Hydration Effect on Low-Frequency Protein Dynamics Observed in Simulated Neutron Scattering Spectra

Yasumasa Joti,*† Hiroshi Nakagawa,‡ Mikio Kataoka,‡§ and Akio Kitao*†

*Institute of Molecular and Cellular Biosciences, University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-0032, Japan; †Japan Science and Technology Agency, Core Research for Evolution Science and Technology, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-0032, Japan; ‡Neutron Biophysics Group, Neutron Biology Research Center, Quantum Beam Science Directorate, Japan Atomic Energy Agency, Tokai, Ibaraki 319-1195, Japan; and §Graduate School of Materials Science, Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara 630-0192, Japan

ABSTRACT Hydration effects on protein dynamics were investigated by comparing the frequency dependence of the calculated neutron scattering spectra between full and minimal hydration states at temperatures between 100 and 300 K. The protein boson peak is observed in the frequency range 1–4 meV at 100 K in both states. The peak frequency in the minimal hydration state shifts to lower than that in the full hydration state. Protein motions with a frequency higher than 4 meV were shown to undergo almost harmonic motion in both states at all temperatures simulated, whereas those with a frequency lower than 1 meV dominate the total fluctuations above 220 K and contribute to the origin of the glass-like transition. At 300 K, the boson peak becomes buried in the quasielastic contributions in the full hydration state but is still observed in the minimal hydration state. The boson peak is observed when protein dynamics are trapped within a local minimum of its energy surface. Protein motions, which contribute to the boson peak, are distributed throughout the whole protein. The fine structure of the dynamics structure factor is expected to be detected by the experiment if a high resolution instrument ($<\sim$20 μeV) is developed in the near future.

INTRODUCTION

The boson peak is a broad peak found in the low frequency region (1–4 meV) of inelastic incoherent neutron and Raman scattering spectra of many glassy materials, such as glass-forming liquids (1), polymers (2), and biological macromolecules (3–9), at cryogenic temperatures below ~200 K. As the temperature rises, the boson peak shifts to a lower frequency and becomes buried in the quasielastic contributions. Moreover, the protein boson peak shifts to higher frequencies upon hydration (5,9); hence, solvent molecules are implicated in the origin of the peak.

A peak corresponding to the boson peak has also been found in simulation studies of hydrated proteins (10–14). Our molecular dynamics (MD) simulation study of proteins in water has revealed that the structured water molecules around a protein molecule increase the number of local minima in the protein energy landscape (15,16), which in turn plays a key role in the origin of the boson peak (12). The peak appears when the protein dynamics are trapped within a local energy minimum at cryogenic temperatures. This trapping causes very low frequency collective motions to shift to higher frequencies. Among the thousands of degrees of freedom of a protein molecule involved with dynamics, only 5% of these need to be considered to understand the origin of the protein boson peak.

The ''dynamical'' or ''glass-like'' transition is another temperature-dependent phenomenon of protein dynamics, which is significantly affected by hydration levels. It is characterized as an increase in the atomic mean-square fluctuations, $\langle \Delta r^2 \rangle$, at a temperature above ~200 K. This phenomenon has been detected by various experimental techniques, such as x-ray crystallography (17–19), Mössbauer spectroscopy (20,21), and incoherent neutron scattering (4,9,22–25). This increase has been interpreted as a result of a transition in protein dynamics from harmonic to diffusive anharmonic motions. The glass-like transition in protein was shown to be suppressed in ''dry'' protein (4,9,24). Interestingly a correspondence between the glass-like transition and the onset of protein activity has been reported, e.g., below the transition temperature, ribonuclease A is unable to bind ligand in its active site (18), and bacteriorhodopsin in a purple membrane is unable to proceed proton pumping (4). It implies that protein function requires activation of diffusive anharmonic motions. Collective motions are often inferred to be important for protein function (26), and a small number of anharmonic collective motions is considered to dominate the total fluctuations (27,28). To find the functionally relevant motions and to gain insight into the protein energy landscape, it is useful to determine collective motions from molecular simulation results by principal component analysis (PCA) (28,29).

Submitted July 23, 2007, and accepted for publication December 31, 2007.

Address reprint requests to Akio Kitao, Institute of Molecular and Cellular Biosciences, University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-0032, Japan. Tel.: 81-3-5841-2297; Fax: 81-3-5841-2297; E-mail: kitao@iam.u-tokyo.ac.jp.

Editor: Arthur G. Palmer III.

0006-3495/08/06/4435/09 $2.00

doi: 10.1529/biophysj.107.118042

Solvent mobility, or the translational dynamics of water, was shown to be the dominant factor in determining protein fluctuations by MD simulations (30–33). Powder samples are commonly used in neutron experiments on biomolecules (3–9). Molecular interactions between proteins should be considered as well as those between protein and solvent to interpret the results of neutron scattering experiments. MD simulations with crystal and pseudopowder models at cryogenic temperature are employed to study the origin of the protein boson peak (11,13,14). In this work we perform MD simulations of crystalline *Staphylococcal* nuclease (SNase) for ''wet'' and ''dry'' samples at six temperatures ranging from 100 to 300 K and compare the frequency dependence of protein dynamics. First we show that the calculated spectra are in good agreement with those of the experiment (9). Then we confirm that our previous observation on the origin of the protein boson peak, in which the trapping of protein dynamics plays a key role (12), is also applicable to the simulation results. Finally we discuss the difference in the frequency dependence of protein dynamics between full hydration state (FHS) and minimal hydration state (MHS) as a function of temperature.

METHODS

Simulation

The results presented here were obtained from constant temperature and pressure MD simulations of the crystalline SNase at six temperatures ranging from 100 to 300 K using the program AMBER9 PMEMD (34). By assuming a powder state as an ensemble of microcrystals, the simulation conditions mimic the powder state of our neutron scattering experiment of SNase (9). The crystal structure of SNase (Protein Data Bank code: 1STN; 149 residues) was used as the initial structure of the simulations. The simulated systems were constructed to imitate the crystal unit cell, which has a space group symmetry of $P4_1$. Both the minimal and full hydration systems contained four protein models in the simulated box. The system in MHS contains only 332 crystal water molecules (included as D_2O) and counterions (32 chloride ions) as a solvent, which mimics a ''dry'' state realized under experimental conditions, assuming these water molecules stick to the protein even after lyophilization. FHS is designed to mimic a fully solvated ''wet'' protein in a microcrystalline state by filling the gaps in the initial structure of MHS with D_2O (1832 molecules total). Exchangeable protons in the systems were exchanged to deuterium as in ordinary neutron scattering experiments. The hydration levels of MHS and FHS are $h = 0.09$ and $h = 0.49$ g D_2O/g protein, respectively. The simulation systems in MHS and FHS contain 10,608 and 15,108 atoms in total, respectively. Periodic boundary conditions were used, and nonbonded interactions were calculated by the particle mesh Ewald method. The AMBER *ff99* force field (35) and TIP3P water model (36) were employed.

MD simulations of MHS and FHS were initiated by 3-ns runs to equilibrate the system to 300 K and 1 bar and gradually relax the restraints. From the equilibrated structures at 3 ns, the system was quenched to 100 K. Raising the temperature from 100 K, simulations were subsequently performed at 140, 180, 220, 260, and 300 K using the structure obtained at the end of the previous temperature simulation. Simulation at each temperature consisted of a 1-ns equilibration run and a 10-ns production run. Each 10-ns trajectory was stored at every 80-fs and was divided into five 2-ns trajectories, and physical quantities (neutron scattering spectra, mean-square fluctuations, etc.) were calculated as the average of the results from the five 2-ns trajectories. This simulation result was also employed for the analysis of hydration dependence of glass-like transition (46).

Inelastic neutron scattering spectra

Neutron scattering experiments essentially measure the total dynamic structure factor, $S(\mathbf{Q},\omega)$, in which $\mathbf{Q}$ and ω correspond to the momentum and energy transfers between incident neutron and sample, respectively. We calculate the incoherent dynamic structure factor, $S_{\mathrm{inc}}(\mathbf{Q},\omega)$, as the Fourier transform of a time correlation function, i.e., the intermediate scattering function, $I_{\mathrm{inc}}(\mathbf{Q},t)$:

$$S_{\mathrm{inc}}(\mathbf{Q},\omega) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \mathrm{d}t \exp(-i\omega t) I_{\mathrm{inc}}(\mathbf{Q},t) \tag{1}$$

$$I_{\mathrm{inc}}(\mathbf{Q},t) = \sum_{\mathrm{a}} b_{\mathrm{inc,a}}^2 \langle \exp(-i\mathbf{Q}\cdot\Delta\mathbf{r}_{\mathrm{a}}(0)) \exp(i\mathbf{Q}\cdot\Delta\mathbf{r}_{\mathrm{a}}(t))\rangle. \tag{2}$$

Here, $b_{\mathrm{inc,a}}$ and $\Delta\mathbf{r}_{\mathrm{a}}(t)$ are the incoherent atomic scattering length and an instantaneous deviation of position vector of atom a from its average position at time t, respectively. It should be noted that we took into account the contribution from all the nuclei in the system in this work although some of the preceding computational works considered only protons. The functions $I_{\mathrm{inc}}(Q,t)$ and $S_{\mathrm{inc}}(Q,\omega)$ reported in this work are the respective rotational averages of $I_{\mathrm{inc}}(\mathbf{Q},t)$ and $S_{\mathrm{inc}}(\mathbf{Q},\omega)$ on the sphere $|\mathbf{Q}| = Q$. To simulate the experimental resolutions, the inelastic neutron scattering spectra calculated from the MD trajectory were broadened by convolution with a Gaussian resolution function. To investigate the resolution dependence of $S_{\mathrm{inc}}(Q,\omega)$, two kinds of Gaussian widths (standard deviations), 200 and 20 μeV, were adopted. The former corresponds to the instrumental resolution of the LAM40 spectrometer at KEK, Japan, used in our experiments (9). The latter resolution corresponds to the protein dynamics analyzer (37,38) to be constructed at J-PARC, Japan. The average value of $S_{\mathrm{inc}}(Q,\omega)$ calculated at seven scattering angles ranging from 16.3° to 112.3° was employed here to compare the results of calculation to that of experiment by LAM40 (9).

Frequency-dependent protein dynamics

In this work, we examine not only $S_{\mathrm{inc}}(Q,\omega)$ but also other kinds of power spectra, $X(\omega)$ and $G(\omega)$. Here, we show their definitions and mutual relationships. First, we introduce a Q-independent spectrum, $X(\omega)$, defined as the power spectrum of the scattering-length weighted coordinate trajectory of atoms:

$$X(\omega) = \frac{1}{3}\sum_{\mathrm{a}} b_{\mathrm{inc,a}}^2 \chi_{\mathrm{a}}(\omega), \tag{3}$$

where

$$\begin{aligned}\chi_{\mathrm{a}}(\omega) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} dt \exp(-i\omega t)\langle \Delta\mathbf{r}_{\mathrm{a}}(0)\cdot\Delta\mathbf{r}_{\mathrm{a}}(t)\rangle \\ &= \frac{1}{2\pi t_{\mathrm{sim}}}\left|\int_0^{t_{\mathrm{sim}}} dt \exp(-i\omega t)\Delta\mathbf{r}_{\mathrm{a}}(t)\right|^2,\end{aligned} \tag{4}$$

and t_{sim} is the simulation time length. When not indicated explicitly, summation in equations is taken over all atoms in the system including both protein and water molecules. From Eq. 2, $I_{\mathrm{inc}}(Q,t)$ is approximated in the small Q-region as

$$\begin{aligned}I_{\mathrm{inc}}(Q,t) \sim &\sum_{\mathrm{a}} b_{\mathrm{inc,a}}^2 \exp\left(-\frac{1}{3}\langle|\Delta\mathbf{r}_{\mathrm{a}}(0)|^2\rangle Q^2\right) \\ &\times\left(1+\frac{1}{3}\langle\Delta\mathbf{r}_{\mathrm{a}}(0)\cdot\Delta\mathbf{r}_{\mathrm{a}}(t)\rangle Q^2 + \cdots\right).\end{aligned} \tag{5}$$

Thus, using Eqs. 1 and 3–5, it is shown that $X(\omega)$ and $S_{\mathrm{inc}}(Q,\omega)$ have the following relationship:

$$X(\omega) = \lim_{\mathrm{Q}\to 0}\frac{S_{\mathrm{inc}}(Q,\omega)}{Q^2}. \tag{6}$$

This function is defined for $\omega \neq 0$. The magnitude of momentum transfer Q depends on the energy transfer ω (39) at each scattering angle and the average over scattering angles is often employed for better statistics. Although the Q dependence of $S_{\mathrm{inc}}(Q,\omega)$ may provide valuable information on protein dynamics, only the frequency dependence of $S_{\mathrm{inc}}(Q,\omega)$ is discussed indirectly in the ordinary neutron scattering experiments (9). In the Results and Discussion section, we compare the frequency dependence of $X(\omega)$ to that of $S_{\mathrm{inc}}(Q,\omega)$ averaged over seven scattering angles.

The temperature-scaled power spectrum, $G(\omega)$, is defined as the spectrum of the mass-weighted coordinate trajectory of atoms. $G(\omega)$, typically called density of states, is related to $\chi_{\mathrm{a}}(\omega)$ as

$$\begin{aligned} G(\omega) &= \frac{1}{k_{\mathrm{B}}T}\sum_{\mathrm{a}} m_{\mathrm{a}} g_{\mathrm{a}}(\omega) \\ &= \frac{\omega^2}{k_{\mathrm{B}}T}\sum_{\mathrm{a}} m_{\mathrm{a}} \chi_{\mathrm{a}}(\omega), \end{aligned} \tag{7}$$

$$\begin{aligned} g_{\mathrm{a}}(\omega) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} dt \exp(-i\omega t)\langle \Delta\dot{\mathrm{r}}_{\mathrm{a}}(0)\cdot\Delta\dot{\mathrm{r}}_{\mathrm{a}}(t)\rangle \\ &= \frac{1}{2\pi t_{\mathrm{sim}}}\left|\int_0^{t_{\mathrm{sim}}} dt \exp(-i\omega t)\Delta\dot{\mathrm{r}}_{\mathrm{a}}(t)\right|^2, \end{aligned} \tag{8}$$

where k_{B}, T, and m_{a} are the Boltzmann constant, temperature, and mass of the ath atom, respectively. Because of the equipartition law of energy, $G(\omega)$ always satisfies the following condition:

$$2\int_0^{\infty} d\omega\, G(\omega) = N_{\mathrm{f}}, \tag{9}$$

where N_{f} is the total degrees of freedom. Similarly, integration of $\chi_{\mathrm{a}}(\omega)$ from $\omega = 0$ to infinity corresponds to the mean-square fluctuation of the ath atom, $\langle \Delta r_{\mathrm{a}}^2\rangle$. Here we decompose $\langle \Delta r_{\mathrm{a}}^2\rangle$ into three parts by the two frequencies, ω_1 and ω_2. Frequency ω_1 is defined as the frequency at which $X(\omega)$ has the lowest minimum below the boson peak (~1 meV). Frequency ω_2 is selected to be the minimum frequency of the hydration-independent range of spectral density $X(\omega)$ (~4 meV). Components from frequency ranges $\omega \le \omega_1$, $\omega_1 \le \omega \le \omega_2$, and $\omega_2 \le \omega$ are considered to be "low", "boson peak", and "harmonic" modes, respectively, as

$$\begin{aligned} \langle \Delta r_{\mathrm{a}}^2\rangle &= 2\int_0^{\infty} d\omega \chi_{\mathrm{a}}(\omega) \\ &= 2\int_0^{\omega_1} d\omega \chi_{\mathrm{a}}(\omega) + 2\int_{\omega_1}^{\omega_2} d\omega \chi_{\mathrm{a}}(\omega) + 2\int_{\omega_2}^{\infty} d\omega \chi_{\mathrm{a}}(\omega) \\ &= \langle \Delta r_{\mathrm{a}}^2\rangle_{\mathrm{low}} + \langle \Delta r_{\mathrm{a}}^2\rangle_{\mathrm{boson}} + \langle \Delta r_{\mathrm{a}}^2\rangle_{\mathrm{harmonic}}. \end{aligned} \tag{10}$$

Thus, the scattering length weighted mean-square fluctuation averaged over protein atoms, $\langle \Delta r^2\rangle$, is decomposed as

$$\begin{aligned} \langle \Delta r^2\rangle &= \frac{\sum_{\mathrm{a}} b_{\mathrm{inc,a}}^2 \langle \Delta r_{\mathrm{a}}^2\rangle_{\mathrm{low}} + \sum_{\mathrm{a}} b_{\mathrm{inc,a}}^2 \langle \Delta r_{\mathrm{a}}^2\rangle_{\mathrm{boson}} + \sum_{\mathrm{a}} b_{\mathrm{inc,a}}^2 \langle \Delta r_{\mathrm{a}}^2\rangle_{\mathrm{harmonic}}}{\sum_{\mathrm{a}} b_{\mathrm{inc,a}}^2} \\ &= \frac{6\int_0^{\omega_1} d\omega X(\omega) + 6\int_{\omega_1}^{\omega_2} d\omega X(\omega) + 6\int_{\omega_2}^{\infty} d\omega X(\omega)}{\sum_{\mathrm{a}} b_{\mathrm{inc,a}}^2} \\ &= \langle \Delta r^2\rangle_{\mathrm{low}} + \langle \Delta r^2\rangle_{\mathrm{boson}} + \langle \Delta r^2\rangle_{\mathrm{harmonic}}. \end{aligned} \tag{11}$$

Principal component analysis and effective frequency

PCA (28,29) is performed by diagonalizing the variance-covariance matrix **A**, defined as

$$\mathbf{A} = \langle \mathbf{q}\mathbf{q}^{\mathrm{T}}\rangle, \tag{12}$$

where **q** is the mass weighted internal displacement of the positional vector expressed as

$$\begin{aligned} \mathbf{q} = (&\sqrt{m_1}\Delta x_1, \sqrt{m_1}\Delta y_1, \sqrt{m_1}\Delta z_1, \cdots, \\ &\sqrt{m_{\mathrm{N}}}\Delta x_{\mathrm{N}}, \sqrt{m_{\mathrm{N}}}\Delta y_{\mathrm{N}}, \sqrt{m_{\mathrm{N}}}\Delta z_{\mathrm{N}})^{\mathrm{T}}. \end{aligned} \tag{13}$$

Diagonalization of **A** gives a diagonal eigenvalue matrix λ and an eigenvector matrix **V**. The diagonal elements of λ are the variances or mean-square fluctuations of the corresponding principal components. From the ith diagonal element of λ, it is possible to define an "effective frequency" (28,29), $\omega_{\mathrm{i}}^{\mathrm{eff}}$, of the ith principal component as

$$\omega_{\mathrm{i}}^{\mathrm{eff}} = \sqrt{\frac{k_{\mathrm{B}}T}{\lambda_{\mathrm{i}}}}. \tag{14}$$

The effective frequency is the frequency of the harmonic oscillator that would give the same mean-square fluctuation. Any kind of anharmonic motion that contributes to mean-square fluctuations of the principal components is reflected in the value of the effective frequency. Thus, the effective frequency is considered to give information on the curvature of the potential surface.

RESULTS AND DISCUSSION

The incoherent dynamic structure factors, $S_{\mathrm{inc}}(Q,\omega)$, as a function of frequency calculated using the results in FHS and MHS at 100 and 300 K, are shown in Fig. 1. As seen in Fig. 1 *a*, the calculated spectrum, which was smoothed by convoluting a 200-μeV resolution function corresponding to the LAM40 spectrometer, has a broad peak at ~3.5 meV at 100 K in FHS, which is higher than that in MHS (~2 meV). The peak frequency of the protein boson peak in FHS is higher than that in MHS at temperatures below 220 K (see Supplementary Material, Data S1). At 300 K, the protein boson peak becomes buried in the quasielastic contributions in FHS but is still observed in MHS. The positions of the peak in the calculated spectra in Fig. 1 *a* agree very well with that of our experiment using the 200-μeV resolution instrument (9). As seen in Fig. 1 *b*, the high resolution $S_{\mathrm{inc}}(Q,\omega)$ (width = 20-μeV) shows fine structure in the lower frequencies below the boson peak except for FHS at 300 K. To observe such fine structures experimentally, prompt construction of a high resolution spectrometer is required. In the simulation study of carboxymyoglobin at 70-μeV resolution (13), a sharp peak around 1 meV was reported at 150 K. We observed three peaks above 1 meV in both FHS and MHS at 100 K. From the results in Fig. 1, *a* and *b*, we judged that ω_1 and ω_2 in Eq. 10 should be 1 and 4 meV, respectively.

From the simulation study, it has been shown that the protein boson peak appears below the glass-transition temperature when the protein conformation is trapped within a local energy minimum and that this trapping causes the peak shift of very low frequency collective motions to higher frequencies (12). However, neither trapping nor the boson peak was observed above the glass-transition temperature. In this study, we observed a boson peak even at 300 K in MHS.

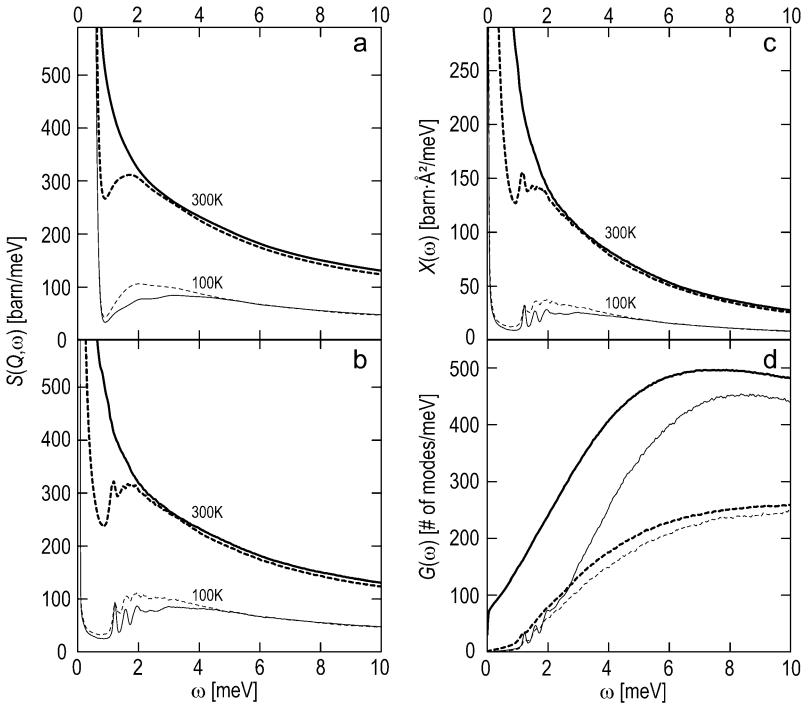

FIGURE 1 (*a* and *b*) Incoherent dynamic structure factors, $S_{inc}(Q,\omega)$, as a function of frequency, ω ($0 < \omega < 10$ meV), calculated using the results in FHS (*solid line*) and MHS (*broken line*) at 100 (*thin line*) and 300 K (*thick line*). The spectra were calculated for all atoms in the system and broadened by convolution with Gaussian resolution functions. Two kinds of Gaussian widths, (*a*) 200 μeV (resolution of LAM 40 spectrometer at KEK) and (*b*) 20 μeV, were adopted here. (*c*) The power spectrum of the scattering-length weighted coordinates, $X(\omega)$, and (*d*) the power spectrum of the mass weighted velocities, $G(\omega)$, defined as Eqs. 3 and 7, respectively. The results in FHS (*solid line*) and MHS (*broken line*) at 100 K (*thin line*) and 300 K (*thick line*) are shown. The spectra in *c* and *d* were calculated for all atoms in the system and broadened by convoluting the Gaussian resolution functions whose width is 20 μeV.

To confirm that the trapping occurs in MHS at 300 K, we examined the effective frequency calculated from PCA against the mode number as shown in Fig. 2. The effective frequencies in MHS at 300 K are significantly higher, suggesting that protein dynamics are trapped within a local minimum even at 300 K in MHS.

$X(\omega)$ and $G(\omega)$, defined as Eqs. 3 and 7 at 20-μeV resolution, are shown in Fig. 1, *c* and *d*, respectively. The frequency dependence of $X(\omega)$ (Fig. 1 *c*) is similar to that of $S_{inc}(Q,\omega)$ averaged over seven scattering angles (Fig. 1 *b*) for all simulation conditions. As discussed in Methods, the averaging of $S_{inc}(Q,\omega)$ taken for different scattering angles can be considered as an operation to effectively deduce the Q-independent function, $X(\omega)$.

Compared with the results in MHS, both $S_{inc}(Q,\omega)$ (Fig. 1, *a* and *b*) and $X(\omega)$ (Fig. 1 *c*) in FHS are in good agreement in the frequency range $\omega > 4$ meV at both temperatures. On the other hand, the shape of $G(\omega)$ in MHS differs entirely from that in FHS at both temperatures. This is due to the fact that the contribution of water to the spectra is negligibly small in the high frequency region of $S_{inc}(Q,\omega)$ and $X(\omega)$, whereas it is significantly large in $G(\omega)$. Fig. 3, *a* and *b*, shows the contribution of protein in $G(\omega)$ and that of solvent. Here, it should be noted that $G(\omega)$ satisfies the normalization given by Eq. 9. Let us first compare the spectral densities between the two models at the same temperature. For protein $G(\omega)$ (Fig. 3 *a*), there is good agreement between FHS and MHS above 4 meV, indicating that the effect of the difference in the hydration levels on protein dynamics does not appear in the frequency range $\omega > 4$ meV. However, solvent $G(\omega)$ in FHS is much larger than in MHS at all temperatures, as the number of solvent degrees of freedom in FHS is about six times larger than that in MHS (Fig. 3 *b*). Next, we compare the temperature change in each model. Protein $G(\omega)$ in FHS at 100 K significantly drops compared to the results in FHS at 300 K, indicating the density shift to higher frequency. Interestingly, solvent $G(\omega)$ in FHS decreases drastically in low temperature in the frequency range shown in this figure, corresponding to that $\langle \Delta r^2 \rangle$ of water in FHS at 100 K (0.14 Å^2) is much smaller than that at 300 K (556.6 Å^2). In Fig. 3 *c*, scaled solvent spectral densities, $G(\omega)/N_{solv}$, where N_{solv} is the number of

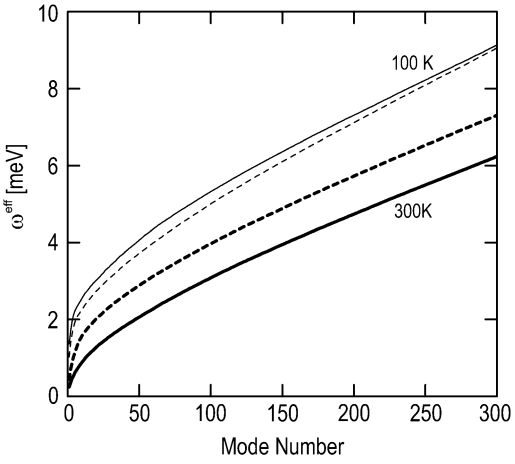

FIGURE 2 Effective frequency calculated from PCA averaged over five 2-ns trajectory plotted against mode number. The results in FHS (*solid line*) and MHS (*broken line*) at 100 K (*thin line*) and 300 K (*thick line*) are shown.

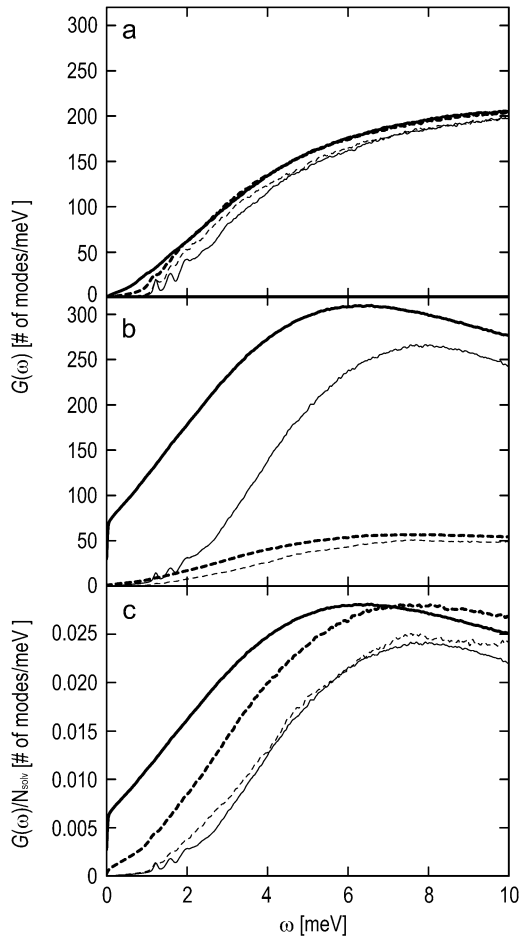

FIGURE 3 Calculated spectral densities, $G(\omega)$, for protein (*a*) and that for solvent (*b*) calculated from MD of FHS (*solid line*) and MHS (*broken line*) in the frequency range $0 < \omega < 10$ meV. (*c*) Scaled solvent spectral densities, $G(\omega)/N_{solv}$, where N_{solv} is the number of degrees of freedom in solvent, are also shown. The results at 100 K (*thin line*) and 300 K (*thick line*) are shown. The spectra were broadened by assuming a frequency resolution of 20 μeV.

degrees of freedom for the solvent, are also shown. $G(\omega)/N_{solv}$ is comparable at 100 K but considerably different at 300 K between FHS and MHS. It is noted that $\langle \Delta r^2 \rangle$ of water in FHS at 100 K (0.14 Å^2) is close to that in MHS (0.19 Å^2). At 300 K, the scaled density in MHS has higher frequency components than FHS, indicating that a higher fraction of water molecules in MHS tends to be restricted in their dynamics. This result is confirmed by the fact that $\langle \Delta r^2 \rangle$ of water in FHS (556.6 Å^2) is one order larger than that in MHS (25.8 Å^2) at 300 K.

As seen in Fig. 3, the large difference of $G(\omega)$ between FHS and MHS is due mainly to the solvent contribution, as the number of water molecules in the system differs by a factor of six. Although we observe this large shift in $G(\omega)$, no large differences are seen in $S_{inc}(Q,\omega)$ and $X(\omega)$ in the frequency ranges $\omega > 4$ meV. This actually is due to the fact that water has no significant contribution to $S_{inc}(Q,\omega)$ and $X(\omega)$ in the high frequency range. $X(\omega)$ is proportional to the scattering-length weighted sum of $\chi_a(\omega)$ over all atoms as seen from Eq. 3. Since the square value of the atomic scattering length of proton ($b_H^2 = 6.36$ barn) is much larger than that of other atoms (e.g., $b_D^2 = 0.16$ barn), the contribution of proton atoms dominates $X(\omega)$. Although the number of D_2O molecules in FHS is six times larger than that in MHS, we confirmed that the contribution of D_2O to $X(\omega)$ is negligible at high frequency ranges, $4 < \omega < 25$ meV (see Data S1). It is noted that the contribution of water to $S_{inc}(Q,\omega)$ and $X(\omega)$ at ~4 meV appears when deuterium atoms of water molecules are exchanged to proton atoms. Paciaroni et al. observed a boson peak at ~4 meV in the experimental incoherent neutron scattering spectra of protein hydration water using H_2O as solvent (7). They also observed a boson peak at ~1.5 meV in the calculated spectra of hydration atoms in protein hydration water (10).

Fig. 4 shows the temperature dependence of $\langle \Delta r^2 \rangle_{low}$, $\langle \Delta r^2 \rangle_{boson}$, and $\langle \Delta r^2 \rangle_{harmonic}$ defined as Eq. 11 calculated for protein atoms. As already mentioned, we chose 1 and 4 meV for ω_1 and ω_2. The frequency ranges, $\omega < 1$ meV and $\omega > 4$ meV correspond to the time ranges of $\tau = 2\pi/\omega > 4$ ps and $\tau < 1$ ps, respectively. As shown in Fig. 3 *a*, protein dynamics above 4 meV are similar to each other in FHS and MHS. The temperature dependence of $\langle \Delta r^2 \rangle_{harmonic}$ is also in good agreement between FHS and MHS and is almost linear. Therefore, protein dynamics in the frequency range higher than 4 meV are considered to be nearly harmonic in both FHS and MHS, as expected. This is also consistent with previous works in which the majority of the protein modes are nearly harmonic in the high frequency range at room temperature (28,40,41). $\langle \Delta r^2 \rangle_{low}$ values in both FHS and MHS are the smallest among three components at 100 K then become larger above ~140 K and dominate $\langle \Delta r^2 \rangle$ above ~220 K. $\langle \Delta r^2 \rangle_{low}$ in FHS is comparable to that in MHS below 220 K but much larger above ~220 K. These results are consistent with the simulation study in which a small number of anharmonic collective motions dominate the total fluctuations and such motions take place at a timescale of $\gg 1$ ps (28). Such slow protein motion is considered to be the origin of the glass-like transition in proteins (42). These results are consistent with the view (43) using PCA, in which the temperature dependence of $\langle \Delta r^2 \rangle$ is decomposed according to the anharmonicity factor (41).

As mentioned, we defined $1 < \omega < 4$ meV as the frequency range of the protein boson peak. The contribution of $\langle \Delta r^2 \rangle_{boson}$ to total $\langle \Delta r^2 \rangle$ is smaller than the other two components in both FHS and MHS at temperatures above 180 K. Fig. 5, *a–d*, shows the frequency dependence of $\chi_a(\omega)$ (Eq. 4) for the hydrogen atoms connected to α-carbons. It should be noted that the magnitude at 300 K is scaled by the temperature ratio of 100 K to 300 K, 1:3. Here, $\chi_a(\omega)$ is smoothed by a 20-μeV resolution function. The sum of $\chi_a(\omega)$ over the hydrogen atoms connected to α-carbons (H^α) are plotted in Fig. 5, *g* and *h*, and the frequency dependence is similar to

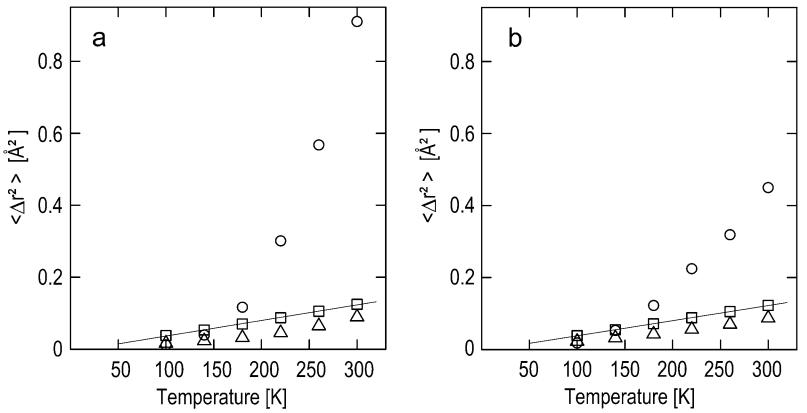

FIGURE 4 Temperature dependence of the decomposed $\langle \Delta r^2 \rangle$ averaged over protein atoms (Eq. 11) using the results in (*a*) FHS and (*b*) MHS at six temperatures ranging from 100 to 300 K. The contributions from the low-frequency dynamics lower than 1 meV, $\langle \Delta r^2 \rangle_{\mathrm{low}}$, (*circles*) that form the frequencies ranging from 1 to 4 meV, $\langle \Delta r^2 \rangle_{\mathrm{boson}}$, (*triangles*) and that form the high-frequency dynamics higher than 4 meV, $\langle \Delta r^2 \rangle_{\mathrm{harmonic}}$, (*squares*) are plotted.

that over all atoms ($X(\omega)$ in Fig. 1 *c*). Sharp peaks between 1 and 2 meV are observed, except in FHS at 300 K. As seen in Fig. 5, *b* and *d*, $\chi_{\mathrm{a}}(\omega)$ in MHS at 300 K is significantly larger than that in MHS at 100 K below 2 meV. Jumping-among-minima motions are expected to take place partly at 300 K. In the results in FHS at 300 K, peaks between 1 and 2 meV disappear and quasielastic contributions dominate $\chi_{\mathrm{a}}(\omega)$ (Fig. 5 *c*) since anharmonic motions are supposed to occur frequently.

At 100 K, $\langle \Delta r^2 \rangle_{\mathrm{boson}}$ in FHS are smaller than that in MHS, as shown in Fig. 5 *e*. The number of hydrogen bonds between protein and water molecules in FHS (360 bonds per single protein) is much larger than that in MHS (~170 per single protein), suggesting the protein in FHS is more restricted in dynamics at 100 K; i.e., the difference in $\langle \Delta r^2 \rangle_{\mathrm{boson}}$ at 100 K between FHS and MHS results in the shift of the protein boson peak (Fig. 1). On the other hand, the magnitude of $\langle \Delta r^2 \rangle_{\mathrm{boson}}$ in FHS at 300 K is comparable to that in MHS as shown in Figs. 3 and 5 *f*. The motions in the frequency range of the boson peak are distributed over whole protein as shown in Fig. 5, *a–d*. In other words, collective motions of protein contribute to the protein boson peak, which is consistent with

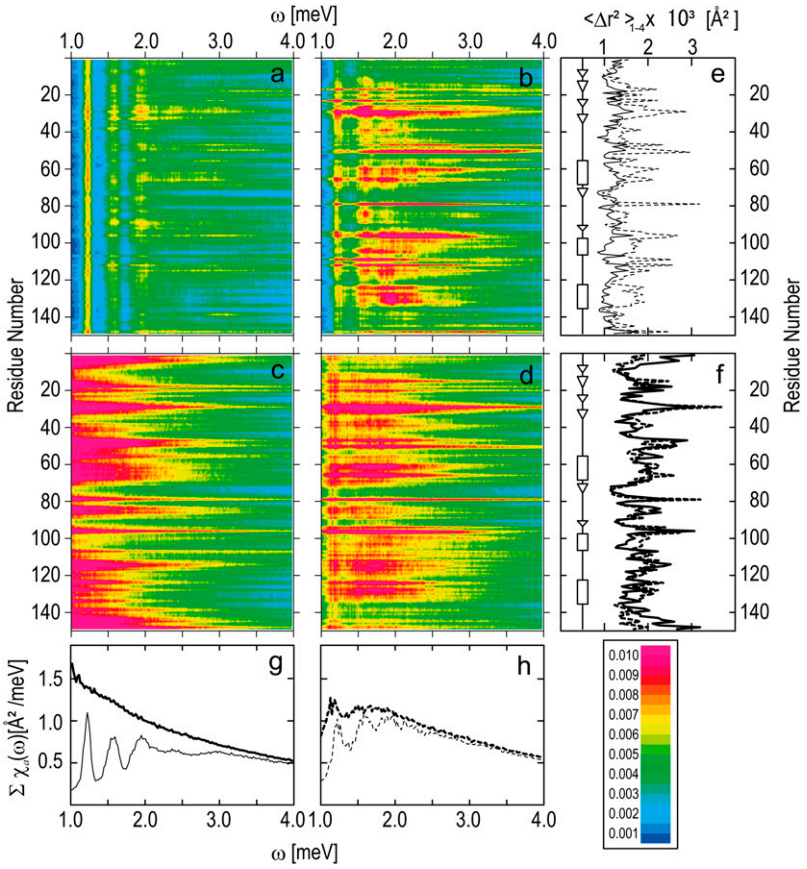

FIGURE 5 (*a–d*) The frequency dependence of the mean-square fluctuations of the H^{α} atoms, $\chi_{\mathrm{a}}(\omega)$, in Eq. 4 are drawn as two-dimensional contour plots. The results in FHS at (*a*) 100 K and (*c*) 300 K and those in MHS (*b*) 100 K and (*d*) 300 K are shown. (*e–h*) Residue dependence of $\langle \Delta r_{\mathrm{a}}^2 \rangle_{\mathrm{boson}}$ of these atoms calculated by Eq. 10 at (*e*) 100 K (*thin line*) and (*f*) 300 K (*thick line*) are shown. The sums of $\chi_{\mathrm{a}}(\omega)$ over the hydrogen atoms connected to α-carbons in (*g*) FHS (*solid line*) and (*h*) MHS (*broken line*) are shown as a function of frequency. Here, the values at 300 K are scaled by 1/3 for comparison with the results at 100 K. A diagram of the secondary structure of SNase is shown in *e* and *f*, with boxes for α-helices and triangles for β-sheets.

the views from the experiments of Kataoka and colleagues (8) and the simulations of Tarek and Tobias (11) and Kurkal-Siebert and Smith (13).

Finally, we discuss the frequency dependence of $X(\omega)$ of protein at frequencies lower than 1 meV. Interestingly, a linear relationship between the logarithm of $X(\omega)$ and that of ω is seen in the frequency range between 0.002 meV and 0.1 meV for all simulation conditions in Fig. 6. Here, no resolution function is applied to the spectra. Thus, $X(\omega)$ in this frequency range can be approximated as

$$X(\omega) = \frac{A}{\omega^{\alpha}}, \tag{15}$$

where A and α depend on the simulation conditions. When only the protein contribution was considered (Fig. 6 *b*), differences in α values are not particularly large. However, the contribution of water molecules is significantly large as seen in the α value difference between 100 and 300 K in FHS. As mentioned, the contribution from a proton ($b_H^2 = 6.36$ barn) to $X(\omega)$ is much larger than that of a deuterium ($b_D^2 = 0.16$ barn). The total numbers of protons and deutriums in the system are comparable (3688 and 4824, respectively). The significant contribution from water is due to the bulk-like water at 300 K in FHS. A total of 95% of water molecules have mean-square fluctuations two orders greater than the average value of the proton in the protein. In the very low frequency region, these bulk-like water molecules contribute to $X(\omega)$ significantly. In other cases, the protein contribution mostly determines the A and α values. A $1/\omega^{\alpha}$ trend was also found for the Fourier transformed autocorrelation function of the potential energy functions of plastocyanin at frequencies between ~0.4 and ~4 meV (44) and for $S_{inc}(Q,\omega)$ of lysozyme at frequencies between ~0.01 and ~0.4 meV (45). The fractional Brownian dynamics model is discussed in both cases (44,45), considering the fractality of the energy landscape. Here, we found a $1/\omega^{\alpha}$ trend in the lower frequency range $\omega <$ ~0.01 meV. Protein dynamics, which occur on a much longer than nanosecond timescale, are related to function and are expected to be investigated from the combination of simulations and experiment.

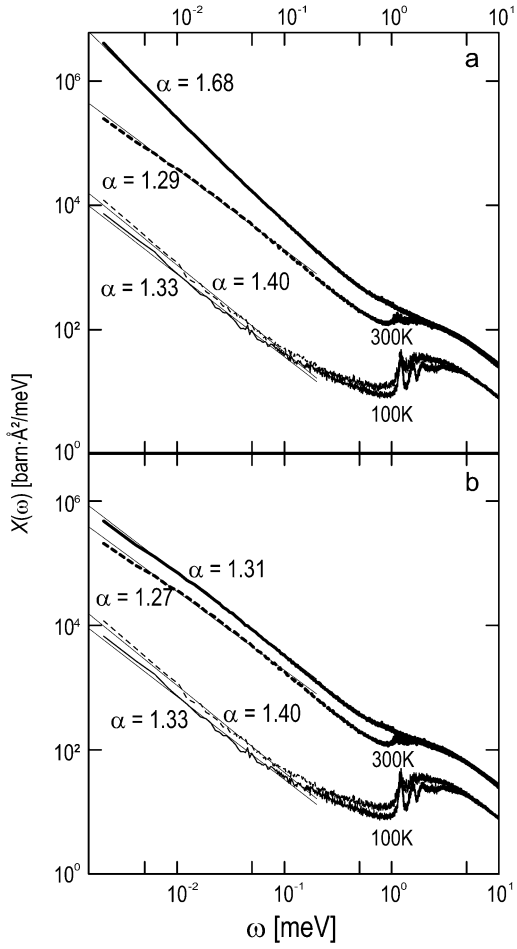

FIGURE 6 The spectra, $X(\omega)$, defined as Eq. 6, as a function of frequency, ω ($0 < \omega < 10$ meV) calculated (*a*) for all atoms in the system and (*b*) for only protein atoms using the results in FHS (*solid line*) and MHS (*broken line*) at 100 (*thin line*) and 300 K (*thick line*) in log-log plots. The regression line using the values between 0.002 and 0.1 meV is drawn for each spectrum.

To summarize, the protein boson peak is observed in the frequency range 1–4 meV at 100 K in both the minimal and FHS. Protein motions with frequencies higher than 4 meV are shown to undergo almost harmonic motion in both states at all temperatures simulated, whereas those with frequencies lower than 1 meV dominate the total fluctuations above 220 K and contribute to the origin of the glass-like transition. At 300 K, the protein boson peak becomes buried in the quasielastic contributions in the FHS but is still observed in the MHS. The peak frequency of the protein boson peak in the MHS is lower than that in the FHS at 100 K. These results are consistent with our experiment (9). Protein motions, which contribute to the protein boson peak, are collective motions expanding over the whole protein. The protein boson peak is observed when protein dynamics are trapped within a local minimum of its energy surface. The fine structure of the dynamics structure, $S_{inc}(Q,\omega)$, is expected to be detected by the experiment if a high resolution instrument ($<$~20 μeV) is developed in the near future.

SUPPLEMENTARY MATERIAL

To view all of the supplemental files associated with this article, visit www.biophysj.org.

The computations were partly performed using the facilities of the Supercomputer Center, Institute for Solid State Physics, University of Tokyo, and the supercomputers at the Research Center for Computational Science, Okazaki Research Facilities, National Institutes of Natural Sciences.

This work was supported by the Next Generation Supercomputing Project, Nanoscience Program to A.K. and Y.J., by a Grant-in-Aid for Young

Scientists (B) to Y.J., a Grant-in-Aid for Scientific Research (B) to A.K., and Grants-in-Aid for Scientific Research on Priority Areas to Y.J., H.N., M.K., and A.K. from the Ministry of Education, Culture, Sports, Science, and Technology of Japan.

REFERENCES

1. Grigera, T. S., V. Martin-Mayor, G. Parisi, and P. Verrocchio. 2003. Phonon interpretation of the 'boson peak' in supercooled liquids. *Nature*. 422:289–292.
2. Frick, B., and D. Richter. 1995. The microscopic basis of the glass transition in polymers from neutron scattering studies. *Science*. 267: 1939–1945.
3. Cusack, S., and W. Doster. 1990. Temperature dependence of the low frequency dynamics of myoglobin. Measurement of the vibrational frequency distribution by inelastic neutron scattering. *Biophys. J*. 58: 243–251.
4. Ferrand, M., A. J. Dianoux, W. Petry, and G. Zaccai. 1993. 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.
5. Diehl, M., W. Doster, W. Petry, and H. Schober. 1997. Water-coupled low-frequency modes of myoglobin and lysozyme observed by inelastic neutron scattering. *Biophys. J*. 73:2726–2732.
6. Fitter, J. 1999. The temperature dependence of internal molecular motions in hydrated and dry α-amylase: the role of hydration water in the dynamical transition of proteins. *Biophys. J*. 76:1034–1042.
7. Paciaroni, A., A. R. Bizzarri, and S. Cannistraro. 1999. Neutron scattering evidence of a boson peak in protein hydration water. *Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics*. 60: R2476–R2479.
8. Kataoka, M., H. Kamikubo, J. Yunoki, F. Tokunaga, T. Kanaya, Y. Izumi, and K. Shibata. 1999. Low energy dynamics of globular proteins studied by inelastic neutron scattering. *J. Phys. Chem. Solids*. 60:1285–1289.
9. Nakagawa, H., M. Kataoka, Y. Joti, A. Kitao, K. Shibata, A. Tokuhisa, I. Tsukushi, and N. Go. 2006. Hydration-coupled protein boson peak measured by incoherent neutron scattering. *Physica B* (*Amsterdam*). 385–86:871–873.
10. Paciaroni, A., A. R. Bizzarri, and S. Cannistraro. 1998. Molecular-dynamics simulation evidences of a boson peak in protein hydration water. *Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics*. 57:R6277–R6280.
11. Tarek, M., and D. J. Tobias. 2001. Effects of solvent damping on side chain and backbone contributions to the protein boson peak. *J. Chem. Phys*. 115:1607–1612.
12. Joti, Y., A. Kitao, and N. Go. 2005. Protein boson peak originated from hydration-related multiple minima energy landscape. *J. Am. Chem. Soc*. 127:8705–8709.
13. Kurkal-Siebert, V., and J. C. Smith. 2006. Low-temperature protein dynamics: a simulation analysis of interprotein vibrations and the boson peak at 150 K. *J. Am. Chem. Soc*. 128:2356–2364.
14. Tarek, M., and D. J. Tobias. 2000. The dynamics of protein hydration water: a quantitative comparison of molecular dynamics simulations and neutron-scattering experiments. *Biophys. J*. 79:3244–3257.
15. Kitao, A., F. Hirata, and N. Go. 1993. Effects of solvent on the conformation and the collective motions of a protein. 2. Structure of hydration in melittin. *J. Phys. Chem*. 97:10223–10230.
16. Kitao, A., F. Hirata, and N. Go. 1993. Effects of solvent on the conformation and the collective motions of a protein. 3. Free energy analysis by the extended RISM theory. *J. Phys. Chem*. 97:10231–10235.
17. Frauenfelder, H., G. A. Petsko, and D. Tsernoglou. 1979. Temperature dependent x-ray diffraction as a probe of protein structural dynamics. *Nature*. 280:558–563.
18. Rasmussen, B. F., A. M. Stock, D. Ringe, and G. A. Petsko. 1992. Crystalline ribonuclease A loses function below the dynamical transition at 220 K. *Nature*. 357:423–424.
19. Joti, Y., M. Nakasako, A. Kidera, and N. Go. 2002. Nonlinear temperature dependence of the crystal structure of lysozyme: correlation between coordinate shifts and thermal factors. *Acta Crystallogr. D. Biol. Crystallogr*. 58:1421–1432.
20. Knapp, E. W., S. F. Fischer, and F. Parak. 1982. Protein dynamics from Mossbauer spectra. The temperature dependence. *J. Phys. Chem*. 86: 5042–5047.
21. Chong, S. H., Y. Joti, A. Kidera, N. Go, A. Ostermann, A. Gassmann, and F. Parak. 2001. Dynamical transition of myoglobin in a crystal: comparative studies of x-ray crystallography and Mossbauer spectroscopy. *Eur. Biophys. J*. 30:319–329.
22. Tsai, A. M., D. A. Neumann, and L. N. Bell. 2000. Molecular dynamics of solid-state lysozyme as affected by glycerol and water: a neutron scattering study. *Biophys. J*. 79:2728–2732.
23. Doster, W., S. Cusack, and W. Petry. 1989. Dynamical transition of myoglobin revealed by inelastic neutron scattering. *Nature*. 337:754–756.
24. Roh, J. H., V. N. Novikov, R. B. Gregory, J. E. Curtis, Z. Chowdhuri, and A. P. Sokolov. 2005. Onsets of anharmonicity in protein dynamics. *Phys. Rev. Lett*. 95:038101.
25. Zaccai, G. 2000. Biochemistry—how soft is a protein? A protein dynamics force constant measured by neutron scattering. *Science*. 288: 1604–1607.
26. Berendsen, H. J. C., and S. Hayward. 2000. Collective protein dynamics in relation to function. *Curr. Opin. Struct. Biol*. 10:165–169.
27. Kitao, A., and N. Go. 1999. Investigating protein dynamics in collective coordinate space. *Curr. Opin. Struct. Biol*. 9:164–169.
28. Kitao, A., S. Hayward, and N. Go. 1998. Energy landscape of a native protein: jumping-among-minima model. *Proteins*. 33:496–517.
29. Kitao, A., F. Hirata, and N. Go. 1991. The effects of solvent on the conformation and the collective motions of protein: normal mode analysis and molecular dynamics simulation of melittin in water and in vacuum. *Chem. Phys*. 158:447–472.
30. Vitkup, D., D. Ringe, G. A. Petsko, and M. Karplus. 2000. Solvent mobility and the protein 'glass' transition. *Nat. Struct. Biol*. 7:34–38.
31. Tournier, A. L., J. C. Xu, and J. C. Smith. 2003. Translational hydration water dynamics drives the protein glass transition. *Biophys. J*. 85: 1871–1875.
32. Tarek, M., and D. J. Tobias. 2002. Role of protein-water hydrogen bond dynamics in the protein dynamical transition. *Phys. Rev. Lett*. 88:138101.
33. Bizzarri, A. R., and S. Cannistraro. 2002. Molecular dynamics of water at the protein-solvent interface. *J. Phys. Chem. B*. 106:6617–6633.
34. Case, D. A., T. A. Darden, T. E. Cheatham III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, K. M. Merz, D. A. Pearlman, M. Crowley, R. C. Walker, W. Zhang, B. Wang, S. Hayik, A. Roitberg, G. Seabra, K. F. Wong, F. Paesani, X. Wu, S. Brozell, V. Tsui, H. Gohlke, L. Yang, C. Tan, J. Mongan, V. Hornak, G. Cui, P. Beroza, D. H. Mathews, C. Schafmeister, W. S. Ross, and P. A. Kollman. 2006. AMBER9. University of California, San Francisco.
35. Wang, J., P. Cieplak, and P. A. Kollman. 2000. 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.
36. Jogensen, W. L., J. Chandrasekhar, and J. D. Madura. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys*. 79:926–935.
37. Niimura, N., K. Shibata, and H. D. Middendorf. 2002. Neutron inelastic scattering spectrometer for protein dynamics analysis: DYANA. *J. Neutron Res*. 10:163–167.
38. Takahashi, N., K. Shibata, T. J. Sato, and M. Arai. 2007. Repetition rate multiplication capability for a high energy resolution mode of DIANA at J-PARC. *J. Neutron Res*. 15:61–67.
39. Smith, J., S. Cusak, U. Pezzeca, B. Brooks, and M. Karplus. 1986. Inelastic neutron scattering analysis of low frequency motions in

proteins: a normal mode study of the bovine pancreatic trypsin inhibitor. *J. Chem. Phys.* 85:3636–3654.

40. Hayward, S., A. Kitao, and N. Go. 1994. Harmonic and anharmonic aspects in the dynamics of BPTI: a normal mode analysis and principal component analysis. *Protein Sci.* 3:936–943.
41. Hayward, S., A. Kitao, and N. Go. 1995. Harmonicity and anharmonicity in protein dynamics: a normal mode analysis and principal component analysis. *Proteins.* 23:177–186.
42. Parak, F. G. 2003. Physical aspects of protein dynamics. *Rep. Prog. Phys.* 66:103–129.
43. Tournier, A. L., and J. C. Smith. 2003. Principal components of the protein dynamical transition. *Phys. Rev. Lett.* 91:208106.
44. Carlini, P., A. R. Bizzarri, and S. Cannistraro. 2002. Temporal fluctuations in the potential energy of proteins: $1/f^{\alpha}$ noise and diffusion. *Physica D.* 165:242–250.
45. Kneller, G. R., and K. Hinsen. 2004. Fractional Brownian dynamics in proteins. *J. Chem. Phys.* 121:10278–10283.
46. Joti, Y., H. Nakagawa, M. Kataoka, and A. Kitao. 2008. Hydration-dependent protein dynamics revealed by molecular dynamics simulation of crystalline staphylococcal nuclease. *J. Phys. Chem. B.* 112: 3522–3528.