf{M}(t)\exp(-i\omega t)dt. \qquad (3)$$

In THz-TDS, the wavelength of the terahertz field pulse is ~0.1 mm, which is much larger than the length of molecular orbitals. The small change in the electron-density distribution originating from quantum effects might not significantly affect the computation of the total dipole moment. However, we expect that the overall features of THz-TDS can be well reproduced by this treatment. Therefore, we computed the dipole moment in the simulation using the partial charge

defined by the AMBER *ff99* force field [24] and TIP3P water model [25] as:

$$\mathbf{M}(t) = \sum_j q_j \mathbf{r}_j(t), \tag{4}$$

where q_j and $\mathbf{r}_j(t)$ are the partial charge and positional vector for the *j*-th atom, respectively. Although electronic polarizability in molecules affects the picosecond dynamics of water [26], this contribution was not considered in this paper. We assume that the significance of the contributions of the mutual and auto-correlations of the atomic fluctuations and water to THz-TDS would not change, although the contribution of electronic polarizability might quantitatively change the spectra. Technical treatment of the periodic boundary condition for the system is discussed in Supplementary Text S1.

The absorption lineshape function of THz-TDS, $I(\omega)$, can be computed from the results of molecular dynamics simulations. It is noted that the energy level of kT at 100 K corresponds to a frequency of ~70 cm^{-1}. Therefore, quantum effects concerning the ratio $\hbar\omega/kT$ are not dominant in protein dynamics above 100 K in the terahertz-frequency regions (~30 cm^{-1}). To eliminate the thermal factor, the reduced absorption cross section (RACS), $\alpha_R(\omega)$, is introduced [15,27] as:

$$\alpha_R(\omega) = \frac{\beta\hbar c\omega}{1 - e^{\beta\hbar\omega}} n(\omega)\alpha(\omega)$$

$$= \frac{4\pi^2}{3}\beta\omega^2 I(\omega), \tag{5}$$

which indicates that RACS is proportional to $\beta\omega^2 I(\omega)$. Hereafter, we focus on the behavior of $\beta\omega^2 I(\omega)$ for biomolecular samples.

Relation between vibrational density of states and THz-TDS

In this subsection, we examine the relation between the vibrational density of states (VDOS) and THz-TDS. The VDOS, $g(\omega)$, for the specimen is expressed using the power spectrum of the mass weighted velocity vector, $\dot{\mathbf{r}}_j(t)$, as shown in [28]:

$$g(\omega) = \beta \sum_j m_j \lim_{\tau\to\infty} \frac{1}{2\pi\tau} \left| \int_0^\tau dt \exp(-i\omega t)\dot{\mathbf{r}}_j(t) \right|^2, \tag{6}$$

where m_j is the mass of the *j*-th atom. The power spectrum of the velocity vector is rewritten using the power spectrum of the instantaneous deviation of the position vector from its average position:

$$\lim_{\tau\to\infty} \frac{1}{2\pi\tau} \left| \int_0^\tau dt \exp(-i\omega t)\dot{\mathbf{r}}_j(t) \right|^2$$

$$= \omega^2 \lim_{\tau\to\infty} \frac{1}{2\pi\tau} \left| \int_0^\tau \Delta\mathbf{r}_j(t)\exp(-i\omega t)dt \right|^2, \tag{7}$$

where $\mathbf{r}_j(t) = \langle\mathbf{r}_j(0)\rangle + \Delta\mathbf{r}_j(t)$. Here, we define $\chi_j(\omega)$ as the Fourier transform of the time-correlation function of $\Delta\mathbf{r}_j(t)$, i.e., the power spectrum of $\Delta\mathbf{r}_j(t)$:

$$\chi_j(\omega) = \frac{1}{2\pi} \int_{-\infty}^\infty \langle\Delta\mathbf{r}_j(0)\cdot\Delta\mathbf{r}_j(t)\rangle \exp(-i\omega t)dt$$

$$= \lim_{\tau\to\infty} \frac{1}{2\pi\tau} \left| \int_0^\tau \Delta\mathbf{r}_j(t)\exp(-i\omega t)dt \right|^2. \tag{8}$$

When the dynamics of the system is harmonic, $\chi_j(\omega)$ is proportional to kT because of the equipartition of energy. Using Eqs. (6) and (8), $g(\omega)$ is expressed as:

$$g(\omega) = \beta\omega^2 \sum_j m_j \chi_j(\omega). \tag{9}$$

Next, we formulate $\beta\omega^2 I(\omega)$ using $\chi_j(\omega)$. When we treat the dipole moment classically as in Eq. (4), the time-correlation function of $\mathbf{M}(t)$ is rewritten as:

$$\langle\mathbf{M}(0)\cdot\mathbf{M}(t)\rangle = \left\langle \sum_j q_j\mathbf{r}_j(0)\cdot\sum_k q_k\mathbf{r}_k(t) \right\rangle$$

$$= |\langle\mathbf{M}(0)\rangle|^2 + \sum_j q_j^2 \langle\Delta\mathbf{r}_j(0)\cdot\Delta\mathbf{r}_j(t)\rangle$$

$$+ \sum_{j\neq k} q_j q_k \langle\Delta\mathbf{r}_j(0)\cdot\Delta\mathbf{r}_k(t)\rangle. \tag{10}$$

A Fourier transform of Eq. (10) yields the absorption lineshape function from THz-TDS,

$$I(\omega) = |\langle\mathbf{M}(0)\rangle|^2 \delta(\omega) + A(\omega) + C(\omega), \tag{11}$$

where $A(\omega) = \sum_j q_j^2 \chi_j(\omega)$ and,

$$C(\omega) = \frac{1}{2\pi} \int_{-\infty}^\infty dt \exp(-i\omega t)\sum_{j\neq k} q_j q_k \langle\mathbf{r}_j(0)\cdot\mathbf{r}_k(t)\rangle. \tag{12}$$

$A(\omega)$ and $C(\omega)$ originate from the auto-correlation of the atomic fluctuation and the mutual correlation (or cross-correlation) of the atomic fluctuations, respectively. It should be noted that $C(\omega)$ can adopt a negative value, though $A(\omega)$ is always positive. Although Eq. (11) is derived based on the atomic partial charge model, a similar formulation is also possible by introducing local dipoles around atoms. From Eq. (11), we can derive the following equation:

$$\beta\omega^2 I(\omega) = \beta\omega^2 A(\omega) + \beta\omega^2 C(\omega). \tag{13}$$

If the absolute value of $C(\omega)$ is negligibly small, the density of states can be estimated from THz-TDS, although the weighting factors for each atom *j* differ between density of state (m_j) and THz-TDS (q_j^2). However, $C(\omega)$ is not negligible as will be shown in the Results and Discussion section. The contribution of the mutual correlation between atomic fluctuations (or $\beta\omega^2 C(\omega)$) to RACS (or $\beta\omega^2 I(\omega)$) has not been evaluated because it is impossible to measure it directly from the experiment. In the Results and Discussion section, we examine the contribution of $\beta\omega^2 C(\omega)$ to $\beta\omega^2 I(\omega)$ using

the results of molecular dynamics simulations of crystalline proteins.

Contributions of protein dynamics, solvent dynamics and their coupling to $\beta\omega^2 I(\omega)$

As described in the Introduction, solvent dynamics play an important role in protein function. In order to clarify the contribution of the solvent to THz-TDS, here we decompose the total dipole moment of the sample, $\mathbf{M}(t)$, into two terms, the dipole moment of the protein molecule and that of the solvent:

$$\mathbf{M}(t) = \mathbf{M}_{pro}(t) + \mathbf{M}_{sol}(t)$$
$$= \sum_{j \in protein} q_j \mathbf{r}_j(t) + \sum_{j \in solvent} q_j \mathbf{r}_j(t). \quad (14)$$

The time-correlation function of $\mathbf{M}(t)$ can be decomposed into three terms:

$$\langle \mathbf{M}(0) \cdot \mathbf{M}(t) \rangle = \langle (\mathbf{M}_{pro}(0) + \mathbf{M}_{sol}(0)) \cdot (\mathbf{M}_{pro}(t) + \mathbf{M}_{sol}(t)) \rangle$$
$$= \langle \mathbf{M}_{pro}(0) \cdot \mathbf{M}_{pro}(t) \rangle + \langle \mathbf{M}_{sol}(0) \cdot \mathbf{M}_{sol}(t) \rangle$$
$$+ \{ \langle \mathbf{M}_{pro}(0) \cdot \mathbf{M}_{sol}(t) \rangle + \langle \mathbf{M}_{sol}(0) \cdot \mathbf{M}_{pro}(t) \rangle \}. \quad (15)$$

Eq. (15) shows that the time-correlation function of $\mathbf{M}(t)$ is rewritten as the sum of the contributions of protein, solvent, and coupling between protein and solvent dynamics. It should also be noted that this decomposition is possible without assuming the atomic partial charge model as shown in the first line of Eq. (14). By performing a Fourier transform of Eq. (15), $\beta\omega^2 I(\omega)$ is decomposed into three parts as:

$$\beta\omega^2 I(\omega) = \beta\omega^2 I_{pro}(\omega) + \beta\omega^2 I_{sol}(\omega) + \beta\omega^2 I_{coupling}(\omega), \quad (16)$$

where $I_{pro}(\omega)$, $I_{sol}(\omega)$, and $I_{coupling}(\omega)$ are the Fourier transforms of the first, second, and third terms in Eq. (15), respectively. In the Results and Discussion section, we examine the contributions of protein dynamics, solvent dynamics and their coupling to $\beta\omega^2 I(\omega)$.

Incoherent neutron scattering spectra

Here, we summarize the formulation of INSS, $S_{inc}(\mathbf{Q},\omega)$, for comparison with THz-TDS (see Supplementary Text S2 for details) where $\mathbf{Q}$ and ω represent the momentum and energy transfers between incident neutron and sample, respectively. $S_{inc}(\mathbf{Q},\omega)$ is defined as the weighted sum of the Fourier transform of the time-correlation functions:

$$S_{inc}(\mathbf{Q},\omega) = \frac{1}{2\pi} \sum_j b_{inc,j}^2 \int_{-\infty}^{+\infty} dt \exp(-i\omega t)$$
$$\times \langle \exp(-i\mathbf{Q} \cdot \Delta\mathbf{r}_j(0)) \exp(i\mathbf{Q} \cdot \Delta\mathbf{r}_j(t)) \rangle \quad (17)$$

where $\langle \exp(-i\mathbf{Q} \cdot \Delta\mathbf{r}_j(0)) \exp(i\mathbf{Q} \cdot \Delta\mathbf{r}_j(t)) \rangle$ indicates the time-

correlation function, and $b_{inc,j}$ and $\Delta\mathbf{r}_j(t)$ are the incoherent atomic scattering length and the instantaneous deviation of the position vector of the j-th atom from its average position at time t, respectively. This equation shows that INSS only consists of the contribution of the atomic auto-correlations, whereas mutual correlations reflect coherent neutron scattering, which is not considered in this work. On the sphere of $|\mathbf{Q}|=Q$, $g(\omega)$, $S_{inc}(Q,\omega)$ and $\chi_j(\omega)$ have the following relation:

$$g(\omega) \propto \lim_{Q \to 0} \frac{\omega^2 S_{inc}(Q,\omega)}{Q^2} = \beta\omega^2 \sum_j b_{inc,j}^2 \chi_j(\omega). \quad (18)$$

Computational procedure

Typically, powder samples are employed to investigate the temperature and hydration dependence of THz-TDS of biomolecules [14–16]. Here, we regarded the powder state as an ensemble of micro-crystals and employed molecular dynamics (MD) simulation trajectories of a crystalline Staphylococcal nuclease (SNase) at six temperatures ranging from 100 to 300 K [12,13] to calculate THz-TDS. In order to examine the solvent effect on protein dynamics, two types of biomolecular states, minimal- and full-hydration states (hereafter, MHS and FHS, respectively), were simulated. The hydration levels for MHS and FHS are $h=0.09$ and $h=0.49\,g$ D_2O/g protein, respectively. The simulations were described in detail in our previous paper [13]. From the results of the simulations, dynamical transitions were observed at around 200 K in both MHS and FHS, which were more significant in FHS. The characteristics of water in FHS were similar to those for bulk water above the transition temperature, which enhances protein dynamics. The protein boson peak was observed in the frequency range 16–32 cm^{-1} at 100 K in both states. The peak frequency in FHS shifted to higher than that in MHS. At 300 K, the boson peak was buried in the quasi-elastic contributions originating from diffusive anharmonic motions in FHS, but was still visible in MHS.

The MD simulations at each temperature generated a 10-ns trajectory. Each 10-ns trajectory was divided into five 2-ns trajectories, and physical quantities such as $I(\omega)$ were calculated as the average of the results from the five 2-ns trajectories. THz-TDS presented here were smoothed by convolution with a resolution of a 0.323 cm^{-1}, which corresponded to 0.1 ns, while INSS were smoothed by convoluting a 1.61 cm^{-1} (~200 μeV) resolution function, which corresponded to the LAM40 spectrometer [7].

Results and Discussion

Figure 1 shows the temperature dependence of the absorption lineshape function, $I(\omega)$, and INSS, $S(Q,\omega)$ at $Q=2$ Å^{-1}, around the THz frequency (33 cm^{-1}) in MHS and FHS (See Supplementary Figure S1 for higher resolution spectra). The temperature dependence of $I(\omega)$ was significantly different from that of $S(Q,\omega)$. No distinct broad peak was observed in
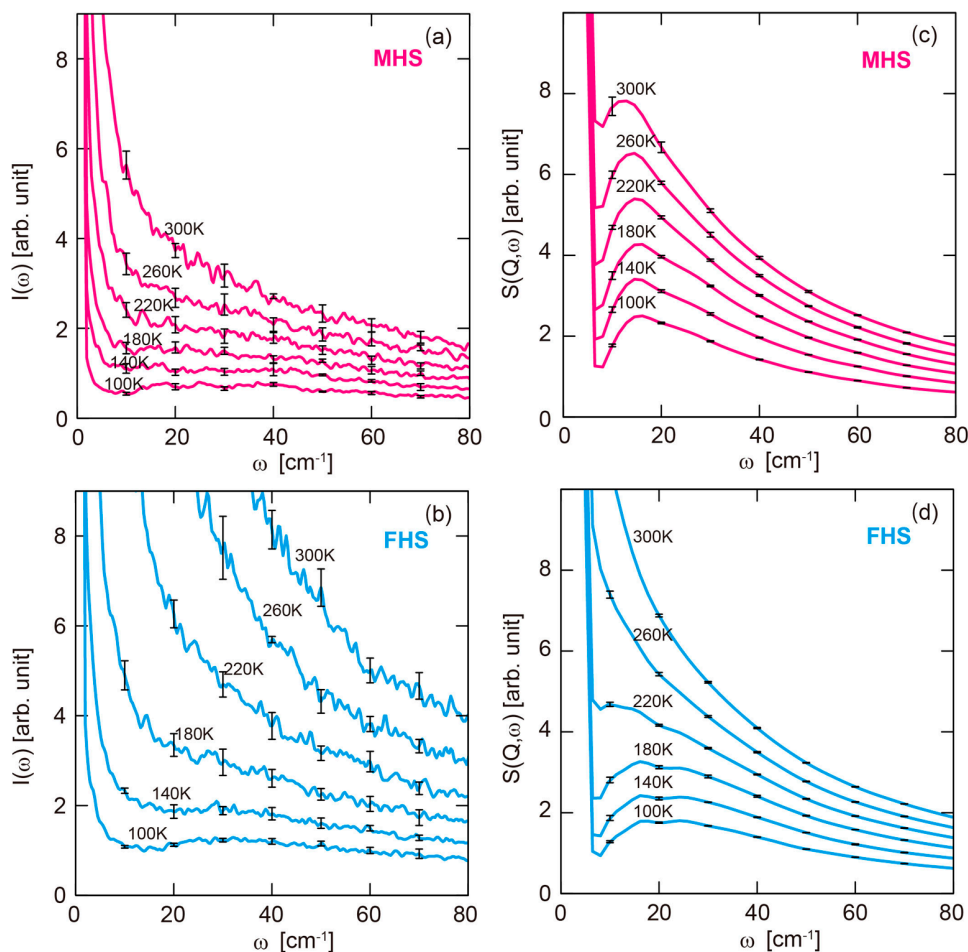
Figure 1 Temperature dependence of the absorption lineshape function, $I(\omega)$, for (a) MHS and (b) FHS, and that of the incoherent neutron scattering spectra (INSS), $S(Q,\omega)$, for (c) MHS and (d) FHS at $Q=2$ Å^{-1}. Error bars show standard deviations.

$I(\omega)$ at any temperature in either FHS and MHS, while the protein boson peak was observed in INSS except for FHS at 260 and 300 K. The sum of $\chi_j(\omega)$ weighted by the atomic scattering length was previously shown to have a peak in the frequency range of 16–32 cm^{-1} except for FHS at 260 and 300 K [12]. Considering this fact and Eq. (11), we expected that the contribution of $C(\omega)$ concealed other components in $I(\omega)$ having a peak in the frequency range 16–32 cm^{-1}, which we will show later. As seen in Figure 2, the temperature-normalized spectra, $\beta I(\omega)$, in the frequency range $\omega>40$ cm^{-1} in MHS were in good agreement at all temperatures, while those in FHS differed particularly at high temperatures. Water molecules in MHS are mainly hydrogen-bonded with the protein, resulting in restricted water movement as compared to the fluctuation of the protein [13]. The loop and terminal regions, which are exposed to solvent in FHS, are restricted in MHS by protein-protein contacts [13]. Because of these restrictions, dynamics in MHS can be considered as almost harmonic. In contrast, most water molecules in FHS behave like bulk water and act as a lubricant for protein dynamics at higher temperatures [12,13], which causes the difference in the temperature dependence of $I(\omega)$ between MHS and FHS.

Figure 3 shows $\beta\omega^2 I(\omega)$ and $g(\omega)$ at 100 and 300 K for MHS and FHS. In FHS, $\beta\omega^2 I(\omega)$ and $g(\omega)$ at 300 K were notably greater than those at 100 K. As shown in our previous paper [12], $g(\omega)$ for protein only in FHS at 300 K was similar to that at 100 K, indicating that the large difference in total $g(\omega)$ between 100 K and 300 K originated from $g(\omega)$ for water. The origin of the large temperature difference in $\beta\omega^2 I(\omega)$ is examined later. On the other hand, in MHS, $\beta\omega^2 I(\omega)$ and $g(\omega)$ at 300 K were only slightly greater than those at 100 K, due to restricted protein fluctuations in MHS [12,13]. In all the hydration and temperature conditions, $\beta\omega^2 I(\omega)$ increased monotonically in the frequency range $\omega<80$ cm^{-1}, whereas $g(\omega)$ became flat in the frequency range $\omega>40$ cm^{-1}. From the comparison between Eqs. (9) and (13), the contribution of $C(\omega)$ was concluded to be the origin of the difference.

Figure 4 shows contributions of the auto-correlation (the first term in Eq. (13)) and the mutual correlation (the second term in Eq. (13)) to total $\beta\omega^2 I(\omega)$ at 100 and 300 K in MHS
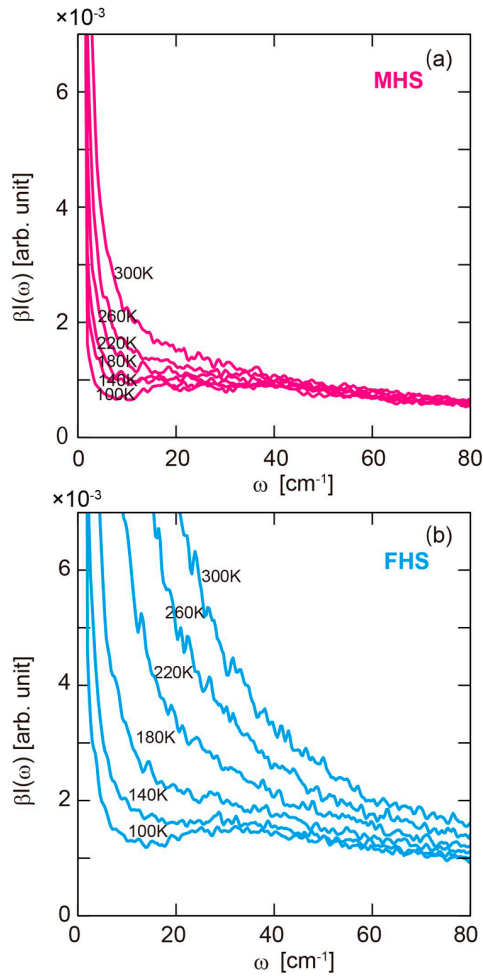
Figure 2 Temperature dependence of the absorption lineshape function normalized with temperature, $\beta I(\omega)$, for (a) MHS and (b) FHS.

Figure 3 (a) $\beta\omega^2 I(\omega)$ at 100 and 300 K for MHS (magenta) and FHS (cyan). (b) Density of states, $g(\omega)$, at 100 and 300 K for MHS (magenta) and FHS (cyan). Error bars show standard deviations.

and FHS. Similar to the ω dependence of $g(\omega)$, the auto-correlation term, $\beta\omega^2 A(\omega)$, becomes almost flat above 40 cm^{-1} in all cases. In contrast, while the ω dependence of the mutual correlation term, $\beta\omega^2 C(\omega)$, had a shape similar to that for the auto-correlation term, the sign was negative, almost a mirror image of the auto-correlation with respect to the abscissa. The negative value of $\beta\omega^2 C(\omega)$ almost canceled the positive value of $\beta\omega^2 A(\omega)$, which resulted in a positive value of the total function $\beta\omega^2 I(\omega)$ and a monotonic increase of RACS. In other words, the physical nature of RACS is significantly different from VDOS because $C(\omega)$ is not negligibly small. If the motion of the j-th atom is independent of that of the k-th atom, $\langle \Delta\mathbf{r}_j(0)\cdot\Delta\mathbf{r}_k(t) \rangle$ in Eq. (10) becomes zero. Our results indicate that the collective motions of the system appear in $C(\omega)$. As theoretically shown above, THz-TDS includes information on both auto- and mutual correlations between atomic fluctuations, whereas INSS originates only from the auto-correlations.

Finally, we examined the contribution of the solvent to THz-TDS. As seen in Figure 1, the intensity of $I(\omega)$ for MHS
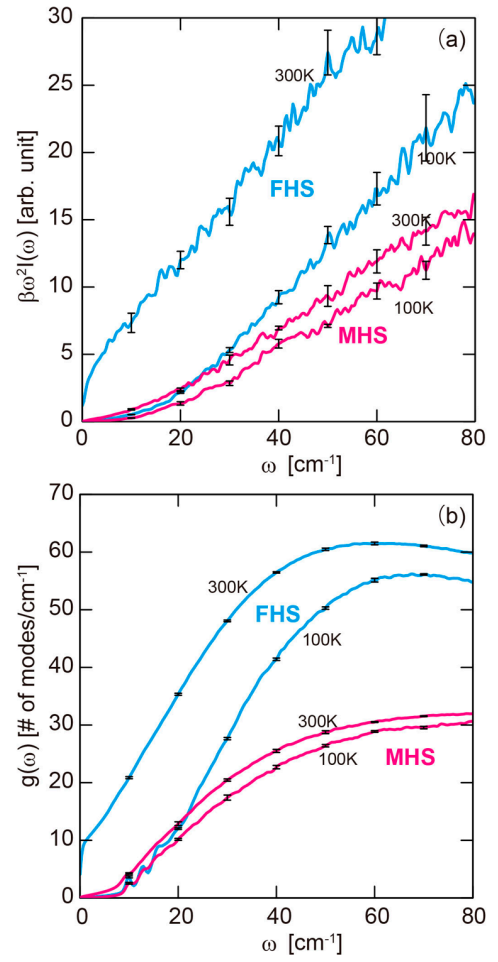
differed entirely from that for FHS at all temperatures. On the other hand, INSS for MHS and FHS were in good agreement above 40 cm^{-1} at all temperatures [12], because the contribution of deuterated water to INSS was negligibly small in the higher frequency region. Figure 5 shows the contributions of protein dynamics, water dynamics and protein-water coupling to $\beta\omega^2 I(\omega)$ at 100 and 300 K for MHS and FHS. $\beta\omega^2 I_{\text{pro}}(\omega)$ in MHS and FHS was comparable at both temperatures. $\beta\omega^2 I_{\text{sol}}(\omega)$ in MHS was lower than that in FHS at both temperatures, and that in FHS was comparable to $\beta\omega^2 I_{\text{pro}}(\omega)$ at 100 K and became almost double that of $\beta\omega^2 I_{\text{pro}}(\omega)$ at 300 K, which indicated that solvent dynamics in FHS dominantly contributed to $\beta\omega^2 I(\omega)$ at 300 K. In all cases, $\beta\omega^2 I_{\text{coupling}}(\omega)$ was small and negative. Because $\beta\omega^2 I_{\text{sol}}(\omega)$ in MHS can nearly cancel out $\beta\omega^2 I_{\text{coupling}}(\omega)$ at both temperatures, $\beta\omega^2 I_{\text{pro}}(\omega)$ and $\beta\omega^2 I(\omega)$ look similar. We should note that the contribution of the solvent to $\beta\omega^2 I(\omega)$ and RACS should be extremely large for a fully hydrated protein. In order to extract information on protein dynamics from THz-TDS of the hydrated proteins, its combination
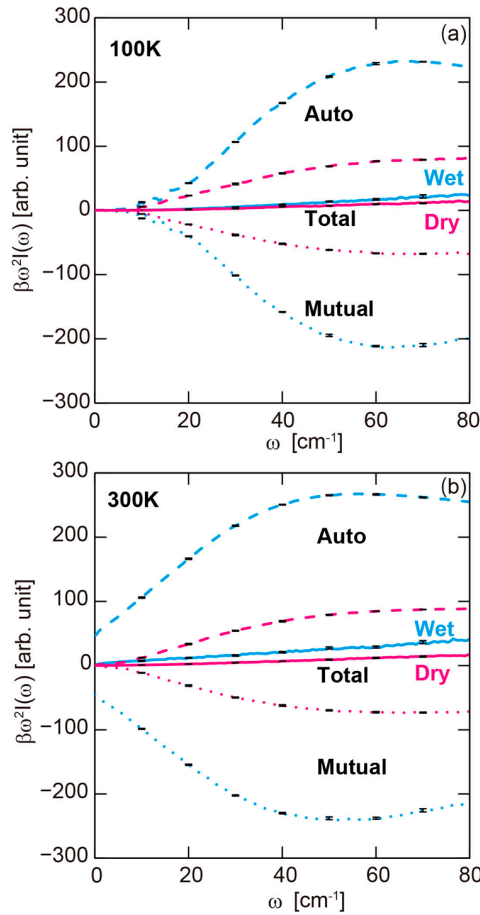
Figure 4 Auto-correlation (broken line) and mutual correlation (dotted line) contributions to total $\beta\omega^2 I(\omega)$ (solid line) at (a) 100 and (b) 300 K for MHS (magenta) and FHS (cyan). This decomposition is defined in Eq. (13). Error bars show standard deviations.

Figure 5 Contributions of $\beta\omega^2 I_{\mathrm{pro}}(\omega)$ (thin solid line), $\beta\omega^2 I_{\mathrm{sol}}(\omega)$ (broken line) and $\beta\omega^2 I_{\mathrm{coupling}}(\omega)$ (dotted line) to $\beta\omega^2 I(\omega)$ (thick solid line) at (a) 100 and (b) 300 K for MHS (magenta) and FHS (cyan). This decomposition is defined in Eq. (16). Error bars show standard deviations.

with other methods such as INSS and MD simulation is essential.

Conclusion

We calculated THz-TDS of SNase in different hydration states using the results of MD simulations conducted at different temperatures. The temperature dependence of the absorption lineshape function was very different from that of INSS because THz-TDS significantly depended on both auto- and mutual correlations between atomic fluctuations. Since the decomposition of THz-TDS into auto- and mutual correlations is experimentally impossible, this work provides the first theoretical and computational analysis of motional components that contribute to THz-TDS. Since the contribution of the mutual correlation of the atomic fluctuations is not negligibly small, RACS is significantly different from VDOS, in contrast to INSS which depends only on the auto-correlation of the atomic fluctuations. Our results clearly demonstrated that the mutual correlation term has a negative
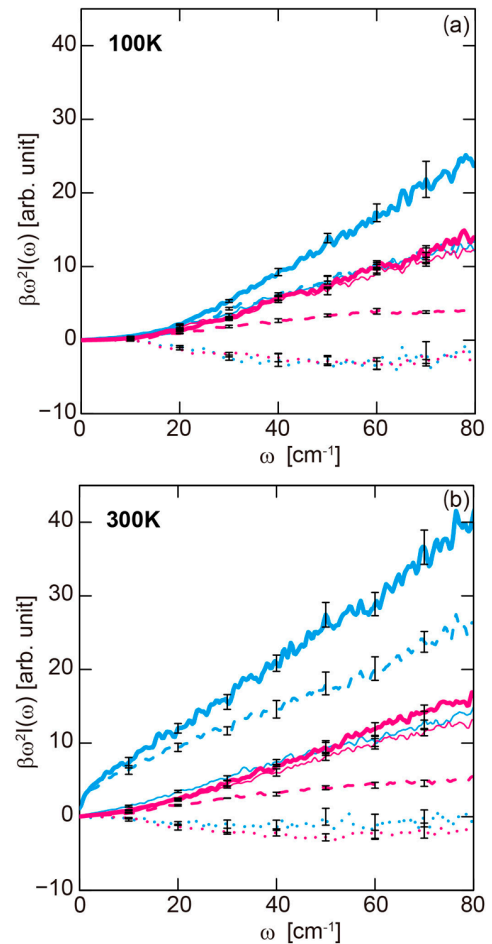
intensity comparable to the auto-correlation term in absolute value. Therefore, the auto- and mutual correlations nearly cancel each other out, and the resulting positive residue can be observed as a monotonically increasing feature of RACS. Because of this cancellation, no distinct broad peak is observed in the absorption lineshape function, whereas the protein boson peak is observed in INSS except for FHS at 260 and 300 K. The behavior of water molecules in FHS is similar to that of bulk water above ~200 K [13], which enhances protein dynamics and results in the large contribution of water molecules to THz-TDS. Therefore, the contribution of water dynamics to THz-TDS at temperatures above 200 K is extremely large in FHS.

As shown above, THz-TDS can be decomposed into many contributions, such as the auto- and mutual correlations of the atomic fluctuations and solvent effects on the picosecond to nanosecond timescale, which are related to protein function. Although it is difficult to extract such information from THz-TDS only, the combination of THz-TDS, INSS and

molecular simulations has the potential to allow the analysis of function-relevant protein dynamics.

Acknowledgements

Conflict of Interest

Y. J. and A. K. declare that they have no conflict of interest.

Author contribution

Y. J. and A. K. directed the entire project and co-wrote the manuscript. Y. J. carried out MD simulations and analyses.

References

[1] Ferrand, M., Dianoux, A. J., Petry, W. & Zaccai, G. Thermal motions and function of bacteriorhodopsin in purple membranes: effects of temperature and hydration studied by neutron scattering. *Proc. Natl. Acad. Sci. USA* **90**, 9668–9672 (1993).

[2] Diehl, M., Doster, W., Petry, W. & Schober, H. Water-coupled low-frequency modes of myoglobin and lysozyme observed by inelastic neutron scattering. *Biophys. J.* **73**, 2726–2732 (1997).

[3] Cusack, S. & Doster, W. Temperature dependence of the low frequency dynamics of myoglobin. Measurement of the vibrational frequency distribution by inelastic neutron scattering. *Biophys. J.* **58**, 243–251 (1990).

[4] Nakagawa, H., Joti, Y., Kitao, A. & Kataoka, M. Hydration affects both harmonic and anharmonic nature of protein dynamics. *Biophys. J.* **95**, 2916–2923 (2008).

[5] Doster, W., Cusack, S. & Petry, W. Dynamical transition of myoglobin revealed by inelastic neutron scattering. *Nature* **337**, 754–756 (1989).

[6] Zaccai, G. How soft is a protein? A protein dynamics force constant measured by neutron scattering. *Science* **288**, 1604–1607 (2000).

[7] Rasmussen, B. F., Stock, A. M., Ringe, D. & Petsko, G. A. Crystalline ribonuclease A looses function below the dynamical transition at 220 k. *Nature* **357**, 423–424 (1992).

[8] Tarek, M. & Tobias, D. J. The dynamics of protein hydration water: a quantitative comparison of molecular dynamics simulations and neutron-scattering experiments. *Biophys. J.* **79**, 3244–3257 (2000).

[9] Paciaroni, A., Bizzarri, A. R. & Cannistraro, S. Molecular-dynamics simulation evidences of a boson peak in protein hydration water. *Phys. Rev. E* **57**, R6277–6280 (1998).

[10] Kurkal-Siebert, V. & Smith, J. C. Low-temperature protein dynamics: A simulation analysis of interprotein vibrations and the boson peak at 150 K. *J. Am. Chem. Soc.* **128**, 2356–2364 (2006).

[11] Joti, Y., Kitao, A. & Go, N. Protein boson peak originated from hydration-related multiple minima energy landscape. *J. Am. Chem. Soc.* **127**, 8705–8709 (2005).

[12] Joti, Y., Nakagawa, H., Kataoka, M. & Kitao, A. Hydration effect on low-frequency protein dynamics observed in simulated neutron scattering spectra. *Biophys. J.* **94**, 4435–4443 (2008).

[13] Joti, Y., Nakagawa, H., Kataoka, M. & Kitao, A. Hydration-dependent protein dynamics revealed by molecular dynamics simulation of crystalline staphylococcal nuclease. *J. Phys. Chem. B* **112**, 3522–3528 (2008).

[14] Yamamoto, N., Kambara, O., Yamamoto, K., Tamura, A., Saito, S., Tominaga, K. Temperature and hydration dependence of low-frequency spectra of poly-L-glutamic acid with different secondary structures studied by terahertz time-domain spectroscopy. *Soft Matter* **8**, 1997–2006 (2012).

[15] Kawaguchi, S., Kambara, O., Shibata, M., Kandori, H. & Tominaga, K. Low-frequency dynamics of bacteriorhodopsin studied by terahertz time-domain spectroscopy. *Phys. Chem. Chem. Phys.* **12**, 10255–10262 (2010).

[16] Yamamoto, N., Andachi, T., Tamura, A. & Tominaga, K. Temperature and Hydration Dependence of Low-Frequency Spectra of Lipid Bilayers Studied by Terahertz Time-Domain Spectroscopy. *J. Phys. Chem. B* **119**, 9359–9368 (2015).

[17] Bellissent-Funel, M. C., Hassanali, A., Havenith, M., Henchman, R., Pohl, P., Sterpone, F., *et al.* Water Determines the Structure and Dynamics of Proteins. *Chem. Rev.* **116**, 7673–7697 (2016).

[18] Nibali, V. C. & Havenith, M. New Insights into the Role of Water in Biological Function: Studying Solvated Biomolecules Using Terahertz Absorption Spectroscopy in Conjunction with Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **136**, 12800–12807 (2014).

[19] Xu, Y. & Havenith, M. Perspective: Watching low-frequency vibrations of water in biomolecular recognition by THz spectroscopy. *J. Chem. Phys.* **143**, 170901 (2015).

[20] Aoki, K., Shiraki, K. & Hattori, T. Salt effects on the picosecond dynamics of lysozyme hydration water investigated by terahertz time-domain spectroscopy and an insight into the Hofmeister series for protein stability and solubility. *Phys. Chem. Chem. Phys.* **18**, 15060–15069 (2016).

[21] Samanta, N., Luong, T. Q., Das Mahanta, D., Mitra, R. K. & Havenith, M. Effect of Short Chain Poly(ethylene glycol)s on the Hydration Structure and Dynamics around Human Serum Albumin. *Langmuir* **32**, 831–837 (2016).

[22] Sun, J., Niehues, G., Forbert, H., Decka, D., Schwaab, G., Marx, D., *et al.* Understanding THz Spectra of Aqueous Solutions: Glycine in Light and Heavy Water. *J. Am. Chem. Soc.* **136**, 5031–5038 (2014).

[23] McQuarrie, D. A. *Statistical Mechanics*; University Science Books, 2000.

[24] Wang, J., Cieplak, P. & Kollman, P. A. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.* **21**, 1049–1074 (2000).

[25] Jogensen, W. L., Chandrasekhar, J. & Madura, J. D. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).

[26] Harder, E., Eaves, J. D., Tokmakoff, A. & Berne, B. J. Polarizable molecules in the vibrational spectroscopy of water. *Proc. Natl. Acad. Sci. USA* **102**, 11611–11616 (2005).

[27] Yamamoto, K., Tominaga, K., Sasakawa, H., Tamura, A., Murakami, H., Ohtake, H., *et al.* Terahertz time-domain spectroscopy of amino acids and polypeptides. *Biophys. J.* **89**, L22–24 (2005).

[28] Hayward, S., Kitao, A., Hirata, F. & Go, N. Effect of solvent on collective motions in globular protein. *J. Mol. Biol.* **234**, 1207–1217 (1993).