THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Water Model Tuning for Improved Reproduction of Rotational Diffusion and NMR Spectral Density

Kazuhiro Takemura[†,‡] and Akio Kitao*[,†,§]

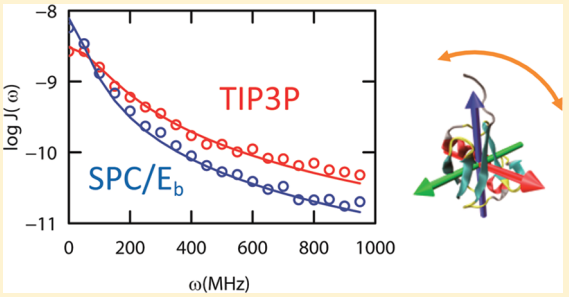
[†]Institute of Molecular and Cellular Biosciences, University of Tokyo, 1-1-1 Yayoi, Bunkyo, Tokyo 113-0032, Japan

[‡]Institute for Molecular Science, National Institute of Natural Sciences, Myodaiji, Okazaki 444-8585, Japan

[§]Core Research for Evolutional Science and Technology, Japan Science and Technology Agency, 1-1-1 Yayoi, Bunkyo, Tokyo 113-0032, Japan

Supporting Information

ABSTRACT: A water model for molecular simulation was optimized to improve the reproduction of translational and rotational diffusion of pure water and proteins. The SPC/E_b model was developed from the original SPC/E model with a slight increase of the O−H bond length of 1%. This tuning has significantly improved the translational and rotational diffusion when compared to the experimental values, whereas only small changes were observed in the other thermodynamic properties examined. The overall tumbling correlation times (τ_p) from ubiquitin, protein G, bovine pancreatic trypsin inhibitor, and barstar C42/80A were successfully reproduced using the SPC/E_b model. Calculated site-specific spectral densities of the main chain amide bond rotation in ubiquitin and protein G were in good agreement with those derived from nuclear magnetic resonance reduced spectral density mapping. The SPC/E_{bT} model was also developed with temperature-dependent bond-length tuning to facilitate reproduction of the experimental τ_p around room temperature.



INTRODUCTION

Molecular dynamics (MD) simulation and nuclear magnetic resonance (NMR) spin relaxation spectroscopy are widely used to investigate protein dynamics. Both methods can be used to probe protein motions in the picosecond to nanosecond range or even longer, and are often employed together because of their mutual complementarity in several aspects. Typical comparisons between NMR relaxation data and MD results are made for the Lipari−Szabo order parameter,[1,2] which generally shows good agreement.[3,4] However, direct comparisons of site-specific NMR spectral density $J_\alpha(\omega)$, defined as the Fourier transformation of the rotational correlation functions, differ significantly. In contrast, very good agreement between the experimental $J_\alpha(\omega)$ and the "back-calculated" spectral densities deduced from a combination of the internal motion in MD and the external motion from NMR data has been reported.[3−5] This suggests that a more direct comparison of $J_\alpha(\omega)$ would be possible if the overall tumbling correlation time of a protein (τ_p) from MD simulation could be improved.

In our previous work,[6] the τ_p of ubiquitin in water obtained using the TIP3P, SPC,[7] SPC/E,[8] and SPC/Fw[9] models was shorter than the experimental value, and this was shown to be related to the following features. (1) The translational diffusion constants of pure water and the protein are linearly dependent on the simulation box size, which is known as the finite size effect.[10,11] (2) The rotational diffusion of pure water and the protein is almost box-size-independent. (3) After finite size correction by extrapolation to infinity, the water translation diffusion constants D_w from these water models were shown to be larger than the experimental value. (4) τ_p was proportional to the inverse of D_w. After publishing this work, we noted that τ_p can be very close to the experimental value of 4.1 ns if D_w takes the value of the experimental diffusion constant (2.4×10^2 Å^2/ns) on the $(D_w)^{-1}-\tau_p$ linear regression line (see the green line in Figure 2b of ref 6). From this consideration, the water model that accurately reproduces the water diffusion constant is also expected to reproduce the protein rotational diffusion. In this work, we first propose the SPC/E_b model, in which the water diffusion constant after a finite size correction D_w and the rotational correlation time τ_w are optimized to reproduce the experimental values at 300 K by a slight modification of the SPC/E model. The τ_p values of several proteins from MD simulations are shown to be very well reproduced using this water model. In addition, we present a comparison between the NMR relaxation and MD simulation derived $J_\alpha(\omega)$ for the case of ubiquitin and protein G. Finally, we propose the temperature-dependent SPC/E_{bT} model, which can be used to reproduce experimental overall tumbling correlation times of proteins around room temperature.

Tuning of the water model conducted in this study was related to several preceding works. Lynden-Bell and co-workers[12−14] investigated the effects of bond-length and

Received: February 2, 2012
Revised: April 19, 2012
Published: May 15, 2012

dx.doi.org/10.1021/jp301100g | J. Phys. Chem. B 2012, 116, 6279−6287

-angle modifications for the SPC/E model on several properties. Höchtl et al.[15] showed that the H−O−H bond angle is a determining factor of the dielectric constant by the simulations of hybrid SPC/TIP3P models. This research mainly focused on the effects of considerable geometry modifications on water properties. In this work, we mainly aim to improve the reproduction of rotational diffusion and NMR spectral density with a slight modification of the SPC/E model.

METHOD

Translational Diffusion Constant of Water and Its Finite Size Correction. To correct the finite size effect for the translation diffusion constant, MD simulations with distinct size boxes with 360, 720, 1080, and 2160 water molecules were performed at 300 K and 1 atm. The slope of the mean square displacement of the center of molecular mass as a function of time between 0 and 0.1 ns was used to determine the size-dependent diffusion constant D. The linear relationship between the cubic root of the simulation volume and D[6,10,11] was extrapolated to obtain the corrected diffusion constant D_w in the infinite box size.

Overall Rotational Correlation Time of Water and Protein. To examine the rotational motion of pure water and protein in water, we focused on the rotational correlation function $C(t)$, defined as the second Legendre polynomial $C(t) = P_2(\cos\theta(t))$, where $\theta(t)$ is the angle formed by a certain bond vector during time interval t. The spectral density for the rotational motion of water $J(\omega)$ is defined as the Fourier transformation of $C(t)$. To calculate the rotational correlation time of pure water τ_w, $J(\omega)$ of the OH vector was calculated from the MD trajectory. It is supposed that pure water undergoes free diffusion and thus $C(t)$ is expected to take a single exponential form $C(t) = \exp(-t/\tau_w)$; therefore, $J(\omega)$ was fitted to a Lorentzian function in the frequency range 0 THz < $\omega \leq$ 1 THz.

The procedure reported by Smith and van Gunsteren[16] was followed to calculate the overall tumbling correlation time of a protein τ_p. The vectors from the center of mass of the protein to C_α atoms were adopted as the axes to define rotational motion. $J(\omega)$ spectra from all the vectors were averaged, and the spectrum was fitted to a Lorentzian function in the frequency range of 0 GHz < $\omega \leq$ 1 GHz to obtain τ_p. To examine anisotropic tumbling of protein, we first calculate the principal axes of interia for the average structure from simulations. Then, the vectors from the center of mass of the protein to C_α atoms can be divided into parallel and perpendicular components to the obtained inertia axes. The perpendicular components of vectors were used to estimate the anisotropic tumbling time around each axis.

Heat of Vaporization. The heat of vaporization (ΔH_{vap}) was calculated by[17]

$$\Delta H_{vap} \cong -(\langle U \rangle + C_{pol})/N + RT \quad (1)$$

where U, N, R, and T are the potential energy, number of water molecules, gas constant, and absolute temperature, respectively. C_{pol} is the depolarization energy when a water molecule is transferred to the gas phase[8]

$$C_{pol} = \frac{N}{2}(\mu_l - \mu_g)^2/\alpha_g \quad (2)$$

where μ_l is the dipole moment of the water model and μ_g and α_g are the dipole moment and the mean polarizability in the gas phase, respectively. C_{pol} was first considered as a term in ΔH_{vap} when the SPC/E model was developed.

Isothermal Compressibility. The isothermal compressibility (κ_T) was calculated using a finite difference expression[9,18,19] as

$$\kappa_T = \frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_T \cong \left(\frac{\ln(\rho_1/\rho_2)}{P_2 - P_1}\right)_T \quad (3)$$

where P and ρ are pressure and density, respectively. In order to calculate κ_T, two additional constant NVT simulations were performed at different densities ($\rho_1 = \rho_0 + 0.04$, $\rho_2 = \rho_0 - 0.04$, with ρ_0 as the equilibrium density at T and 1 atm).

Isobaric Heat Capacity. The isobaric heat capacity (c_p) was calculated by numerical differentiation of enthalpies

$$c_p = \left(\frac{\partial H}{\partial T}\right)_P \cong \left(\frac{H_2 - H_1}{T_2 - T_1}\right)_P \quad (4)$$

with H as the enthalpy per water molecule. To calculate c_p at T, two additional constant NPT simulations were performed at different temperatures ($T_1 = T - 10$ K, $T_2 = T + 10$ K). Following the preceding works,[9,17,18] −2.22 cal/mol/K was added to c_p as the correction term of the missing quantum effects.

Thermal Expansion Coefficient. The thermal expansion coefficient (α_p) was calculated using a finite difference expression as

$$\alpha_p = \left(\frac{\partial V}{\partial T}\right)_P \cong \left(\frac{V_2 - V_1}{T_2 - T_1}\right)_P \quad (5)$$

with V as the volume of the simulation box. To calculate α_p at T, two additional constant NPT simulations were performed at different temperatures ($T_1 = T - 10$ K, $T_2 = T + 10$ K).

Static Dielectric Constant. The static dielectric constant ($\varepsilon(0)$) is determined by the magnitude and density of the molecular dipole moments, which is expressed as

$$\varepsilon(0) = 1 + 4\pi\frac{\langle \mathbf{M}^2 \rangle - \langle \mathbf{M} \rangle^2}{3\langle V \rangle k_B \langle T \rangle} \quad (6)$$

where $\mathbf{M}$ is the system's total dipole moment.

Water Model Tuning. The calculated τ_p is expected to be closer to the experimental value if D_w calculated after finite size correction is close to the experimental value. In this work, tuning of the water model was conducted to improve the reproduction of D_w and τ_w. The SPC/E model was selected as the base model for tuning, which showed D_w closest to the experimental value in our previous work.[6] Tuning was conducted by slightly adjusting the bond length or bond angle of the SPC/E model. MD simulations of pure water were conducted for 12 ns (2 ns for the equilibration run and 10 ns for the production run), while configurations were recorded every 0.05 ps.

Protein Simulation and Calculation of Overall Tumbling Correlation Time. Improvement of both D_w and τ_w in pure water are also expected to contribute to a better reproduction of τ_p in the MD simulation. MD simulations of several proteins were performed using the tuned water model, and the experimental and calculated τ_p were compared. The ff99SB force field[20] was used for MD simulation of the proteins, and protein configurations were recorded every 0.1 ps for analysis.

dx.doi.org/10.1021/jp301100g | J. Phys. Chem. B 2012, 116, 6279−6287

Site Specific Rotational Spectra of Protein. Site-specific rotational spectra $J_\alpha(\omega)$ for the NH bond vectors of the αth main chain amides in protein were calculated three different ways. $J_\alpha^{\mathrm{direct}}(\omega)$ is defined as the direct calculation of $J_\alpha(\omega)$ from the MD trajectories. The back-calculated $J_\alpha(\omega)$[3−5] ($J_\alpha^{\mathrm{back}}(\omega)$) was obtained using the procedure reported by Showalter and Brüschweiler.[4] Assuming that the internal and external motion of the protein can be treated independently, the rotational correlation function $C_\alpha(t)$ can be expressed as

$$C_\alpha(t) = C_\alpha^{\mathrm{int}}(t)C^{\mathrm{ext}}(t) \tag{7}$$

Site-dependent internal correlation functions $C_\alpha^{\mathrm{int}}(t)$ were calculated after removing overall tumbling and translation by superimposing each snapshot to the average structure. The correlation functions were then fitted to a multiexponential function:

$$C_\alpha^{\mathrm{int}}(t) = A_{\alpha 0} + \sum_{i=1}^{5} A_{\alpha i} \exp(-t/\tau_{\alpha i}) \tag{8}$$

with the conditions $\sum A_{\alpha i} = 1$, $A_{\alpha i} \geq 0$, and $\tau_{\alpha i} \geq 0$. $A_{\alpha 0}$ is equivalent to the Lipari−Szabo order parameter (S^2).[1,2] Fitting from eq 8 was performed in the time range of $t \leq 6$ ns. The external correlation is described by a single exponential function with the experimentally derived relaxation time $\tau_{\mathrm{p}}^{\mathrm{exp}}$ = 4.1 ns[21,22] (ubiquitin) as

$$C^{\mathrm{ext}}(t) = \exp(-t/\tau_{\mathrm{p}}^{\mathrm{exp}}) \tag{9}$$

After substitution of eqs 8 and 9 into eq 7, the Fourier transformation of $C_\alpha(t)$ yields $J_\alpha^{\mathrm{back}}(\omega)$. This back calculation procedure includes two effects; the correction of overall tumbling motion and smoothing of the internal motion.

To examine the smoothing effects, $J_\alpha^{\mathrm{direct}}(\omega)$ was fitted to a multi-Lorentzian function:

$$J_\alpha^{\mathrm{direct}}(\omega) = \sum_{i=1}^{5} \frac{2A_i}{1 + (\omega\tau_i)^2} \tag{10}$$

where $J_\alpha^{\mathrm{smooth}}(\omega)$ is defined as the fitted $J_\alpha(\omega)$.

The obtained $J_\alpha^{\mathrm{direct}}(\omega)$, $J_\alpha^{\mathrm{back}}(\omega)$, and $J_\alpha^{\mathrm{smooth}}(\omega)$ were compared to the experimental $J_\alpha(\omega)$ and $J_\alpha^{\mathrm{exp}}(\omega)$, which were calculated using reduced spectral density mapping.[23−25] NMR relaxation parameters were taken from the Biological Magnetic Resonance Data Bank (bmr4245[22] for ubiquitin and bmr5569[26] for protein G).

Other Simulation Details. All the MD simulations were performed using the PMEMD module of AMBER 10.[27] The system was brought to thermodynamic equilibrium at the target temperature and 1 atm using a weak coupling thermostat and barostat. The target temperature was 300 K unless otherwise specified. The equations of motion were integrated with a time step of 2 fs. The long-range Coulomb energy was evaluated using the particle mesh Ewald (PME) method. For each simulation, initial velocities were generated to obey the Maxwell distribution without the external velocity of the total system.

RESULTS AND DISCUSSION

Relationship between Protein Rotation and Water Diffusion. The water model dependency of τ_{p} in ubiquitin on D_{w} and τ_{w} is shown in Figure 1 for the TIP3P,[28] TIP4P,[28] SPC,[7] and SPC/E[8] models. All the τ_{p} values were calculated from 60 ns MD trajectories, which is 1 order longer than τ_{p} values, as suggested[4,29] (see Figure S1 in the Supporting Information for the simulation time dependence of τ_{p}). The solid lines represent the results of linear regression, and the broken lines are from NMR spin relaxation experiments.[21,22] As a water molecule moves faster (larger D_{w} and smaller τ_{w}), ubiquitin also moves faster (smaller τ_{p}). The solid lines in Figure 1 intersect with the broken lines at $D_{\mathrm{w}} = 2.39 \times 10^2$ Å^2/ns (Figure 1a) and τ_{w} = 2.06 ps (Figure 1b), which are close to the experimental values 2.40 × 10^2 Å^2/ns[30] and 1.95 ps,[31,32] respectively. The water model that reproduces the experimental D_{w} and τ_{w} is more suitable for the simulation of protein diffusion.

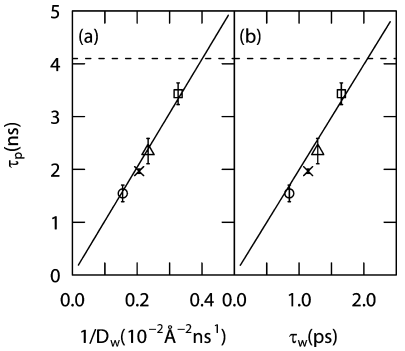


Figure 1. Dependence of the rotational correlation time of ubiquitin τ_{p} on (a) the translational diffusion constant D_{w} and (b) the rotational correlation time τ_{w} using the TIP3P (circle), TIP4P (triangle), SPC (cross), and SPC/E (square) models. Error bars were calculated as the standard deviations in three 60 ns MD simulations for SPC and TIP4P and ten 60 ns MD simulations for TIP3P and SPC/E. τ_{p} = 1.55 ± 0.16 (TIP3P), 2.35 ± 0.24 (TIP4P), 1.97 ± 0.04 (SPC), and 3.43 ± 0.21 (SPC/E) ns.

Tuning of SPC/E. Among the water models examined in Figure 1, SPC/E shows the closest D_{w} and τ_{w} to the experimental values. Using SPC/E as the base model, we examined the effects of bond length and angle modification on several properties of the water models at 300 K and 1 atm. Bond length variations were examined from 1.00 Å (original) to 1.04 Å with an interval of 0.01 Å. All the models exhibit a simulation box size dependence of the translational diffusion constant (Figure 2a). The models with longer bond length slow down the diffusion (smaller D_{w} and larger τ_{w}). The broken lines in Figure 2b and c indicate the experimental values.[30−32] The

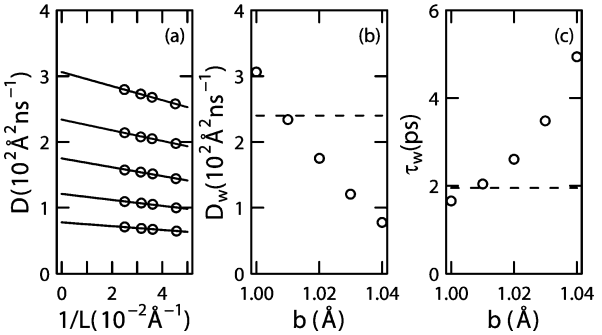


Figure 2. Effect of bond length modification based on the SPC/E model. (a) Box size dependence of the translational diffusion constant. Bond lengths are 1.00, 1.01, 1.02, 1.03, and 1.04 Å from the top. Bond length dependence of (b) the corrected translational diffusion constant D_{w} and (c) the rotational correlation time, τ_{w}. L and b indicate the cubic root of the simulation box volume and the O−H bond length of water, respectively.

dx.doi.org/10.1021/jp301100g | J. Phys. Chem. B 2012, 116, 6279−6287

modified SPC/E model with 1% longer bond length (hereafter referred to as SPC/E_b) gives the closest values to the experimental D_w and τ_w. The effects of bond angle modification are shown in Figure 3. H−O−H bond angles are examined for

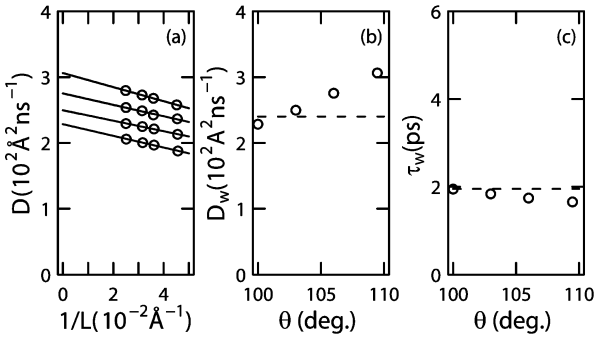


Figure 3. Effect of bond angle (θ) modification based on the SPC/E model. (a) Box size dependence of the translational diffusion constant. Bond angles are 109.47, 106, 103, and 100° from the top. Bond angle dependence of (b) the corrected translational diffusion constant D_w, and (c) the rotational correlation time, τ_w.

109.47 (original), 106, 103, and 100°. The simulation box size dependence of the translational diffusion constant was observed again, and the model with the smaller bond angle revealed slower diffusion. Among the three variations examined, the modified SPC/E model with a bond angle of 103° (SPC/E_θ) best reproduces the experimental D_w and τ_w shown by the broken lines. The parameters of SPC/E_b, SPC/E_θ, and the other water models are summarized in Table 1, and the properties of the water models are compared to the experimental values in Table 2. Both the SPC/E_b and SPC/E_θ models result in a significant improvement of the D_w and τ_w values; however, the deviations from the target density and ΔH_{vap} are slightly worse than the original SPC/E. In the case of SPC/E_b, the effects of the modification on the density (1%) and ΔH_{vap} (4%) are small compared to those in D_w (22%) and τ_w (23%). The situation is similar for SPC/E_θ, with the effects on density (2%) and ΔH_{vap} (1%). The other properties of SPC/E_b, i.e., the isothermal compressibility (κ_T), the isobaric heat capacity (c_p), the thermal expansion coefficient (α_p), and the static dielectric constant ($\varepsilon(0)$), do not show large deviation from the original SPC/E model (α_p and $\varepsilon(0)$ are closer to experimental values than the original model). The situation is similar in SPC/E_θ; however, due to the larger dipole moment, SPC/E_θ shows very large $\varepsilon(0)$. Considering these results, it was concluded that the SPC/E_b model is the most suitable among the models examined to reproduce water diffusion properties. It was also confirmed that the radial distribution functions for SPC/E_b and SPC/E_θ do not significantly differ from that for SPC/E (see Figure S2, Supporting Information).

Overall Tumbling Correlation Time of Proteins with SPC/E_b. We first examined τ_p of ubiquitin using the SPC/E_b model. As expected, the obtained $\tau_p^{calc} = 4.30 \pm 0.21$ ns is much closer to the experimental value $\tau_p^{exp} = 4.1$ ns[21,22] than the results shown in Figure 1 or Figure 4 (filled symbols). The development of SPC/E_b was carried out based on simulations of ubiquitin; therefore, calculations of τ_p were also conducted for other proteins (protein G, bovine pancreatic trypsin inhibitor (BPTI), and barstar C42/80A), as shown in Table 3 and Figure 4. Protein concentrations were set to be the same as in the NMR experiments. Error bars were calculated as the standard deviations from MD simulations of protein G and BPTI for 50 ns and of ubiquitin and barstar C42/80A for 60 ns. Simulation times were chosen after monitoring the convergence behavior of τ_p^{calc} (see Figure S1, Supporting Information). If complete overlap of the symbols with the broken line is observed, this indicates perfect agreement between the experimental and simulated τ_p. The results presented in Figure 4 indicate that the SPC/E_b model can reproduce overall tumbling of these proteins very well.

Spectral Density of Ubiquitin. Here we examine if improvement of the external protein motion in the MD simulation also contributes to a better reproduction of site-specific $J_\alpha(\omega)$ for the protein. The residue dependence of $J_\alpha^{direct}(\omega)$ from MD and $J_\alpha^{exp}(\omega)$ from the reduced spectral density mapping of the NMR relaxation data at 0, ω_N (50 MHz) and ω_H (500 MHz), respectively, is shown in Figure 5. $J_\alpha^{exp}(\omega)$ and $J_\alpha^{direct}(\omega)$ obtained using TIP3P and SPC/E_b are shown by boxes and by red and green symbols, respectively. The error bars were calculated as the standard deviation of ten 60 ns MD simulations. The deviation between $J_\alpha^{exp}(\omega)$ and simulated $J_\alpha^{calc}(\omega)$ ($J_\alpha^{direct}(\omega)$, $J_\alpha^{back}(\omega)$, or $J_\alpha^{smooth}(\omega)$) is calculated as

$$R(\omega) = \frac{\sum_\alpha |J_\alpha^{exp}(\omega) - J_\alpha^{calc}(\omega)|}{\sum_\alpha J_\alpha^{exp}(\omega)} \quad (11)$$

which is similar to the R factor in X-ray crystallography.

The obtained $R(\omega)$ values are summarized in Table 4. Figure 5 and Table 4 indicate that $J_\alpha^{direct}(\omega)$ obtained with the SPC/E_b model has better agreement with $J_\alpha^{exp}(\omega)$ than $J_\alpha^{direct}(\omega)$ obtained with the TIP3P and SPC/E models. The back-calculated $J_\alpha^{back}(\omega)$ was also examined; Table 4 (second to last line) shows that $J_\alpha^{back}(\omega)$ correlates very well with $J_\alpha^{exp}(\omega)$ at all frequencies examined. $J_\alpha^{back}(\omega)$ obtained using the other water models also agree similarly (data now shown), which suggests that the water model mainly affects overall tumbling. To

Table 1. Parameters of the Water Models

model	b (Å)	θ (deg)	q_O (e)	ε_O (kcal/mol)	σ_O (Å)	k_b (kcal/mol/Å^2)	k_θ (kcal/mol/rad^2)
TIP3P	0.9572	104.52	−0.834	0.1520	3.1508		
TIP4P[a]	0.9572	104.52	−1.04	0.1550	3.1537		
SPC	1.0	109.47	−0.82	0.1553	3.1657		
SPC/E	1.0	109.47	−0.8476	0.1553	3.1657		
SPC/Fw	1.012	113.24	−0.82	0.1554	3.1655	1059.162	75.90
SPC/E_b	1.01	109.47	−0.8476	0.1553	3.1657		
SPC/E_θ	1.0	103	−0.8476	0.1553	3.1657		

[a]Negative charge is located at 0.15 Å from the oxygen atom for TIP4P.

dx.doi.org/10.1021/jp301100g | J. Phys. Chem. B 2012, 116, 6279−6287

Table 2. Characteristics of the Water Models

	Models							
properties[a]	TIP3P	TIP4P	SPC	SPC/E	SPC/Fw	SPC/E_b	SPC/E_θ	expt
density (g/cm^3)	0.983	0.992	0.975	0.997	1.006	1.005	1.016	0.9965[33]
D_w (10^2 Å^2/ns)	6.43	4.28	4.85	3.06	3.18	2.34	2.50	2.4[30]
τ_w (ps)	0.85	1.29	1.14	1.65	1.75	2.04	1.83	1.95[31,32]
ΔH_{vap}[b] (kcal/mol/K)	8.95 (10.14)	9.94 (10.46)	9.66 (10.52)	10.53 (11.73)	9.27 (10.69)	10.97 (12.29)	10.46 (12.72)	10.50[33]
κ_T (10^{-5}/atm)	5.74	5.12	5.42	4.52	4.50	4.40	4.30	4.505[34]
c_p (cal/mol/K)	16.76	18.83	17.68	18.48	24.98	19.02	18.67	18.00[33]
α_p (10^{-4}/K)	9.19	5.76	7.36	5.22	5.04	4.18	4.32	2.747[34]
$\varepsilon(0)$	96.77	51.65	65.01	70.40	75.92	71.13	135.5	77.74[35]

[a]All the properties except for D_w and τ_w were calculated from the 2160 water box. [b]Values in brackets show the values without polarization corrections.

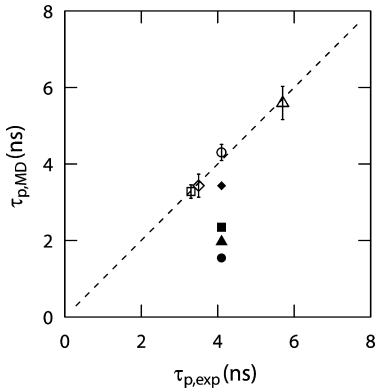


Figure 4. Open symbols represent τ_p of ubiquitin (circle), BPTI (diamond), protein G (square), and barstar C40/82A (triangle) using the SPC/E_b model. Filled symbols indicate τ_p of ubiquitin using the TIP3P (circle), TIP4P (square), SPC (triangle), and SPC/E (diamond) models.

determine the difference between $J_\alpha^{back}(\omega)$ and $J_\alpha^{direct}(\omega)$, the ω dependence of $J_\alpha(\omega)$ was investigated. Examples of $J_\alpha(\omega)$ for several residues are shown in Figure 6 (see Figure S3 in the Supporting Information for all the residues). The ω dependences of $J_\alpha^{exp}(\omega)$ (crosses), $J_\alpha^{back}(\omega)$ (solid lines), and $J_\alpha^{direct}(\omega)$ (points with error bars) obtained with SPC/E_b are quite similar. The difference between $J_\alpha^{direct}(\omega)$ obtained with SPC/E_b and $J_\alpha^{exp}(\omega)$ is within the range of error. One main reason for the good correlation between $J_\alpha^{back}(\omega)$ and $J_\alpha^{exp}(\omega)$ is the use of an experimentally derived overall tumbling time. Another reason is the smoothing in eq 8, which reduces the noise. To examine only the smoothing effects, smoothed $J_\alpha(\omega)$, $J_\alpha^{smooth}(\omega)$, was calculated using eq 10. The obtained $J_\alpha^{smooth}(\omega)$ reproduced $J_\alpha^{exp}(\omega)$ very well (the last line in Table 4). Thus, the smoothing effect has a significant contribution to the good correlation between the simulated $J(\omega)$ and $J_\alpha^{exp}(\omega)$. Compared to $J_\alpha^{back}(\omega)$, $J_\alpha^{smooth}(\omega)$ has similar agreement with $J_\alpha^{exp}(\omega)$ at ω_N and ω_H but not at zero frequency. The agreement is significantly dependent on the simulation length and statistics. The results shown in Figure 5 and Table 3 were calculated from ten 60 ns MD simulations. When $J_\alpha^{smooth}(\omega)$ is calculated from thirty 20 ns MD simulations (summarized in Table S1, Supporting Information), better agreement was observed for $J(\omega_H)$ but not for $J(0)$. The latter is related to the higher resolution in the frequency domain achieved with longer simulation, and the former is due to less statistics. Therefore, further improvement in $J(\omega)$, especially in the low ω range, could be assisted by longer MD simulation.

Spectral Density of Protein G and Anisotropic Rotational Diffusion. Protein G is known as an anisotropic tumbler,[29,42] which is a suitable target for examining the reproduction of $J_\alpha^{exp}(\omega)$. The residue dependence of $J_\alpha^{smooth}(\omega)$ and $J_\alpha^{exp}(\omega)$ from the reduced spectral density mapping of the NMR relaxation data at $\omega = 0$, ω_N (80 MHz), and ω_H (800 MHz), respectively, is shown in Figure 7. At all the frequencies, $J_\alpha^{smooth}(\omega)$ are in good agreement with $J_\alpha^{exp}(\omega)$. The average $R(\omega)$ is 0.18 (values at all frequencies examined are summarized in Table S2, Supporting Information), which is comparable to the value calculated for ubiquitin. Thus, the current procedure (SPC/E_b model and smoothing) successfully reproduces $J_\alpha^{exp}(\omega)$ of relatively anisotropic protein. To examine the box size effect on the anisotropic tumbling, we performed two additional simulations with distinct protein concentrations, 5.1 and 14.6 mol/L, in which the distances between the outermost protein atom and the closest simulation box face in the initial setup are 20 and 10 Å, respectively. Anisotropic rotation of protein G was represented by the overall tumbling times around the principal axes of inertia. The obtained results are summarized in Table 5. In a relatively large simulation box (1.6 and 5.1 mol/L), tumbling times and anisotropy A_τ (see Table 5 for the definition) show similar values. However, in the small box size with very high concentration (14.6 mol/L), the protein rotates more slowly (larger τ_p) and less anisotropically. Therefore, anisotropic rotational diffusion is nearly box size independent except for the MD in the smallest box. The anisotropy calculated by the program r2r1_diffusion[21] using 800 MHz NMR relaxation data (bmr5569[26]) and protein structure (1GB1[39]) was 1.34, which is slightly larger than 1.21 ± 0.07 reported in Table 5. It should be noted that protein is

Table 3. Details of the Protein Simulation Setup

protein	PDB ID	concentration (mM)	simulation time	τ_p^{exp} (ns)	τ_p^{calc} (ns)
ubiquitin	1D3Z[36]	4.0	60 ns × 10	4.1[21,22]	4.30 ± 0.21
BPTI	1PIT[37]	3.0	50 ns × 4	3.5[38]	3.43 ± 0.30
protein G	1GB1[39]	1.6	50 ns × 4	3.3[40]	3.28 ± 0.18
barstar C40/82A	1AB7[41]	3.5	60 ns × 5	5.7[41]	5.59 ± 0.44

dx.doi.org/10.1021/jp301100g | J. Phys. Chem. B 2012, 116, 6279−6287

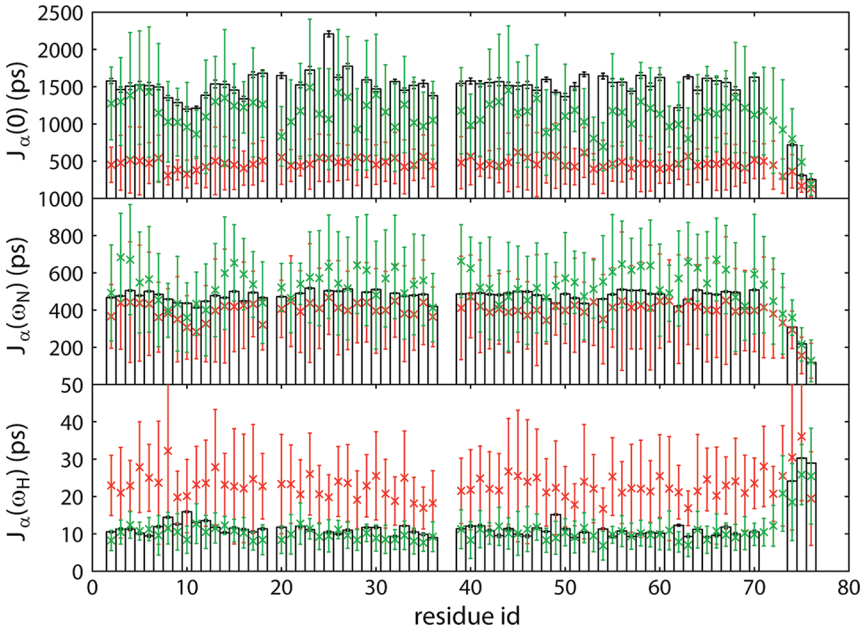


Figure 5. Site-specific spectral densities of ubiquitin $J_\alpha^{direct}(\omega)$ as a function of the residue number. Red and green crosses with error bars are from MD simulation with the TIP3P and SPC/E_b models, respectively. Boxes represent the values obtained from experimental NMR reduced spectral density mapping.[22]

Table 4. Correlation between Simulated and Experimental $J(\omega)$

water model	method[a]	$R(0)$	$R(\omega_N)$	$R(\omega_H)$	$\langle R(\omega)\rangle$
TIP3P	direct	0.69	0.16	0.96	0.60
SPC/E	direct	0.32	0.12	0.15	0.20
SPC/E_b	direct	0.24	0.16	0.15	0.18
SPC/E_b	back	0.10	0.10	0.12	0.11
SPC/E_b	smooth	0.20	0.09	0.14	0.15

[a]Direct, direct calculation; back, back calculation; smooth, smoothing calculation (see text).

regarded as a rigid body in r2r1_diffusion, whereas it is flexible in the MD simulation. Also, considering the statistical error, the difference can be judged as insignificant. In total, we conclude that MD simulation with the SPC/E_b water model reproduces the anisotropic rotational behavior of protein G in solution very well.

Temperature-Dependent SPC/E_{bT} Model. The parameter of the SPC/E_b model was tuned at 300 K and 1 atm with a bond length interval of 0.01 Å. The SPC/E_b model proved very effective in reproducing the overall tumbling correlation time and NMR spectral densities at 300 K. Therefore, the effect of bond length modification on D_w was further examined at several temperatures (287, 300, 309, and 326 K) using a smaller bond length interval of 0.001 Å. Some characteristics of the modified models that best reproduce D_w at each temperature are summarized in Table 6. Although the bond length was tuned to reproduce experimental D_w, other properties are also reproduced well. The bond-length dependency of D_w at several temperatures is shown in Figure 8a. As the temperature increases (from bottom to top in Figure 8a), the smaller bond length models fit to the experimental values (broken lines). Solid lines are the results of linear regressions for D_w from four models with different bond lengths. The intersections of the solid and broken lines indicate the optimized bond length b_T, of which the temperature dependence is shown in Figure 8b. The

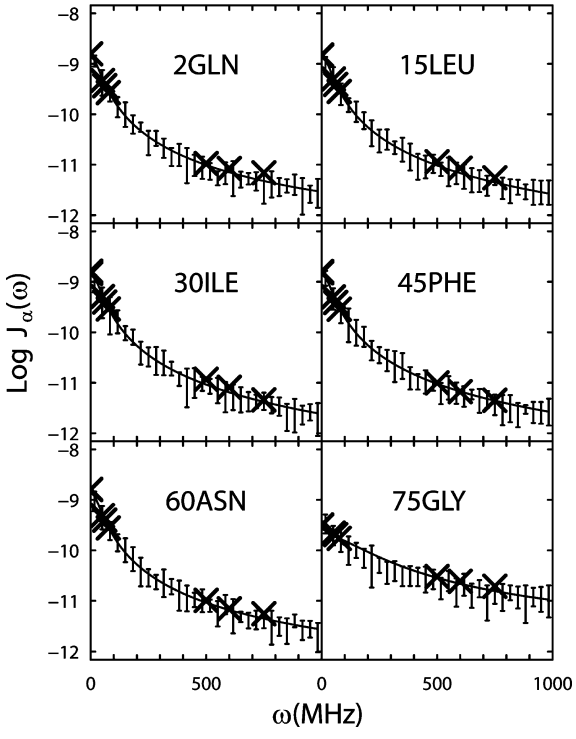


Figure 6. Examples of $J_\alpha^{exp}(\omega)$ (cross), $J_\alpha^{direct}(\omega)$ (point with error bar), and $J_\alpha^{back}(\omega)$ (line) dependence on ω.

solid line in Figure 8b is the result of linear regression, $b_T = (-7.16 \times 10^{-5} \times T + 1.0307)$ Å, where T is temperature in kelvins). In the temperature range not significantly deviated from 300 K, the overall tumbling correlation time and NMR spectral densities are expected to be reproduced well using the temperature dependent bond length model, SPC/E_{bT} (SPC/E with the bond length b_T).

dx.doi.org/10.1021/jp301100g | J. Phys. Chem. B 2012, 116, 6279−6287

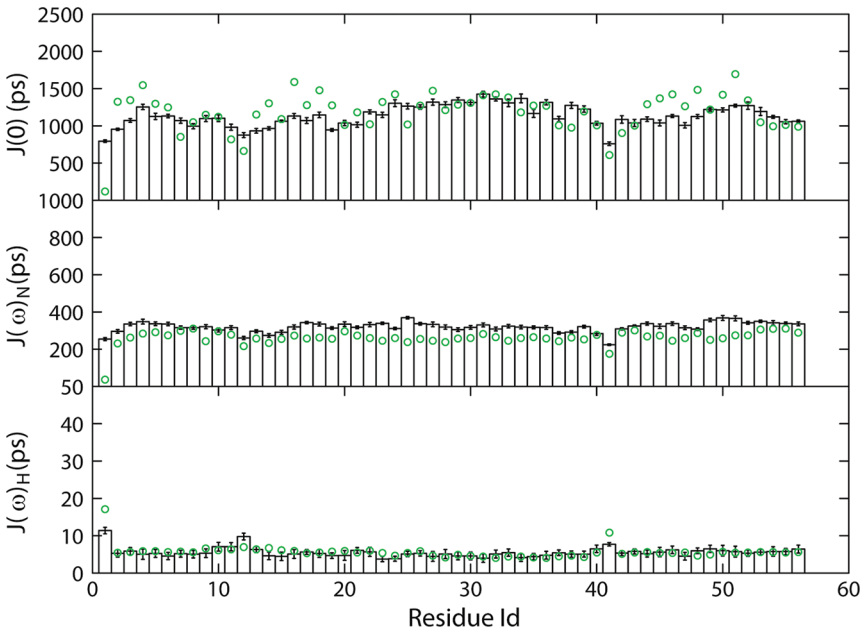


Figure 7. Smoothed site-specific spectral densities of protein G $J_\alpha^{\text{smooth}}(\omega)$ as a function of the residue number from MD simulation with the SPC/E_b model (green circles with error bars). Boxes represent the values obtained from experimental NMR reduced spectral density mapping.[26]

Table 5. Correlation between Simulated and Experimental $J(\omega)$ of protein G

c^a (mol/L)	τ_p (ns)	$\tau_X{}^b$ (ns)	τ_Y (ns)	τ_Z (ns)	$A_\tau{}^c$
1.6	3.28 ± 0.18	3.39 ± 0.20	3.63 ± 0.35	2.89 ± 0.13	1.21 ± 0.07
5.1	3.18 ± 0.23	3.38 ± 0.27	3.56 ± 0.29	2.91 ± 0.32	1.20 ± 0.10
14.6	3.62 ± 0.35	3.76 ± 0.30	4.09 ± 0.21	3.58 ± 0.47	1.10 ± 0.12

[a]Protein concentration. [b]Tumbling time around three principal axes of inertia. [c]A_τ is calculated by $(\tau_X + \tau_Y)/2\tau_Z$.

Table 6. Characteristics of the SPC/E_{bT} Model at Temperatures around 300 K

	temperatures (K)			
properties[a,b]	287	300	309	326
b (Å)	1.010	1.009	1.009	1.007
density (g/cm^3)	1.009 (0.999)	1.004 (0.997)	1.000 (0.994)	0.988 (0.987)
D_w (10^2 Å^2/ns)	1.72 (1.71)	2.43 (2.40)	2.95 (2.96)	4.22 (4.17)
τ_w (ps)	2.61 (2.82)	2.00 (1.95)	1.73 (1.54)	1.36 (1.08)
$\Delta H_{vap}{}^c$ (kcal/mol/K)	11.14 (10.63)	10.93 (10.50)	10.81 (10.40)	10.52 (10.23)
κ_T (10^{-5}/atm)	4.17 (4.70)	4.36 (4.51)	4.43 (4.44)	4.80 (4.42)
c_p (cal/mol/K)	19.65 (18.14)	19.25 (18.13)	18.74 (18.13)	18.26 (18.13)
α_p (10^{-4}/K)	2.95 (1.37)	4.28 (2.74)	4.98 (3.53)	6.06 (4.77)
$\varepsilon(0)$	72.92 (82.51)	70.18 (77.74)	69.96 (74.60)	65.99 (69.00)

[a]Except for D_w and τ_w, all the properties were calculated from the 2160 water box. [b]Values in brackets are from experiments.[30−35] [c]ΔH_{vap} values were calculated including polarization correction.

CONCLUSION

The SPC/E_b water model was developed with a slight increase of the O−H bond length of 1% from the original SPC/E model to better reproduce the translational and rotational diffusion of pure water by MD simulation. This tuning significantly improved the translational diffusion constant D_w and rotational relaxation time τ_w with respect to the experimental values, whereas only small changes were observed in the other properties. Experimental overall tumbling correlation times for four proteins with distinct sizes were successfully reproduced using the SPC/E_b model.

Calculated site-specific spectral densities of the main chain amide bond rotation in ubiquitin $J_\alpha(\omega)$ were compared to $J_\alpha^{\text{exp}}(\omega)$ derived from the reduced spectral density mapping of NMR relaxation data. $J_\alpha^{\text{direct}}(\omega)$ calculated directly from MD simulation using the SPC/E_b model reproduced $J_\alpha^{\text{exp}}(\omega)$ well; however, the agreement with $J_\alpha^{\text{exp}}(\omega)$ was better for $J_\alpha^{\text{back}}(\omega)$ calculated from the fitted internal correlation functions and experimentally derived external correlation function used in previous work.[3−5] The smoothed spectral density $J_\alpha^{\text{smooth}}(\omega)$ also agreed with $J_\alpha^{\text{exp}}(\omega)$ very well and was comparable to $J_\alpha^{\text{back}}(\omega)$. Although estimation of $J_\alpha^{\text{back}}(\omega)$ from MD simulation was shown to be very useful to interpret the results of NMR experiments,[3−5] two conditions were required: independence of the internal and external rotation and information on the experimental overall tumbling time. $J_\alpha^{\text{exp}}(\omega)$ could be reproduced with good agreement comparable to $J_\alpha^{\text{back}}(\omega)$ using the SPC/E_b model without these two requirements, which is

dx.doi.org/10.1021/jp301100g | J. Phys. Chem. B 2012, 116, 6279−6287

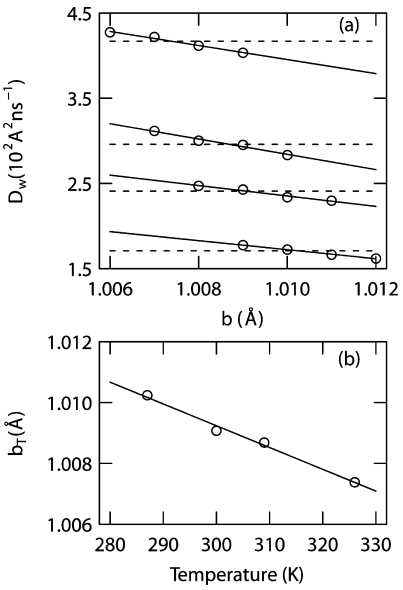


Figure 8. (a) Bond length dependence of D_w at 287, 300, 309, and 326 K (from bottom). Circles are from MD simulations. Broken lines indicate the experimental diffusion constants. Solid lines are the results of linear regression. (b) Temperature dependence of the optimized bond length. Circles were obtained from the intersections in part a. The solid line is the result of linear regression, $b_T = (-7.16 \times 10^{-5} \times T + 1.0307)$ Å.

advantageous in the study of flexible proteins with significant internal−external coupling or considerably anisotropic proteins. The combination of the SPC/E_b model and the smoothed spectral density $J_\alpha^{smooth}(\omega)$ of protein G were calculated and showed good agreement with $J_\alpha^{exp}(\omega)$. Thus, we concluded that MD simulation with the SPC/E_b water model reproduces the behavior of protein G in solution very well including anisotropy.

In this work, we focused on tuning of the water model and demonstrated improvement of the calculated NMR spectral density of a protein. Significant improvements have also been achieved for the protein force field. Li and Brüschweiler[43,44] have recently developed an NMR-based optimization to improve the ff99SB force field and reported significant improvement of the NMR chemical shifts and order parameters. The combination of their optimized force field ff99SB_$\phi\varphi$ with SPC/E_b will be useful to further improve the reproduction of $J_\alpha(\omega)$.

Finally, the temperature-dependent SPC/E_{bT} model was also developed, of which the bond length is linearly dependent on the temperature with the relation $b_T = (-7.16 \times 10^{-5} \times T + 1.0307)$ Å, where the temperature is in kelvins. This model will be useful to improve the overall tumbling correlation time and NMR spectral densities around room temperature.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Simulation time dependence of overall tumbling time of protein, radial distribution functions obtained with SPC/E_b and SPC/E_θ, site-specific spectral densities of all residues, and the effects of simulation length, statistics, and protein concentration on $J(\omega)$ reproduction. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: kitao@iam.u-tokyo.ac.jp.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This research was supported by the Strategic Programs for Innovative Research (SPIRE) of the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan, and the Computational Materials Science Initiative (CMSI), Japan, Japan Society for the Promotion of Science (JSPS) Grants-in-Aid for Science Research (B), and MEXT Grants-in-Aid for Science Research in Priority Areas to A.K. The computation in this work was conducted mainly using the facilities of the Supercomputer Center, Institute for Solid State Physics, University of Tokyo and our PC clusters, and partly with the supercomputers at the Research Center for Computational Science, Okazaki Research Facilities, National Institute of Natural Science.

■ REFERENCES

(1) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4559.
(2) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546.
(3) Nederveen, A. J.; Bonvin, A. M. J. J. *J. Chem. Theory Comput.* **2005**, *1*, 363.
(4) Showalter, S. A.; Brüschweiler, R. *J. Chem. Theory Comput.* **2007**, *3*, 961.
(5) Pfeiffer, S.; Fushman, D.; Cowburn, D. *J. Am. Chem. Soc.* **2001**, *123*, 3021.
(6) Takemura, K.; Kitao, A. *J. Phys. Chem. B* **2007**, *111*, 11870.
(7) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. In *Intermolecular Forces*; Pullman, B., Ed.; Reidel: Dordrecht, The Netherlands, 1981; p 331.
(8) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269.
(9) Wu, Y.; Tepper, H. L.; Voth, G. A. *J. Chem. Phys.* **2006**, *124*, 024503.
(10) Dunweg, B.; Kremer, K. *J. Chem. Phys.* **1993**, *99*, 6983.
(11) Yeh, I. C.; Hummer, G. *J. Phys. Chem. B* **2004**, *108*, 15873.
(12) Lynden-Bell, R. M. *J. Phys.: Condens. Matter* **2010**, *22*, 284107.
(13) Bergman, D. L.; Lynden-Bell, R. M. *Mol. Phys.* **2001**, *99*, 1011.
(14) Chatterjee, S.; Debenedetti, P. G.; Stillinger, F. H.; Lynden-Bell, R. M. *J. Chem. Phys.* **2008**, *128*, 124511.
(15) Höchtl, P.; Boresch, S.; Bitomsky, W.; Steinhauser, O. *J. Chem. Phys.* **1998**, *109*, 4927.
(16) Smith, P. E.; van Gunsteren, W. F. *J. Mol. Biol.* **1994**, *236*, 629.
(17) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. *J. Chem. Phys.* **2004**, *120*, 9665.
(18) Glättli, A.; Daura, X.; van Gunsteren, W. F. *J. Chem. Phys.* **2002**, *116*, 9811.
(19) Motakabbir, K. A.; Berkowitz, M. *J. Phys. Chem.* **1990**, *94*, 8359.
(20) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins* **2006**, *65*, 712.
(21) Tjandra, N.; Feller, S. E.; Pastor, R. W.; Bax, A. *J. Am. Chem. Soc.* **1995**, *117*, 12562.
(22) Lee, A. L.; Wand, A. J. *J. Biomol. NMR* **1999**, *13*, 101.
(23) Peng, J. W.; Wagner, G. *Biochemistry (Moscow)* **1992**, *31*, 8571.
(24) Peng, J. W.; Wagner, G. *J. Magn. Reson.* **1992**, *98*, 308.
(25) Lefèvre, J. F.; Dayie, K. T.; Peng, J. W.; Wagner, G. *Biochemistry (Moscow)* **1996**, *35*, 2674.
(26) Idiyatullin, D.; Nesmelova, I.; Daragan, V. A.; Mayo, K. H. *J. Mol. Biol.* **2003**, *325*, 149.
(27) Case, D. A.; Darden, T. A.; Cheatham, T. E., I. I. I.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Crowley, M.; Walker, R. C.; Zhang, W.; Merz, K. M.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.;

Kolossvary, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Kollman, P. A. *AMBER 10*; University of California: San Francisco, CA, 2008.
(28) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
(29) Wong, V.; Case, D. A. *J. Phys. Chem. B* **2007**, *112*, 6013.
(30) Holz, M.; Heil, S. R.; Sacco, A. *Phys. Chem. Chem. Phys.* **2000**, *2*, 4740.
(31) Vandermaarel, J. R. C.; Lankhorst, D.; Debleijser, J.; Leyte, J. C. *Chem. Phys. Lett.* **1985**, *122*, 541.
(32) Ludwig, R. *Chem. Phys.* **1995**, *195*, 329.
(33) Wagner, W.; Pruss, A. *J. Phys. Chem. Ref. Data* **2002**, *31*, 387.
(34) Kell, G. S. *J. Chem. Eng. Data* **1975**, *20*, 97.
(35) Fernandez, D. P.; Goodwin, A. R. H.; Lemmon, E. W.; Sengers, J. M. H. L.; Williams, R. C. *J. Phys. Chem. Ref. Data* **1997**, *26*, 1125.
(36) Cornilescu, G.; Marquardt, J. L.; Ottiger, M.; Bax, A. *J. Am. Chem. Soc.* **1998**, *120*, 6836.
(37) Berndt, K. D.; Guntert, P.; Orbons, L. P. M.; Wuthrich, K. *J. Mol. Biol.* **1992**, *227*, 757.
(38) Beeser, S. A.; Goldenberg, D. P.; Oas, T. G. *J. Mol. Biol.* **1997**, *269*, 154.
(39) Gronenborn, A. M.; Filpula, D. R.; Essig, N. Z.; Achari, A.; Whitlow, M.; Wingfield, P. T.; Clore, G. M. *Science* **1991**, *253*, 657.
(40) Barchi, J. J.; Grasberger, B.; Gronenborn, A. M.; Clore, G. M. *Protein Sci.* **1994**, *3*, 15.
(41) Wong, K. B.; Fersht, A. R.; Freund, S. M. *J. Mol. Biol.* **1997**, *268*, 494.
(42) Hall, J. B.; Fushman, D. *J. Biomol. NMR* **2003**, *27*, 261.
(43) Li, D. W.; Brüschweiler, R. *Angew. Chem., Int. Ed. Engl.* **2010**, *49*, 6778.
(44) Li, D. W.; Brüschweiler, R. *J. Chem. Theory Comput.* **2011**, *7*, 1773.