Rotational dynamics of proteins from spin relaxation times and molecular dynamics simulations

O. H. Samuli Ollila*
Insititute of Biotechnology, University of Helsinki and
Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, Prague 6, Czech Republic

Harri Heikkinen and Hideo Iwaï
Insititute of Biotechnology, University of Helsinki
(Dated: October 12, 2017)

Conformational fluctuations and rotational tumbling of proteins can be experimentally accessed with nuclear spin relaxation experiments. Interpretation of molecular dynamics from the experimental data is, however, often complicated, especially for molecules with anisotropic shape. Here we apply classical molecular dynamics simulations to interpret conformational fluctuations and rotational tumbling of proteins with arbitrarily anisotropic shape. The direct calculation of spin relaxation times from simulation data did not reproduce the experimental data. This was successfully corrected by scaling the overall rotational diffusion constants around protein intertia axes with a constant factor. The achieved good agreement with experiments allowed the interpretation of internal and overall dynamics of proteins with significantly anisotropic shape. The overall rotational diffusion was found to be brownian, having only a short subdiffusive region below 0.12 ns. The presented methodology can be applied to interpret rotational dynamics and conformation fluctuations of proteins with arbitrary anisotropic shape. However, for intrinsically disordered proteins a water model with more realistic dynamical properties is probably required.

I. INTRODUCTION

Conformational fluctuations and the entropy of proteins play a significant role in their functionality and interactions with other biomolecules. Conformational fluctuations and the overall brownian tumbling of proteins are experimentally accessible through the spin relaxation times of ^{15}N and ^{13}C nucleai measured with nuclaer magnetic resonance (NMR) techniques [1–7]. The spin relaxation rates have been used to, for example, analyze conformational entropies [1, 8–10], binding entropies [1, 11], resolve sampled structures [3–5, 12] and validate molecular dynamics simulations [13–17]. These analyzes are almost exclusively based on the separation of the internal conformational fluctuations and the overall rotational tumbling [18, 19]. Also the isotropic overall diffusion is often assumed, while analysis of anisotropic molecules is significantly more complicated [1, 2, 20–23]. Thus, the new approaches are needed to intepret spin relaxation times measured from anisotropic or intrinsically disordered molecules.

Classical molecular dynamics simulation methods are promising tools to interpretate spin relaxation experiments for molecules with significantly anisotropic shape or correlations between internal and overall rotational motions. Practical applications are, however, limited by inaccuracies in the force field descriptions and the available time scales in the simulations [16, 17, 24–26]. The main issues have been the overestimated overall rotational diffusion of proteins due to inaccuracies in water models [25] and the insufficient accuracy of correlation functions calculated from single molecules in MD simulations [26, 27].

In this work we overcome these issues by assuming that the overall rotational dynamics of protein follows anisotropic rigid body diffusion. Diffusion coefficients around inertia axes are directly calculated from angular displacements. The diffusion coefficients are then used to determine the contribution of the overall rotational tumbling to the rotational correlation functions of N-H bonds. This reduces the required simulation length for the accurate determination of rotational correlation functions. Furthermore, the overestimated overall brownian tumbling rates due to the inaccurate water model can be corrected during the correlation function calculation by scaling the diffusion coefficients in all directions with a constant factor. The corrected correlation functions can be used to interpret the spin relaxation experiments for proteins with arbitrarily anisotropic shapes.

The developed approach is demonstrated by interpreting the experimental spin relaxation data for C-terminal domains of TonB proteins from *H. pyroli* (*Hp*TonB) [28] and from *Pseudomonas aeruginosa* (*Pa*TonB). Both proteins have significantly anisotropic shape, which would complicate the standard spin relaxation data analysis [1, 2, 20–23].

II. METHODS

A. Spin relaxation experiments and rotational dynamics of molecules

Molecular dynamics of the protein backbone residues and spin relaxation experiments can be connected by using the spectral density $J(\omega)$

$$J(\omega) = 2\int_0^\infty C(t)\cos(\omega t)\mathrm{d}t, \qquad (1)$$

* samuli.ollila@helsinki.fi

which is the Fourier transformation of the second order rotational correlation function for N-H bond vector

$$C(t) = \langle \frac{3}{2}\cos^2\theta_{t'+t} - \frac{1}{2}\rangle_{t'}, \tag{2}$$

where $\theta_{t'+t}$ is the N-H bond angle between times t' and $t'+t$ and angular brackets refer to the ensemble average. Connection to the experimentally measured spin relaxation times T_1, T_2 and T_{NOE} is given by the Redfield equations [29, 30]

$$\frac{1}{T_1} = \frac{d_{\mathrm{NH}}^2 N_{\mathrm{H}}}{20}\left[J(\omega_{\mathrm{H}}-\omega_{\mathrm{N}}) + 3J(\omega_{\mathrm{N}}) + 6J(\omega_{\mathrm{N}}+\omega_{\mathrm{H}})\right] + \frac{(\sigma\omega_{\mathrm{N}})^2}{15}J(\omega_{\mathrm{N}}), \tag{3}$$

$$\frac{1}{T_2} = \frac{1}{2}\frac{d_{\mathrm{NH}}^2 N_{\mathrm{H}}}{20}\left[4J(0) + 3J(\omega_{\mathrm{N}}) + J(\omega_{\mathrm{H}}-\omega_{\mathrm{N}}) + 6J(\omega_{\mathrm{H}}) + 6J(\omega_{\mathrm{N}}+\omega_{\mathrm{H}})\right] + \frac{(\sigma\omega_{\mathrm{N}})^2}{90}[4J(0)+3J(\omega_{\mathrm{N}})], \tag{4}$$

$$\frac{1}{T_{\mathrm{NOE}}} = 1 + \frac{d_{\mathrm{NH}}^2 N_{\mathrm{H}}}{20}\left[6J(\omega_{\mathrm{N}}+\omega_{\mathrm{H}}) + J(\omega_{\mathrm{H}}-\omega_{\mathrm{N}}))\right]\frac{\gamma_{\mathrm{H}}T_1}{\gamma_{\mathrm{N}}}, \tag{5}$$

where ω_{N} and ω_{H} are the Larmor angular frequencies of ^{15}N and ^{1}H respectively, and the number of bound protons $N_{\mathrm{H}} = 1$ for N-H bonds. The dipolar coupling constant is given by

$$d_{\mathrm{NH}} = -\frac{\mu_0 \hbar \gamma_{\mathrm{H}}\gamma_{\mathrm{N}}}{4\pi\langle r_{\mathrm{NH}}^3\rangle},$$

where μ_0 is the magnetic constant or vacuum permeability, $\hbar$ is the reduced Planck constant, γ_{N} and γ_{H} are the gyromagnetic constants of ^{15}N and ^{1}H, respectively. The average cubic length is calculated as $\langle r_{\mathrm{NH}}^3\rangle = (0.101\mathrm{nm})^3$ and the value of $\Delta\sigma = -160$ ppm is used for the chemical shift anisotropy of N-H bonds in proteins [30, 31].

Spin relaxation experiments are typically interpreted for proteins by assuming that the motions related to the overall brownian tumbling and conformational fluctuations are independent. The rotational correlation function for chemical bonds can be then written as [1, 2, 18, 19, 32]

$$C(t) = C_I(t)C_O(t), \tag{6}$$

where $C_I(t)$ and $C_O(t)$ are correlation functions for internal and overall rotations, respectively. Conformational fluctuations can be described in this approximation by using the square of the order parameter respect to molecular axes S^2, which is given by the plateau of the internal rotational correlation function. Timescales for the fluctuations can be characterized by using the effective correlation time

$$\tau_{\mathrm{eff}} = \int_0^\infty C_I'(t)\mathrm{d}t, \tag{7}$$

where $C_I'(t) = \frac{C_I - S^2}{1-S^2}$ is the reduced correlation function [19].

The overall rotational correlation function is often described by approximating the protein as a rigid body. For arbitrarily anisotropic molecules, the correlation functions can be presented as a sum of five exponentials [2, 20]

$$C_O(t) = \sum_{j=1}^{5} A_j e^{-t/\tau_j}, \tag{8}$$

where the time constants τ_j are related to the diffusion constants around three principal axes of a molecule (D_{xx}, D_{yy} and D_{zz}) [33] and the prefactors A_j depend on the directions of chemical bonds respect to the molecular axes [20, 22].

The simplest approach to extract molecular dynamics from the experimental data is the original "model free analysis" [19], where an isotropic diffusion is assumed for the overall rotation of the protein. This reduces Eq. 8 to a mono-exponential form and the overall rotational dynamics can be described with a single time constant τ_c. Also the internal correlation functions for each residue are assumed to decay exponentially with a single time constant τ_{eff} toward to the square of the order parameter S^2. The three parameters (τ_c, τ_{eff} and S^2) can be then successfully resolved from a fit to the experimental data. However, the amount of parameters to be fitted increases if the protein experiences an anisotropic overall diffusion or has several timescales for internal motions. In this case the fitting becomes often ambiguous, even if the experimental data would be measured with multiple magnetic field strengths [1, 22, 34]. The anitropic rotational diffusion is sometimes described with hydrodynamical calculations, but they are quite sentitive for the assumptions about the hydration shell of the protein [35].

Rough estimate for the timescale of overall rotational dynamics is often given by using the T_1/T_2 ratio [30]. This is based on the assumptions that T_1 and T_2 are essentially independent of the internal motions and that the overall dynamics is isotropic. The spectral density then reduces to

$$J'(\omega) = S^2\frac{\tau_c'}{1+(\omega\tau_c')^2} \tag{9}$$

and the correlation time describing the overall rotational motion, τ_c', can be then estimated by numerically minimizing the equation

$$\frac{T_1}{T_2} \approx \frac{\frac{1}{2}\frac{d_{\mathrm{NH}}^2 N_{\mathrm{H}}}{20}\left[4J'(0) + 3J'(\omega_{\mathrm{N}}) + J'(\omega_{\mathrm{H}}-\omega_{\mathrm{N}}) + 6J'(\omega_{\mathrm{H}}) + 6J'(\omega_{\mathrm{N}}+\omega_{\mathrm{H}})\right] + \frac{(\sigma\omega_{\mathrm{N}})^2}{90}[4J'(0)+3J'(\omega_{\mathrm{N}})]}{\frac{d_{\mathrm{NH}}^2 N_{\mathrm{H}}}{20}\left[J'(\omega_{\mathrm{H}}-\omega_{\mathrm{N}}) + 3J'(\omega_{\mathrm{N}}) + 6J'(\omega_{\mathrm{N}}+\omega_{\mathrm{H}})\right] + \frac{(\sigma\omega_{\mathrm{N}})^2}{15}J'(\omega_{\mathrm{N}})} \tag{10}$$

with respect to the experimentally measured T_1/T_2 ratio.

B. Rotational dynamics from molecular dynamics simulations

A classical molecular dynamics simulation gives a trajectory for each atom in the system as a function of time. Rotational correlation functions for each bond can be then directly calculated from the trajectories by using Eq. 2 and used to calculate the spin relaxation times through Eqs. 1-5. The resulting values can be compared to experimental data in order to assess simulation model quality [13–17, 36] or to interpret experiments [36].

The direct comparison with experiments is, however, often complicated by insufficient statistics for the calculated correlation functions and the overestimated rotational diffusion due to inaccuracies in water models [25, 26]. Here we show that the statisitical accuracy of the contribution of the overall tumbling to the correlation functions, $C_0(t)$ in Eq. 6, can be increased for rigid proteins by directly calculating the diffusion coefficients for intertia axes. The rotational diffusion coefficients can be related to timescales in the correlation function for anisotropic rigid body rotation in Eq. 8 [20, 33].

The rotational diffusion coefficients are calculated by fitting a linear slope to the square angle deviation of intertia axes (see below). This requires less simulation data for the good statistics than a direct fit of multiexponential sum in Eq. 8 to the rotational correlation function calculated from MD simulation. In addition, the overestimated rotational diffusion due to water model [25] can be corrected by scaling the diffusion coefficients around all intertia axes by a constant factor. This approach takes into account the anisotoropic shape of the molecule. This is a significant advancement to the studies, which assume isotropic rotational diffusion with a single exponential rotational correlation function [10, 14–16, 37] or use order parameters to compare simulations with experimental data [13, 16, 17, 37].

The practical analysis can be divided into seven steps:
1) The total rotational correlation functions $C(t)$ for the protein N-H bond vectors are directly calculated from MD simulation trajectory by applying Eq. 2.
2) The rotational correlation functions for internal dynamics $C_I(t)$ are calculated from MD simulation trajectory by removing overall rotation of the protein.
3) The overall and internal motions are assumed to be independent and the overall rotational correlation function is calculated from Eq. 6 as $C_O(t) = C(t)/C_I(t)$.
4) The mean square angle deviations of rotation around protein intertia axes are calculated from MD simulation trajectory.
5) Rotational diffusion constants D_x, D_y and D_z are calculated by fitting a straight line to the mean square angle deviations

$$\begin{aligned}\langle\Delta\alpha^2_{t'+t}\rangle_{t'} &= 2D_x t\\ \langle\Delta\beta^2_{t'+t}\rangle_{t'} &= 2D_y t\\ \langle\Delta\gamma^2_{t'+t}\rangle_{t'} &= 2D_z t,\end{aligned} \tag{11}$$

where $\langle\Delta\alpha^2_{t'+t}\rangle_{t'}$, $\langle\Delta\beta^2_{t'+t}\rangle_{t'}$ and $\langle\Delta\gamma^2_{t'+t}\rangle_{t'}$ are the mean square angle deviations from the shortest protein intertia axis to the longest, respectively.
6) Contribution of the overall rotational tumbling to all correlation functions is assumed to follow Eq. 8 with timescales τ_j calculated from the rotational diffusion constants [33]. Weighting factors A_j are determined by fitting the equation with new timescales to the overall rotational correlation functions calculated from MD simulations in step 3.
7) The new correlation functions are calculated by substituting internal correlation functions, $C_I(t)$, from step 2 and anisotropic rigid body rotational correlation functions, $C_O(t)$, from step 6 to Eq. 6 giving

$$C_N(t) = C_I(t)\sum_{j=1}^{5} A_j e^{-t/\tau_j}. \tag{12}$$

These correlation functions are then used to calculate spin relaxation times from Eqs. 1-5. The incorrect overall rotational diffusion due to water model can be corrected at this point by scaling the rotational diffusion coefficients, i.e. timescales τ_j, with a constant factor before calculating new correlation functions from Eq. 12.

C. Simulation and analysis details

All simulations were ran using Gromacs 5 [38] and Amber ff99SB- ILDN [39] force field for proteins . The proteins were solvated to tip3p[40], tip4p [40] or OPC4 [41] water models. Initial structures were taken from NMR structures of *Hp*TonB [28] and *Pa*TonB [**?**]. The temperature was coupled to the desired value with v-rescale thermostat [42] and the pressure was isotropically set to 1 bar using Parrinello-Rahman barostat [43]. Timestep was 2 fs, Lennart-Jones interactions were cut-off at 1.0 nm, PME [44, 45] was used for electrostatics and LINCS was used to constraint all bond lengths [46]. Simulation trajectories and the related files are available at [**?**]. The simulated systems are listed in Table I

The rotational correlation functions are calculated with *gmx rotacf* from Gromacs package [47]. The overall rotation was removed for $C_I(t)$ calculation by using a fit option of the *gmx trjconv* tool in Gromacs package [47]. The order parameters S^2 were determined by averaging rotational correlation functions from oriented trajectory, $C_I(t)$, over the lag times above 50 ns. The effective correlation times were then calculated by using Eq. 7. Inertia axes of proteins were calculated with *compute_inertia_tensor* function from MDTraj python library [48].

Spectral density was calculated by fitting a sum of 471 exponentials with timescales from 1 ps to 50 ns with logarithmic spacing

$$C_N(t) = \sum_{i=1}^{N} \alpha_i e^{-t/\tau_i} \tag{13}$$

to the new correlation function from Eq. 12 by using the

TABLE I. Simulated systems and rotational diffusion coefficients (rad$^2 \cdot 10^7$/s) calculated from simulations.

Protein	Water model	T (K) [a]	$t_{\rm sim}$ (ns) [b]	$t_{\rm anal}$ (ns) [c]	D_{xx}	D_{yy}	D_{zz}	$D_{\|\|}/D_\perp$ [d]	$D_{\rm av}$ [e]	files
*Pa*TonB	tip4p	298	400	390	1.81 ± 0.01	2.06 ± 0.03	4.55 ± 0.03	2.35 ± 0.04	2.80 ± 0.02	[**?**]
*Pa*TonB	tip4p	310	400	390	2.60 ± 0.02	2.22 ± 0.05	5.0 ± 0.1	2.07 ± 0.09	3.26 ± 0.07	[**?**]
*Pa*TonB	OPC4	310	1200	1190	2.01 ± 0.01	2.19 ± 0.01	5.01 ± 0.03	2.39 ± 0.02	3.07 ± 0.01	[**?**]
*Hp*TonB	tip3p	310	570	370	8.25 ± 0.05	7.67 ± 0.06	15.9 ± 0.3	1.99 ± 0.06	10.6 ± 0.2	[**?**]
*Hp*TonB	tip3p	303	800	790	6.24 ± 0.02	7.04 ± 0.03	11.9 ± 0.2	1.80 ± 0.03	8.40 ± 0.07	[**?**]
*Hp*TonB	tip4p	310	470	370	3.6 ± 0.1	3.24 ± 0.01	6.3 ± 0.3	1.8 ± 0.1	4.4 ± 0.2	[**?**]
*Hp*TonB	tip4p	303	400	200	2.7 ± 0.1	2.71 ± 0.02	5.6 ± 0.5	2.1 ± 0.2	3.7 ± 0.2	[**?**]
*Hp*TonB	OPC4	310	800	790	2.85 ± 0.01	2.70 ± 0.01	5.56 ± 0.01	2.00 ± 0.01	3.70 ± 0.01	[**?**]

[a] Simulation temperature
[b] Total simulation time
[c] Analyzed simulation time
[d] $D_{||} = D_{zz}$, $D_\perp = \frac{1}{2}(D_{xx} + D_{yy})$
[e] $D_{\rm av} = \frac{1}{3}(D_{xx} + D_{yy} + D_{zz})$

lsqnonneg routine in MATLAB [49]. The Fourier transform was then calculated by using the analytical function for the sum of exponentials

$$J(\omega) = 4 \sum_{i=1}^{N} \alpha_i \frac{\tau_i}{1 + \omega^2 \tau_i^2}. \qquad (14)$$

A similar approach has been previously used for the lamellar lipid and surfactant systems in combination with solid state NMR experiements [50, 51]. All the computer programs used for the analysis are available at [**?**].

D. Spin relaxation experiments

NMR experiments were recorded on a Bruker Avance III 850 NMR spectrometer equipped with cryogenic probe head. The longitudinal (T_1), transverse (T_2), and ^{1}H-^{15}N-heteronuclear NOE ($T_{\rm NOE}$) spin relaxation times for the backbone ^{15}N atoms of *Hp*TonB [28] were collected at 303 K using the well-established NMR pulse sequences described previously [30, 52]. The similarly detected spin relaxation data for *Pa*TonB at 298 K is reported in a another publication [**?**]. The T_1 and T_2 relaxation times were measured using the following delay times: 10, 50, 100, 200, 300, 500, 800, 1000, 1200 and 2000 ms, and 16, 64, 96, 128, 156, 196, 224 and 256 ms for CPMG pulse train with 1.0 ms interval for T_2 relaxation times, respectively. The relaxation rates ($R_1 = 1/T_1$, $R_2 = 1/T_2$) were calculated as an exponential fit of a single exponential decay to peak intensity values: $I(t) = I_0 \exp(-t/T_1)$, or $I(t) = I_0 \exp(-t/T_2)$, where $I(t)$ is the peak volume at a time t. The ^{15}N$\{^1$H$\}$-NOE measurements were carried out with a relaxation delay of 5 s with and without saturation of the amide protons. The ^{15}N$\{^1$H$\}$-NOE values were derived from the volumes of the HSQC peaks using the equation of $\nu = I/I_0$. The relaxation data was processed and analysed using Bruker Dynamic Center software (version 2.1.8).

III. RESULTS

A. Global rotational dynamics of protein

The mean square angle deviations for the rotation of *Pa*TonB protein construct around inertia axes in the simulation with OPC4 water model are shown in Fig. 1. This is the longest simulation data set in this work (1.2μs) and the linear behaviour of the mean square angle deviations are observed for the lag times up to one hundreth of the total simulation length (12 ns), which is expected to be the maximum lag time for the good statistics of rotational dynamics analyzed from a single molecule in MD simulations [27]. Deviations from the linear behaviour are only seen with the lag times longer than this limit, as also demonstrated for the shorter simulations in Figs. 8 and 9 with tip4p water at two different temperatures. The plots with log-log scale in Figs. 1, 8 and 9 B) reveals a weakly subdiffusive region only below very short timescales of approximately 0.12 ns. Thus, we conclude that the protein experiences the brownian rotational tumbling with a good approximation. The diffusion coefficients can be then calculated from the slope of the mean square angle deviations according to Eq. 11 by using the lag times less than one hundreth of the total MD simulation length. The error bars were calculated varying the lag time with 1 ns to both directions. The data from *Hp*TonB protein (not shown) led to similar conclusions.

The resulting rotational diffusion constants from different simulations are shown in Table I. As expected, the rotational diffusion coefficients increase with the temperature and the decreasing size of a protein. The values are, however, larger than would be expected from the experimental T_2/T_1 ratio analysed with Eq. 10 and from the previosly reported values for proteins with similar sizes [53], especially when tip3p water model is used. Similar results were previously explained by the overestimated water self-diffusion of tip3p water model [25].

The analysis leading to the new correlation functions in Eq. 12 (see section II B) is exemplified in Fig. 2 for three

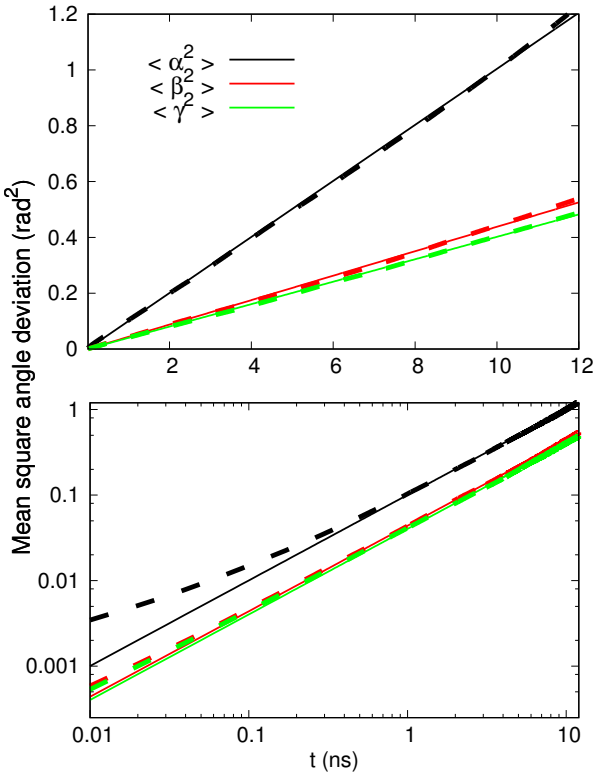

FIG. 1. Mean square angle deviations of intertia tensor axes calculated from *Pa*TonB simulation with OPC water model. The data shown with linear (top) and logarithmic scale (bottom).

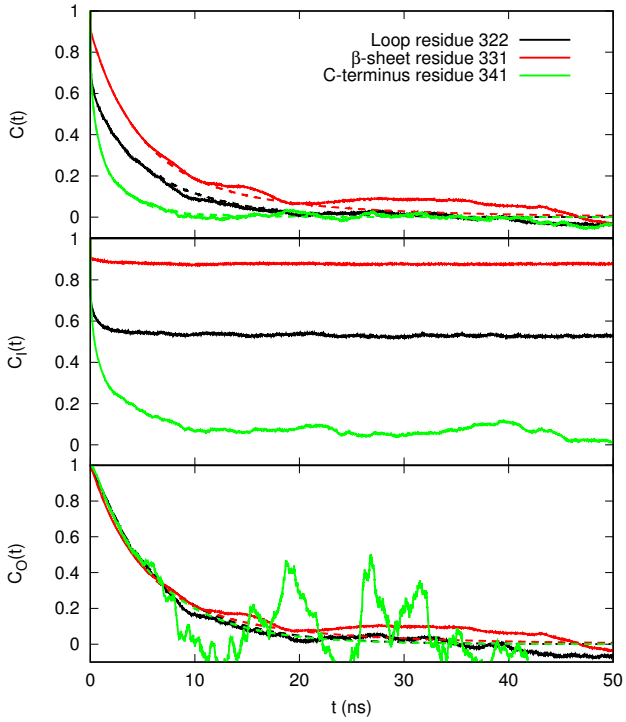

FIG. 2. Rotational correlation functions calculated from MD simulations of *Pa*TonB with tip4p water model at 298K for residues at different regions. (top) total correlation functions $C(t)$ calculated from MD simulation (solid lines) and new correlation functions determined from Eqs. 6 and 8 by using rotational diffusion constants and fitted prefactors (dashed lines), (middle) correlation functions for internal motions calculated from simulation with removed overall protein rotation (bottom) correlation function for overall motions determined as $C_O(t) = C(t)/C_I(t)$ (solid lines) and by fitting to Eq. 8 with timescales from rotational diffusion coefficients in Table I (dashed lines).

residues locating in different domains of *Pa*TonB protein with different characteristic rotational dynamics. The flexible C-terminus is represented by the residue 341, more rigid β-sheet by the residue 331 and a flexible loop between two sheets by the residue 322 (see the labeling in Fig. 6). The total correlation functions $C(t)$ of all the residues in Fig. 2 (top, solid lines). decay toward zero within ∼10-50 ns. The internal correlation functions $C_I(t)$ in Fig. 2 (middle) decay to a plateau value, which defines the square of the order parameter S^2. As expected, the internal correlation function for the residue 331 in the rigid β-sheet rapidly decays to the largest order parameter value, while the correlation functions of the residues in the loop and C-terminus decay slower to the smaller order parameter values due to the larger conformational ensemble sampled by these regions.

The overall rotational correlation functions, $C_O(t) = C(t)/C_I(t)$, are shown in Fig. 2 (bottom, solid lines). Also the correlation functions of anisotropic rigid body rotation from Eq. 8 are shown in Fig. 2 (bottom, dashed lines). The timescales for the latter, τ_i, are given by the rotational diffusion coefficients from the simulation and the relation in Ref. 33, and the prefactors, A_j, are determined by fitting Eq. 8 to the overall rotation correlation functions, $C_O(t)$, calculated the MD simulation. The new correlation functions, determined from Eq. 12 and shown in Fig. 2 (top, dashed lines), are indistinguishable from the correlation functions calculated from the original MD simulations with the lag times shorter than one hundredth of total simulation time (approximately 4-12 ns for the studied systems), which is the expected limit for the good statistics in single molecule MD simulations [27]. This suggests that the anisotropic rigid body diffusion model (Eq. 8) and the separation of internal and global motions (Eq. 6) are good approximations for the studied system. The analytical description of the overall rotation with Eq. 8 in the new correlation functions clearly reduces the statistical fluctuations with the long lag times in Fig. 2. The effect is most visible for the flexible C-terminus (residue 341) having the smallest, thus the least detectable, contribution from the overall rotation of the protein due to the small order parameters values.

B. Global rotational dynamics in simulations and experiments

The spin relaxation times of the *Hp*TonB construct are compared between the experiments and simulations using two different water models in Fig. 3. Simulations with tip3p water model are significantly off from the experimental data and un-

derestimate the T_1/T_2 ratio, suggesting too fast overall rotational diffusion dynamics [54]. This is in agreement with the previous study, where the overestimated rotational diffusion was attributed to the self-diffusion of tip3p [25]. On the other hand, simulation results with tip4p water model show significantly better agreement with the experimental data in Fig. 3.

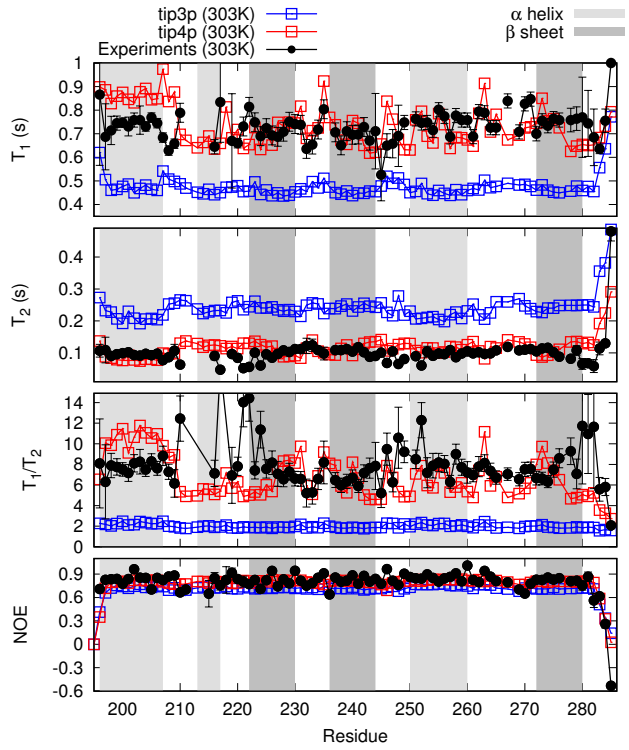

FIG. 3. ^{15}N Spin relaxation times for *Hp*TonB from experimental data (circles) and MD simulations with different water models (squares).

To see if the discrepancy in spin relaxation times for simulations with tip3p water model could be explained by the overestimated overall diffusion of the protein, the diffusion coefficients were divided with a constant factor of 2.9 before applying Eq. 12 to calculate the new correlation functions. The spin relaxation times calculated from the new correlation functions after scaling the rotational diffusion coefficients are, indeed, in good agreement with the experimental data in Fig. 4

Similar comparison for the spin relaxation times of the *Pa*TonB construct between experiments and simulations with tip4p and OPC4 water models is shown in Fig. 5. The experimentation of the OPC4 water model was inspired by the recent study reporting significant improvements in lipid monolayer simulations when this water model was used [57]. Since tip3p water model was observed to significantly overestimate the rotational diffusion of protein, it was not used in *Pa*TonB simulations. The underestimation of T_1/T_2 ratio was also observed in simulations of *Pa*TonB with tip4p and OPC4 water models when compared with experiments. The discrepancy is, however, less severe than for *Hp*TonB simulated with tip3p in Fig. 3, suggesting that also the required correction

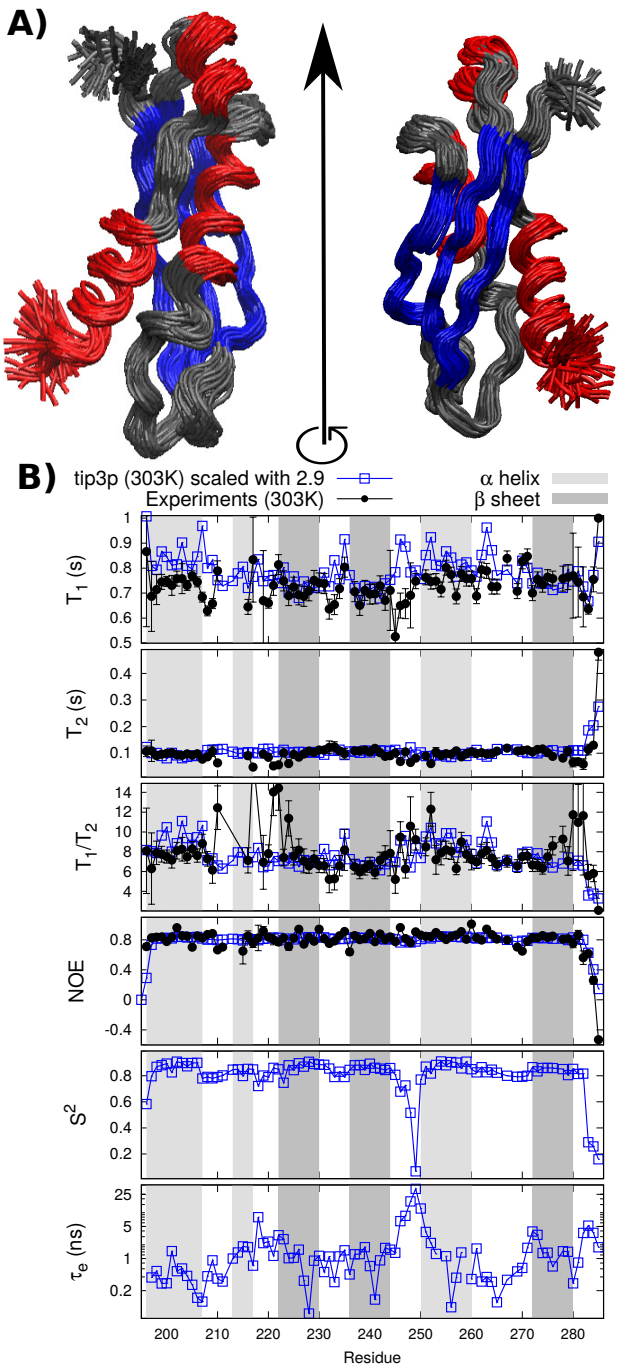

FIG. 4. A) Structures of *Hp*TonB from the MD simulations with tip3p at 303 K (100 structures taken from 400ns long trajectory). Secondary structures are colour-labelled with Visual Molecular dynamics [55, 56]; α-helices are red and β-sheets are blue. B) Spin relaxation times from experiments (circles) and tip3p simulations (squares) with rotational diffusion coefficients divided by a constant factor of 2.9 at 303 K. Order parameters and effective internal correlation times calculated from simulations

for the overall rotational diffusion should be smaller. Indeed, the spin relaxation times calculated from *Pa*TonB simulation with tip4p water model were in good agreement with experiments in Fig. 6 after the diffusion coefficients were divided with a constant factor of 1.2, which is smaller than 2.9 used for *Hp*TonB simulation above. Notably, the temperature difference of 12 degrees has significantly smaller effect on spin relaxation times than the scaling of the diffusion coefficient or the differences between simulations and experiments.

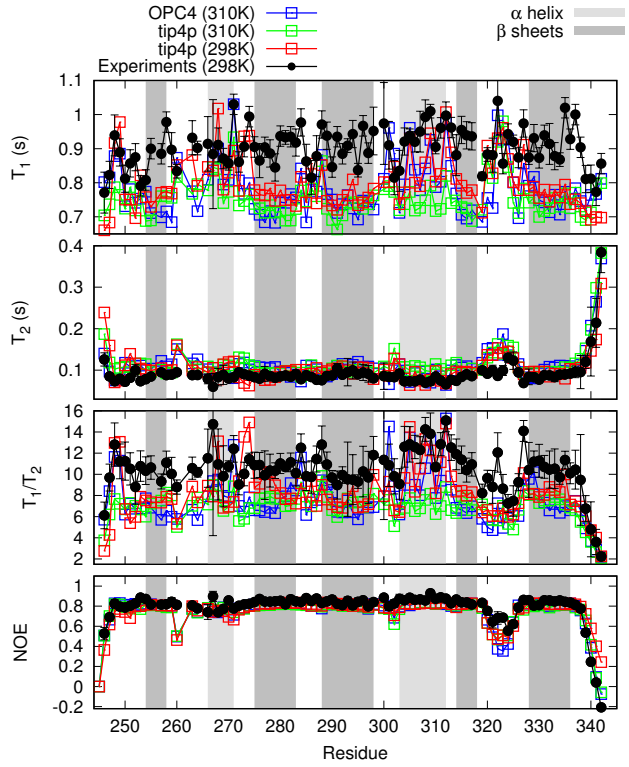

FIG. 5. Plots of experimental (circles) and simulated (squares) spin relaxation times for *Pa*TonB.

The correction of the overall rotational diffusion coefficients with a constant factor led to a good agreement with experimental spin relaxation data for both studied proteins simulated with different water models, as seen in Figs. 4 and 6. The good agreement with experiments suggests that the scaled rotational diffusion coefficients from MD simulations can be considered as a interpretation of the anisotropic rotational motion in NMR experiments. The scaled rotational diffusion coefficients from simulations giving the best agreement with the experimental data are collected in Table II. In contrast to the unscaled diffusion constants in Table I, these results are in line with the previously reported values for proteins with similar sizes [53]. Also the timescales, τ_c', estimated from Eq. 10 are close to the average diffusion coefficient, $\tau_c = (6D_{av})^{-1}$, in Table I.

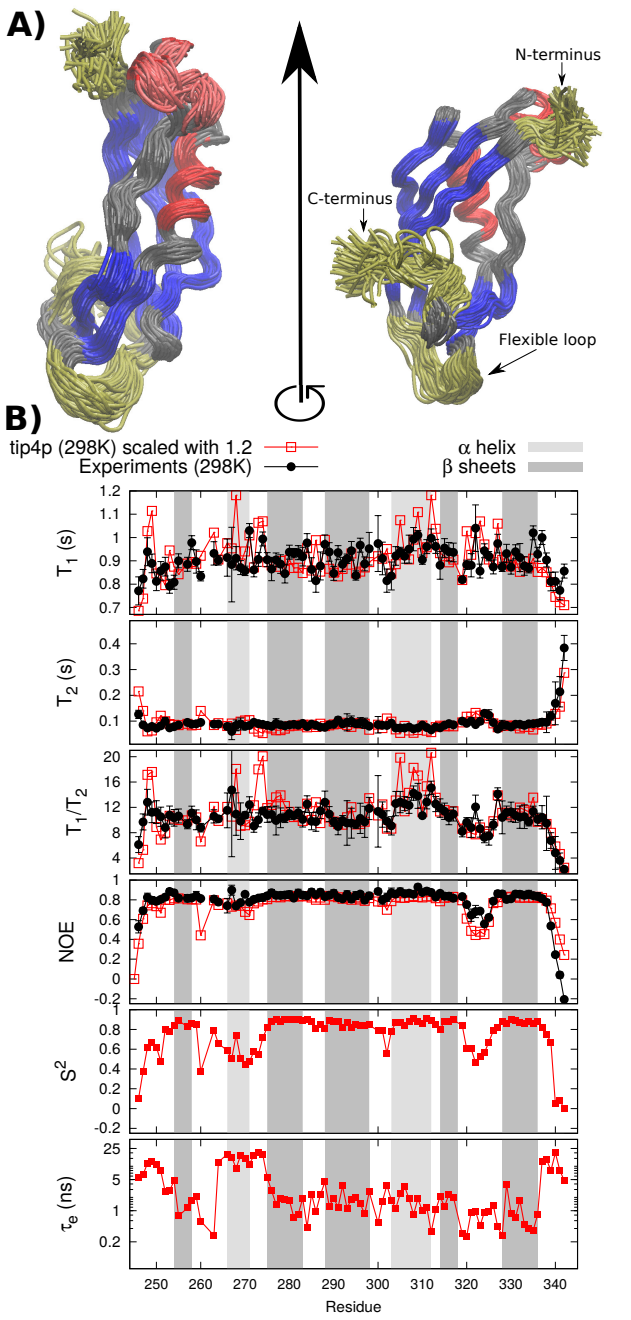

FIG. 6. A) Structures sampled by *Pa*TonB from MD simulations with tip4p at 298 K (100 structures from 400ns long trajectory). Secondary structures are colour labelled with Visual Molecular dynamics [55, 56]; α-helixes are red and β-sheets are blue. Residues 246-251, 320-326 and 338-342 with increased internal dynamics are yellow and α-helix fluctuations between two orientations (residues 266-270) is pink in the left column. B) Spin relaxation times from experiments (circles) and tip4p simulations (squares) with rotational diffusion coefficients divided by a constant factor of 1.2 at 298 K. Order parameters and effective internal correlation times calculated from simulations.

TABLE II. Rotational diffusion coefficients giving the best agreement with experimental spin relaxation data. For *Hp*TonB construct the values calculated from simulation with tip3p were scaled with 2.9 (spin relaxation data in Fig. 4) and for *Pa*TonB the values from tip4p simulation at 298K were scaled by 1.2 (spin relaxation data in Fig. 6).

	*Hp*TonB	*Pa*TonB
D_{xx}	2.15 ± 0.01	1.51 ± 0.01
D_{yy}	2.43 ± 0.01	1.72 ± 0.03
D_{zz}	4.10 ± 0.01	3.79 ± 0.03
D_{av}	2.90 ± 0.03	2.3 ± 0.02
τ_c(ns)[a]	5.7 ± 0.1	7.2 ± 0.1
τ_c'(ns)[b]	5.8 ± 0.1	6.9 ± 0.1

[a] $(6D_{av})^{-1}$
[b] Average over all residues given by Eq. 10

C. Interpretation of protein internal relaxation from MD simulations

The good agreement of the spin relaxation times between simulations with the scaled overall rotational diffusion coefficients and experiments (Figs. 4 and 6) suggest that the simulations can be used to interpret the internal dynamics of proteins from the experimental data.

Only small variations between different residues are observed for spin relaxation times of *Hp*TonB construct in Fig. 3. This indicates a rather rigid protein structure, which is also seen in MD simulation snapshots overlayed in Fig. 3 A). Only few residues in the terminal ends show slightly enhanced conformational fluctuations in the MD simulation and in spin relaxation data. In addition, some deviations from the average spin relaxation times are observed in the experiemental data close to the residues 210-222, where difficulties in the peak assignment was also observed [28]. Simulations of *Hp*TonB construct do not offer any explanation for this observation, however, similar region in *Pa*TonB simulation shows fluctuations between two orientations of α-helix (see discussion below). Exceptionally low order parameters and long effective correlation times are observed in simulations for the residues 245-250 of *Hp*TonB. Also small T_1 times are experimentally observed close to this region, but the interpretation is not straightforward as the low T_1 values are not reproduced by MD simulations.

More variety in the internal dynamics between residues is observed for *Pa*TonB protein and the segments with enhanced conformational fluctuations are labelled with yellow colour in Fig. 6. Larger number of sampled conformations in both terminal ends are characterized by low order parameters and long effective internal correlation times observed in simulations. Enhanced conformational fluctuations are also observed for residues between 320-326, which correspond the loop between two β-sheets. MD simulations predict low order parameters and long internal effective correlation times also for residues between 260-274, which can be explained by two different orientations sampled by the α-helix in this region (colour labelled with pink in Fig. 6 A). The orientational fluctuations of the similar short helix could also explain the above mentioned lower spectral resolution and deviations of spin relaxation times for residues 210-222 of *Hp*TonB [28].

MD simulations can be used to suggest how the rotational dynamics of the individual N-H bonds builds up from different components. In this work we have fitted a sum of 471 different timescales to the correlation functions according to Eq. 13. Most of the prefactor values are zero in all the correlation functions, thus the remaining timescales with non-zero prefactors are considered as the different components of the total relaxation process. The prefactors are shown in Fig. 7 for the same residues of *Pa*TonB, which were used to exemplify the correlation functions in Fig. 2. As expected for the residue 322 in the rigid β-sheet with large order parameter value, the rotational relaxation is dominated by timescales of $\sim$5.5 ns and $\sim$8 ns, matching with the protein overall rotation. Also the dynamics of the residue 322 in the flexible loop is dominated by the timescales around $\sim$8 ns corresponding to the protein overall rotation, however, the fast motions from internal dynamics are more significant than for the rigid β-sheet residue. This is in agreement with lower order parameter value in the flexible loop residues. On the other hand, the rotational dynamics of the residue 341 in the flexible N-terminus is dominated by timescales below 3 ns, most likely related to the internal dynamics of the protein. Contributions from timescales around $\sim$13 ns to the dynamics of the residue 341 probably arise from the slow conformational fluctuations of the N-terminus, rather than the overall rotational dynamics. This supports the conclusion that the large amount of sampled conformations lead to the small order parameters and large effective correlation times observed in Fig. 5. While fractionalization of the rotational dynamics to different components gives intuitively understandable results, it should be kept in mind that this is based on the fitting of a multiexponential sum to the simulation data and the solution of such fit is not unique.

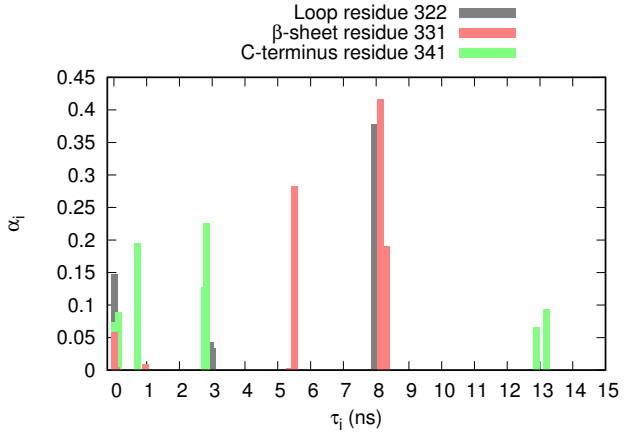

FIG. 7. Prefactors α_i corresponding different timescales τ_i resulting from a fit of Eq.13 to correlation functions from MD simulation of *Pa*TonB at 298K. The used correlation functions give a good agreement with experimental spin relaxation times as shown in Fig. 6.

IV. DISCUSSION AND CONCLUSIONS

The experimental spin relaxation data for the protein backbone N-H bonds was successfully reproduced by using the classical molecular dynamics simulations for two different proteins. Thus, the simulation trajectories give an atomic resolution interpretation for the protein dynamics measured with NMR experiments. Interpretation of the overall and internal dynamics was demonstrated for two proteins with anisotropic molecular shape and some flexible domains. Interpretation of the spin relaxation data measured from such proteins has been very challenging with the previously available methods [22, 52].

The overall rotation of the studied proteins was found to be brownian having only a small subdiffusive behaviour with short timescales below $\sim$0.12 ns, which could be constrasted with crowded environments, where anomalous diffusion is expected to be more significant [58]. The direct analysis of classical molecular dynamics trajectories did not, however, reproduce the experimental spin relaxation data. Comparison between the rotational diffusion coefficients and spin relaxation times between simulations and and experiments suggested that the overall brownian tumbling of proteins is too rapid in simulations, in agreement with the previous study suggesting that the discrepance arises from the inaccuracies in water models [25]. Scaling down the anisotropic diffusion coefficients in simulation data led to the good agreement with experiments.

The overall rotational diffusion coefficients were overestimated by a factor of $\sim$3 in *Hp*TonB simulations with tip3p water model, in agreement with previous studies [24–26]. Simulations with tip4p and opc4 water models gave the spin relaxation times in reasonable agreement with experiments with scaling factors of $\sim$1-1.2, which is significantly less than for tip3p. It should be noted, however, that the correct scaling factors for the two different proteins, *Hp*TonB and *Pa*TonB, simulated with the same water model, tip4p, seems to be 1 and 1.2, respectively. This suggests that the correction factor is not fully determined by the bulk water properties and that the the hydration layer and water-protein interactions play a significant role in rotational dynamics of proteins, as also suggested by the hydrodynamical calculations [35].

Similarity between the correlation functions from the original MD trajectory and the new correlation functions from Eq. 12 suggest that the usage of the intertia axes and the separation of internal and the overall rotational motions (Eq. 6) are good approximations for the investigated proteins, which is in agreement with previous studies of similar proteins [10, 25]. However, it remains to be seen how well this and other related approaches [24, 26] will succeed for intrinsically disordered proteins without the well defined shape. Since the correction of the incorrect overall rotational diffusion due to water model may become higly complicated for such proteins, it may be necessary to employ a water model giving correct overall rotational diffusion coefficients for biomolecules.

ACKNOWLEDGMENTS

We acknowledge CSC-IT center for science for computational resources Academy of Finland (#) and N. N. are acknowledged for the financial support to complete this work. The Finnish Biological NMR Center is supported by Biocenter Finland.

[1] V. A. Jarymowycz and M. J. Stone, Chemical Reviews **106**, 1624 (2006).
[2] D. Korzhnev, M. Billeter, A. Arseniev, and V. Orekhov, Progress in Nuclear Magnetic Resonance Spectroscopy **38**, 197 (2001).
[3] F. A. Mulder, A. Mittermaier, B. Hon, F. W. Dahlquist, and L. E. Kay, Nat Struct Mol Biol **8**, 932 (2001).
[4] E. Z. Eisenmesser, O. Millet, W. Labeikovsky, D. M. Korzhnev, M. Wolf-Watz, J. J. Bosco, Daryl A. Skalicky, L. E. Kay, and D. Kern, Nature **438**, 117 (2005).
[5] V. den Bedem H and F. JS., Nat. Methods **12**, 307 (2015).
[6] J. R. Lewandowski, M. E. Halse, M. Blackledge, and L. Emsley, Science **348**, 578 (2015).
[7] J. M. Lamley, M. J. Lougher, H. J. Sass, M. Rogowski, S. Grzesiek, and J. R. Lewandowski, Phys. Chem. Chem. Phys. **17**, 21997 (2015).
[8] D. Yang and L. E. Kay, Journal of Molecular Biology **263**, 369 (1996).
[9] V. Kasinath, K. A. Sharp, and A. J. Wand, Journal of the American Chemical Society **135**, 15092 (2013).
[10] O. Allnér, N. Foloppe, and L. Nilsson, The Journal of Physical Chemistry B **119**, 1114 (2015).
[11] M. Akke, R. Brüschweiler, and A. G. Palmer, Journal of the American Chemical Society **115**, 9832 (1993).
[12] C. Sanchez-Medina, A. Sekhar, P. Vallurupalli, M. Cerminara, V. Muoz, and L. E. Kay, Journal of the American Chemical Society **136**, 7444 (2014).
[13] R. B. Best and M. Vendruscolo, Journal of the American Chemical Society **126**, 8090 (2004).
[14] S. A. Showalter and B. Rafael, Journal of Chemical Theory and Computation **3**, 961 (2007).
[15] S. A. Showalter, E. Johnson, M. Rance, and R. Brüschweiler, Journal of the American Chemical Society **129**, 14146 (2007).
[16] P. Maragakis, K. Lindorff-Larsen, M. P. Eastwood, R. O. Dror, J. L. Klepeis, I. T. Arkin, M. . Jensen, H. Xu, N. Trbovic, R. A. Friesner, A. G. Palmer, and D. E. Shaw, The Journal of Physical Chemistry B **112**, 6155 (2008).
[17] N. Trbovic, B. Kim, R. A. Friesner, and A. G. Palmer, Proteins: Structure, Function, and Bioinformatics **71**, 684 (2008).
[18] H. Wennerstroem, B. Lindman, O. Soederman, T. Drakenberg, and J. B. Rosenholm, Journal of the American Chemical Society **101**, 6860 (1979).
[19] G. Lipari and A. Szabo, J. Am. Chem. Soc. **104**, 4546 (1982).
[20] D. E. Woessner, The Journal of Chemical Physics **37**, 647 (1962).
[21] H. Shimizu, The Journal of Chemical Physics **37**, 765 (1962).
[22] P. Luginbühl, K. V. Pervushin, H. Iwaï, and K. Wüthrich, Biochemistry **36**, 7305 (1997).
[23] J. Blake-Hall, O. Walker, and D. Fushman, "Characterization of the overall rotational diffusion of a protein from 15n relax-

ation measurements and hydrodynamic calculations," in *Protein NMR Techniques*, edited by A. K. Downing (Humana Press, Totowa, NJ, 2004) pp. 139–159.
[24] J. J. Prompers and R. Brüschweiler, Journal of the American Chemical Society **124**, 4522 (2002).
[25] V. Wong and D. A. Case, The Journal of Physical Chemistry B **112**, 6013 (2008).
[26] J. S. Anderson and D. M. LeMaster, Biophysical Chemistry **168**, 28 (2012).
[27] C.-Y. Lu and D. A. V. Bout, The Journal of Chemical Physics **125**, 124701 (2006).
[28] A. Ciragan, A. S. Aranko, I. Tascon, and H. Iwaï, Journal of Molecular Biology **428**, 4573 (2016).
[29] A. Abragam, *The Principles of Nuclear Magnetism* (Oxford University Press, 1961).
[30] L. E. Kay, D. A. Torchia, and A. Bax, Biochemistry **28**, 8972 (1989).
[31] Y. Hiyama, C. H. Niu, J. V. Silverton, A. Bavoso, and D. A. Torchia, Journal of the American Chemical Society **110**, 2378 (1988).
[32] B. Halle, The Journal of Chemical Physics **131**, 224507 (2009).
[33] $\tau_1 = (4D_{xx} + D_{yy} + D_{zz})^{-1}$, $\tau_2 = (D_{xx} + 4D_{yy} + D_{zz})^{-1}$, $\tau_3 = (D_{xx} + D_{yy} + 4D_{zz})^{-1}$, $\tau_4 = [6(D + (D^2 - L^2)^{-1/2}]^{-1}$, $\tau_5 = [6(D - (D^2 - L^2)^{-1/2}]^{-1}$, $D = \frac{1}{3}(D_{xx} + D_{yy} + D_{zz})$ and $L^2 = \frac{1}{3}(D_{xx}D_{yy} + D_{xx}D_{zz} + D_{yy}D_{zz})$.
[34] P. Dosset, J.-C. Hus, M. Blackledge, and D. Marion, Journal of Biomolecular NMR **16**, 23 (2000).
[35] J. G. de la Torre, M. Huertas, and B. Carrasco, Journal of Magnetic Resonance **147**, 138 (2000).
[36] O. Fisette, P. Lage, S. Gagn, and S. Morin, Journal of Biomedicine and Biotechnology **2012**, 254208 (2012).
[37] Y. Gu, D.-W. Li, and R. Brüschweiler, Journal of Chemical Theory and Computation **10**, 2599 (2014).
[38] M. J. Abraham, T. Murtola, R. Schulz, S. Pll, J. C. Smith, B. Hess, and E. Lindahl, SoftwareX **12**, 19 (2015).
[39] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror, and D. E. Shaw, Proteins: Structure, Function, and Bioinformatics **78**, 1950 (2010).
[40] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).
[41] S. Izadi, R. Anandakrishnan, and A. V. Onufriev, The Journal of Physical Chemistry Letters **5**, 3863 (2014).
[42] G. Bussi, D. Donadio, and M. Parrinello, J. Chem. Phys. **126** (2007).
[43] M. Parrinello and A. Rahman, J. Appl. Phys. **52**, 7182 (1981).
[44] T. Darden, D. York, and L. Pedersen, J. Chem. Phys. **98**, 10089 (1993).
[45] U. L. Essman, M. L. Perera, M. L. Berkowitz, T. Larden, H. Lee, and L. G. Pedersen, J. Chem. Phys. **103**, 8577 (1995).
[46] B. Hess, J. Chem. Theory Comput. **4**, 116 (2008).
[47] M. Abraham, D. van der Spoel, E. Lindahl, B. Hess, and the GROMACS development team, *GROMACS user manual version 5.0.7* (2015).
[48] R. T. McGibbon, K. A. Beauchamp, M. P. Harrigan, C. Klein, J. M. Swails, C. X. Hernández, C. R. Schwantes, L.-P. Wang, T. J. Lane, and V. S. Pande, Biophysical Journal **109**, 1528 (2015).
[49] "Matlab, r2016a, the mathworks, inc., natick, massachusetts, united states.".
[50] A. Nowacka, N. Bongartz, O. Ollila, T. Nylander, and D. Topgaard, J. Magn. Res. **230**, 165 (2013).
[51] T. M. Ferreira, O. H. S. Ollila, R. Pigliapochi, A. P. Dabkowska, and D. Topgaard, J. Chem. Phys. **142**, 044905 (2015).
[52] G. Barbato, M. Ikura, L. E. Kay, R. W. Pastor, and A. Bax, Biochemistry **31**, 5269 (1992).
[53] V. Krishnan and M. Cosman, Journal of Biomolecular NMR **12**, 177 (1998).
[54] W. R. Carper and C. E. Keller, The Journal of Physical Chemistry A **101**, 3246 (1997).
[55] D. Frishman and P. Argos, Proteins: Structure, Function, and Bioinformatics **23**, 566 (1995).
[56] W. Humphrey, A. Dalke, and K. Schulten, Journal of Molecular Graphics **14**, 33 (1996).
[57] M. Javanainen, A. Lamberg, L. Cwiklik, I. Vattulainen, and O. S. Ollila, Langmuir **0**, null (0), http://dx.doi.org/10.1021/acs.langmuir.7b02855.
[58] F. Hfling and T. Franosch, Reports on Progress in Physics **76**, 046602 (2013).

SUPPLEMENTARY INFORMATION

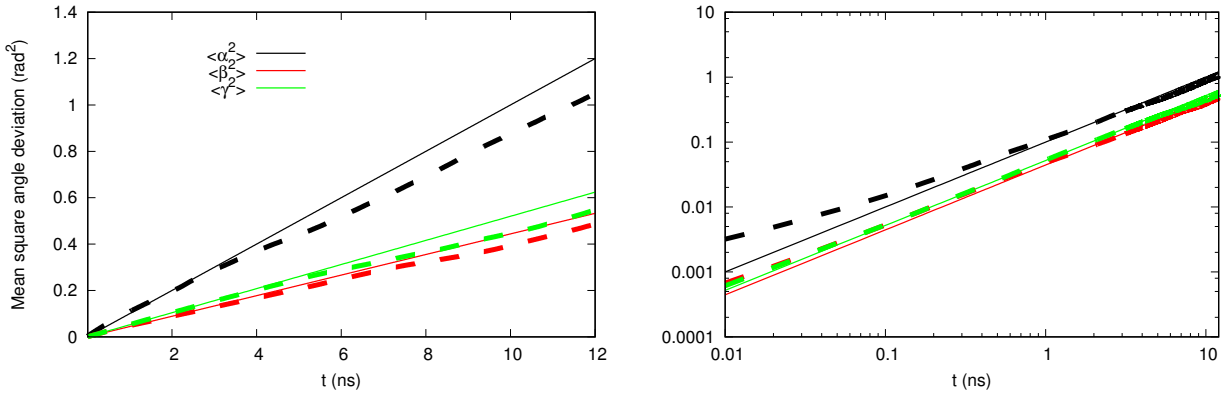

FIG. 8. The intertia tensor angles as a function of time and mean square angular deviations for *Pa*TonB simulation with tip4p water model at 310K.

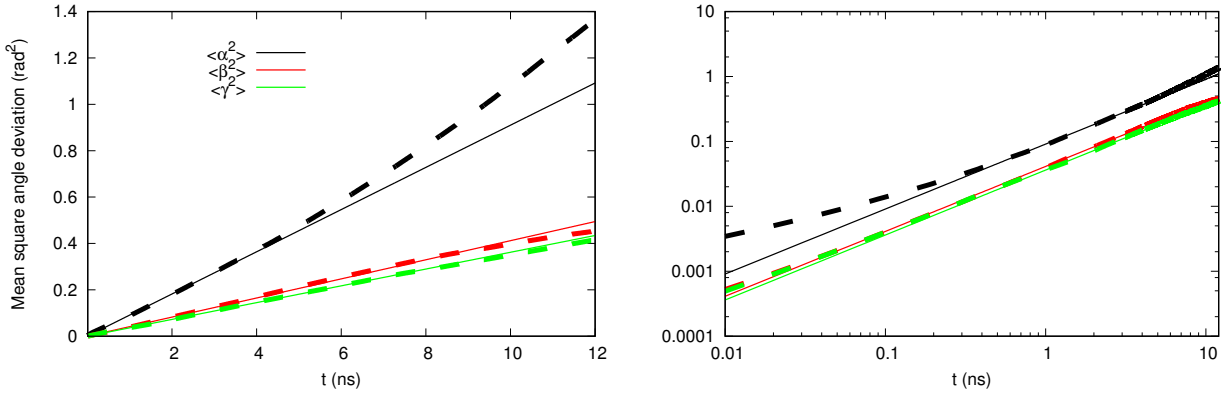

FIG. 9. The intertia tensor angles as a function of time and mean square angular deviations for *Pa*TonB simulation with tip4p water model at 298K.