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

Hideo Iwai
Insititute of Biotechnology, University of Helsinki
(Dated: June 22, 2017)

I. INTRODUCTION

Conformational sampling and entropy of proteins play a significant role in functionality and interactions with other biomolecules. In addition to internal dynamics due to conformation sampling, the proteins also experience overall brownian tumbling. These properties are experimentally accessible through spin relaxation times of ^{15}N and ^{13}C nucleai measured with nuclaer magnetic resonance (NMR) techniques [1–5]. Spin relaxation rate experiments have been used to, for example, analyze conformational entropies [1, 6, 7], binding entropies [1, 8], resolve sampled structures [3] and validate molecular dynamics simulations [9–13]. These analyzes are almost exclusively based on the separation of internal conformational sampling and overall rotational tumbling [14, 15] and isotropic overall diffusion is often assumed, while analysis of anisotropic molecules is significantly more complicated [1, 2, 16, 17]. Thus, 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 time experiments for molecules with significant anisotropy or correlations between internal and overall rotational motions. Practical applications are, however, limited by inaccuracies in the force field descriptions and available time scales in the simulations [12, 13, 18–20]. The overestimation of overall rotational diffusion due to inaccuracies in water models [19] and very long simulations required to calculate rotational correlation functions with sufficient accuracy from single molecules in MD simulations [20, 21] have been the main complications.

In this work we overcome these issues by assuming that the overall rotational dynamics of protein follows anisotropic rigid body diffusion around inertia axes. Diffusion coefficients are directly calculated from angular displacements of protein intertia axes and used to determine overall rotation part of backbone N-H bonds correlation functions. This reduces the simulation length required for accurate determination of rotational correlation functions. Furthermore, the overestimated rotational diffusion due to water model can be anisotropically corrected by scaling the diffusion coefficients in all directions with a constant factor. The approach is demonstrated by interpreting the experimental spin relaxation times of HpTonB-92 [22] and PsTonB [**?**] protein constructs, both having significantly anisotropic shape. These protein segments are critical for iron transport into Helicbacteri Pyroli and Pseudomonas bacteria, respectively.

* samuli.ollila@helsinki.fi

II. METHODS

A. Spin relaxation experiments and rotational dynamics of molecules

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

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

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 experimentally measured spin relaxation times T_1, T_2 and T_{NOE} is given by Redfield equations [23, 24]

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

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

$$\frac{1}{T_{\mathrm{NOE}}} = 1 + \frac{d_{\mathrm{NH}}^2 N_{\mathrm{H}}}{20}\Big[6J(\omega_{\mathrm{N}}+\omega_{\mathrm{H}}) + J(\omega_{\mathrm{H}}-\omega_{\mathrm{N}}))\Big]\frac{\gamma_{\mathrm{H}}}{\gamma_{\mathrm{N}}R_1}, \tag{5}$$

where ω_{N} and ω_{H} are the Larmor angular frequencies of ^{15}N and ^{1}H respectively, and N_{H} is the number of bound protons.

The dipolar coupling constant is given by

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

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. Average cubic length is calculated as $\langle r^3_{\mathrm{CN}} \rangle = (0.101\mathrm{nm})^3$ and the value of $\Delta\sigma = -160$ ppm is used for the chemical shift anisotropy of N-H bonds in proteins [24, 25].

Spin relaxation experiments are typically interpreted by dividing the rotational dynamics in overall tumbling and internal relaxation due to conformational sampling and assuming that the motions are independent [1, 2, 14, 15]. The rotational correlation function for chemical bonds can be then written as

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

where $C_I(t)$ and $C_O(t)$ are correlation functions for internal and overall rotations, respectively. Conformational sampling can be described in this approximation by using the square of order parameter respect to molecular axes S^2, which is given by the plateau of the internal rotational correlation function, and 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 [15].

The overall rotational correlation function is often described by approximating protein as a rigid body, which gives a sum of five exponentials for fully anisotropic molecule [2, 26]

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

where time constants τ_j are related to the diffusion constants around three principal axes of a molecule (D_{xx}, D_{yy} and D_{zz}) [27] and prefactors A_j to the directions of chemical bonds respect to the molecular axes [16, 26].

The simplest approach to extract molecular dynamics from experimental data is the original "model free analysis", where isotropic diffusion for overall rotational correlation function and monoexponential decay with time constant τ_{eff} for internal correlation function are assumed [15]. Equation 8 reduces to monoexponential for isotropic diffusion and the overall dynamics can be described with a single time constant τ_c. Dynamics for each residue can be described with internal rotational relaxation time τ_{eff} and the order parameter S^2. For this simple model the parameters can be successfully fitted to the experimental data, however the fitting becomes significantly more difficult for proteins with anisotropic overall diffusion or several internal timescales [1, 16, 28]. Alternative approach is to describe overall protein diffusion with hydrodynamical calculations [29]. These are less problematic for anisotropic systems but are sentitive for the assumptions about protein hydration shell [29].

B. Rotational dynamics from molecular dynamics simulations

Classical molecular dynamics simulation gives a trajectory for each atom in a system as a function of time. Rotational correlation functions for each bond can be directly calculated from the trajectories by using Eq. 2 and then 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 [9–13, 30] or interpret experiments [30]. The comparison is often complicated by statistical fluctuations of calculated correlation functions with time scales corresponding the overall rotational diffusion. Here we reduce the fluctuations by calculating the overall rotational diffusion coefficients directly from angular displacements of intertia axes. The overall rotation is then assumed to have rotational correlation function of anisotropic rigid body rotation, Eq. 8, and the timescales are determined from rotational diffusion coefficients [27]. Calculation of diffusion coefficients is based on fitting a slope on linear mean square angle deviation (see below), thus it is numerically more robust and requires less simulation data than a direct fit of multiexponential sum in Eq. 8 to the rotational correlation function calculated from MD simulation.

The analysis can be divided in essentially six steps:
1) Total rotational correlation functions $C(t)$ for protein N-H bonds are directly calculated from MD simulation trajectory by applying Eq. 2.
2) Rotational correlation functions for internal dynamics $C_I(t)$ are calculated from MD simulation trajectory from where the overall rotation of protein is removed.
3) The overall and internal motions are assumed to be independent and overall rotational correlation function is calculated as $C_O(t) = C(t)/C_I(t)$ according to Eq. 6.
4) The protein axes of inertia and mean square angle deviations of rotation around the axes are calculated as function of time 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 of inertia axes

$$\begin{aligned} \langle \Delta\alpha^2_{t'+t} \rangle_{t'} &= 2D_x t \\ \langle \Delta\beta^2_{t'+t} \rangle_{t'} &= 2D_y t \\ \langle \Delta\gamma^2_{t'+t} \rangle_{t'} &= 2D_z t, \end{aligned} \tag{9}$$

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 mean square angle deviations of rotation around protein intertia axes from shortest to longest, respectively.
6) Rigid body rotational correlation functions for anisotropic motion are determined from Eq. 8. Timescales τ_j are given by rotational diffusion constants from previous step [27] and weighting factors A_j are determined by fitting the equation to overall rotational correlation functions from MD simulations $C_0(t)$ determined in step 3.
7) Rigid body rotational correlation functions from previous

step are combined with internal rotational correlation functions from MD simulations (step 2) by using Eqs. 6 and 8 to determine new correlation functions

$$C_N(t) = C_I(t) \sum_{j=1}^{5} A_j e^{-t/\tau_j}. \quad (10)$$

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 by scaling the rotational diffusion coefficients, i.e. timescales τ_j, with a constant factor before calculating new correlation functions from Eq. 10.

C. Simulation and analysis details

All simulations were ran using Gromacs 5 [31] and Amber ff99SB- ILDN [32] force field for proteins . The proteins were solvated to tip3p[33], tip4p [33] or OPC4 [34] water models. Initial structures were taken from NMR structures published elsewhere for PsTonB [**?**] and HpTonB-92 [22]. Temperature was coupled to desired value with v-rescale thermistat [35] and pressure was isotropically set to 1 bar using Parrinello-Rahman barostat [36]. Timestep was 2 fs, Lennart-Jones interactions were cut-off at 1.0 nm, PME [37, 38] was used for electrostatics and LINCS was used to constraint all bond lengths [39]. Simulation trajectory and related files are available at [**?**]. The simulated systems are listed in Table I

Rotational correlation functions are calculated with *gmx rotacf* from Gromacs package [40]. Overall rotation was removed for $C_I(t)$ calculation by using fit option in *gmx trjconv* from Gromacs package [40]. The order parameters S^2 were determined by averaging rotational correlation functions from oriented trajectory, $C_I(t)$, over lag times above 50 ns. Effective correlation times were then calculated by using Eq. 7. Inertia axes of protein were calculated with *compute_inertia_tensor* function from MDTraj python library [41].

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

$$C_N(t) = \sum_{i=1}^{N} \alpha_i e^{-t/\tau_i} \quad (11)$$

to the new correlation function from Eq. 10 by using the *lsqnonneg* routine in MATLAB [42]. The Fourier transform was then calculated by using analytical function for the sum of exponentials

$$J(\omega) = 4 \sum_{i=1}^{N} \alpha_i \frac{\tau_i}{1 + \omega^2 \tau_i^2}. \quad (12)$$

Similar approach is used previously for lamellar systems in combination with solid state NMR experiements [43, 44].

III. RESULTS

A. Global rotational dynamics of protein

Mean square angle deviations for rotation around inertia axes of PsTonB protein in simulation with OPC4 water model are shown in Fig. 1. This is the longest simulation data set in this work (1.2μs) and linear behaviour of 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 a good statistics for rotational dynamics analyzed from a single molecule in MD simulations [21]. Deviations from linear behaviour are only seen with lag times longer than this limit, as also demonstrated for shorter simulations in Figs. 8 and 9. Plot with log-log scale in Fig. 1 B) reveals a weakly subdiffusive region only below very short timescales of approximately 0.12 ns. Thus, we conclude that the protein experiences brownian rotational tumbling with a good approximation and that the diffusion coefficients can be calculated from the slope of mean square angle deviations according to Eq. 9 by using lag times less than one hundreth of the total MD simulation length.

The resulting rotational diffusion constants from different simulations are shown in Table I. The values are larger than would be expected from experiments for different proteins with similar size [45], especially when tip3p water model is used. This is in line with previously reported simulation results, where the results were explained by overestimated water self diffusion [19]. As expected, rotational diffusion coefficients increase with the temperature and decreasing size of a protein.

The rotational diffusion coefficients are used to determine the timescales τ_j for anisotropic rigid body rotation in Eq. 8 [27], which used as an approximation for the overall rotational dynamics of a protein. The prefactors A_j are determined by fitting the equation to overall rotational correlation functions, $C_O(t)$, calculated from MD simulations. The resulting rigid body rotational correlation function is then combined with internal correlation function, $C_I(t)$, to determine the new correlation functions from Eq. 10.

This analysis is exemplified in Fig. 2 for residues in different parts of PsTonB protein having different features in rotational dynamics. Flexible C-terminus is represented by residue 341, more rigid β-sheet by residue 331 and a flexible loop between two sheets by residue 322. The total correlation functions $C(t)$ in Fig. 2 A) (solid lines), calculated from original MD trajectories, decay toward zero within $\sim$10-50 ns for all residues. Internal correlation functions $C_I(t)$ in Fig. 2 B), calculated from trajectory with removed overall rotation of the protein, decay to a plateau value, which defines the square of the order parameter S^2. As expected, the β-sheet residue 331 has the largest order parameter value and fastest decay to it, while order parameters for the loop and C-terminus residues are significantly smaller and deacay is slower due to larger conformational ensemble sampled by these regions. The overall rotational correlation functions $C_O(t)$ for a protein, determined as $C_O(t) = C(t)/C_I(t)$, are shown in Fig. 2 C) (solid lines).

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

Protein	Water model	T (K)	$t_{\rm sim}$ (ns)	$t_{\rm anal}$ (ns)	D_{xx}	D_{yy}	D_{zz}	$D_{||}/D_{+}$	D_{av}	files
PaTonB	tip4p	298	400	390	1.81 ± 0.01	2.06 ± 0.03	4.55 ± 0.03	2.35 ± 0.04	2.80 ± 0.02	[**?**]
PaTonB	tip4p	310	400	390	2.60 ± 0.02	2.22 ± 0.05	5.0 ± 0.1	2.07 ± 0.09	3.26 ± 0.07	[**?**]
PaTonB	OPC4	310	1200	1190	2.01 ± 0.01	2.19 ± 0.01	5.01 ± 0.03	2.39 ± 0.02	3.07 ± 0.01	[**?**]
HpTonB-92	tip3p	310	570	370	8.25 ± 0.05	7.67 ± 0.06	15.9 ± 0.3	1.99 ± 0.06	10.6 ± 0.2	[**?**]
HpTonB-92	tip3p	303	800	790	6.24 ± 0.02	7.04 ± 0.03	11.9 ± 0.2	1.80 ± 0.03	8.40 ± 0.07	[**?**]
HpTonB-92	tip4p	310	470	370	3.6 ± 0.1	3.24 ± 0.01	6.3 ± 0.3	1.8 ± 0.1	4.4 ± 0.2	[**?**]
HpTonB-92	tip4p	303	400	200	2.7 ± 0.1	2.71 ± 0.02	5.6 ± 0.5	2.1 ± 0.2	3.7 ± 0.2	[**?**]
HpTonB-92	OPC4	310	800	790	2.85 ± 0.01	2.70 ± 0.01	5.56 ± 0.01	2.00 ± 0.01	3.70 ± 0.01	[**?**]

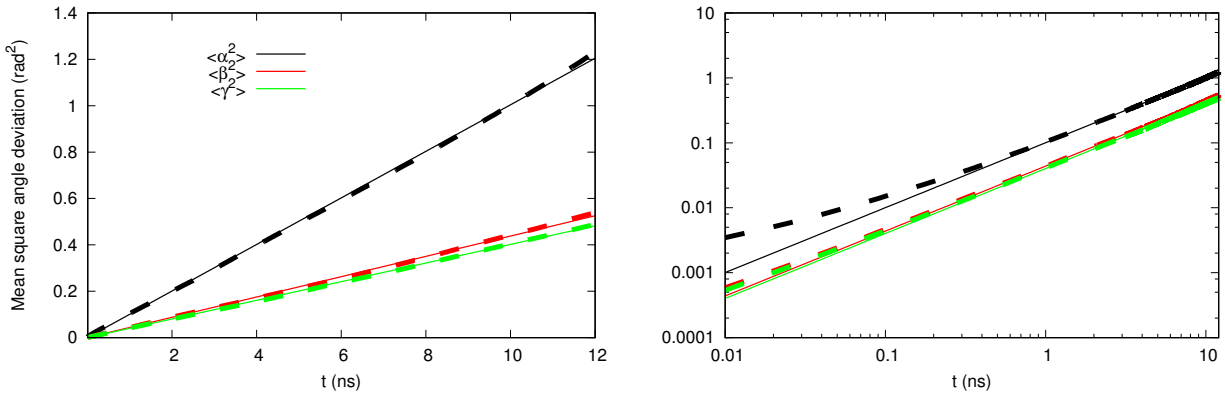

FIG. 1. The intertia tensor angles as a function of time and mean square angular deviations for PsTonB simulation with OPC water model.

The determined anisotropic rigid body rotational correlation functions from Eq. 8 and new correlation functions from Eq. 10 are shown in Figs. 2 C) and A) (dashed lines), respectively. The correlation functions are indistinguishable from original MD simulation results with lag times shorter than one hundredth of total simulation time (approximately 4-12ns for the studied systems), which is the expected limit for a good statistics in single molecule MD simulations [21]. This suggests that the anisotropic rigid body diffusion model (Eq. 8) and separation of internal and global motions (Eq. 6) are good approximations for the studied system. Thus, we conclude that the new correlation functions can be used to reduce the statistical fluctioations with long lag times and correct overall rotational dynamics by scaling the diffusion coefficints with constant factor for spin relaxation time calculations from MD simulations. Exceptionally large statistical fluctuations for overall rotational correlation function of flexible C-terminus (residue 341) with longer lag times are observed, because small contribution of overall rotation dynamics is difficult to detect with high accuracy for segments with small order parameters.

B. Global rotational dynamics in simulations and experiments

Spin relaxation times from HpTonB-92 simulations and experiments are compared in Fig. 3. Simulation results with tip4p water model are in good agreement with experiments. However, simulations with tip3p water model are significantly off from experiments and underestimate T_1/T_2 ratio, suggesting too fast overall rotational diffusion dynamics [46] in agreement with previous studies [19]. Indeed, division of diffusion constants by a constant factor of 2.9 before applying Eq. 10 brings spin relaxation times in good agreement with experiments, as shown in Fig. 4.

Spin relaxation times from PsTonB simulations and experiments are compared in Fig. 5. Simulations with tip4p or OPC4 water models systemically underestimate T_1 values and T_1/T_2 ratio when compared with experiments, but much less extent than in HpTonB-92 simulations with tip3p above. Dividing the diffusion coefficinents with a constant factor of 1.2 gives a good agreement for spin relaxation rates between tip4p simulation and experiments at 298K, as seen in Fig. 6. The effect of temperature difference of 12 degrees is significanlty smaller than the difference between simulations and experiments or the effect of the diffusion constant scaling.

The good agreement between spin relaxation data from experiments and simulations with scaled diffusion coefficients suggests that the values can be used to interpret the anisotropic

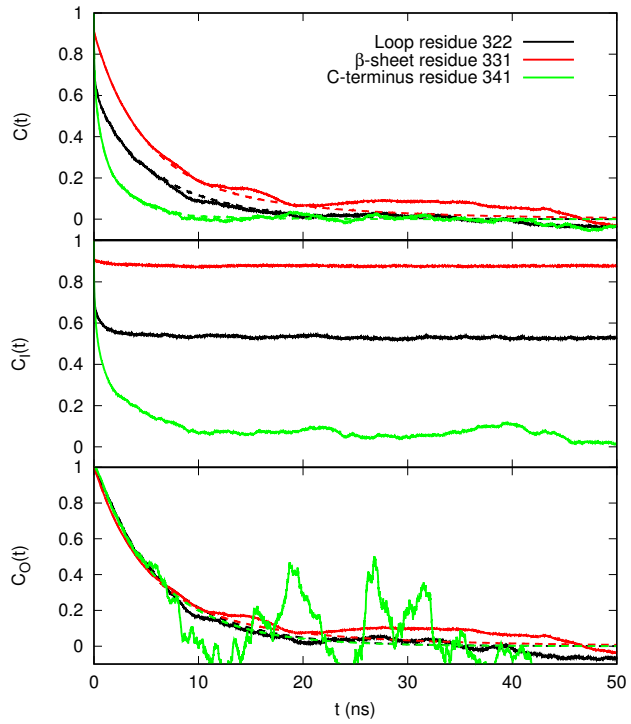

FIG. 2. Rotational correlation functions calculated from MD simulations of PsTonB with tip4p water model at 298K for residues at different regions. A) total correlation functions $C(t)$ calculated from MD simulation (solid lines) and new correlation functions determined from Eqs. 6 and 8 by using rotational diffusion constants and fitted prefactors (dashed lines), B) correlation functions for internal motions calculated from simulation with removed overall protein rotation C) 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).

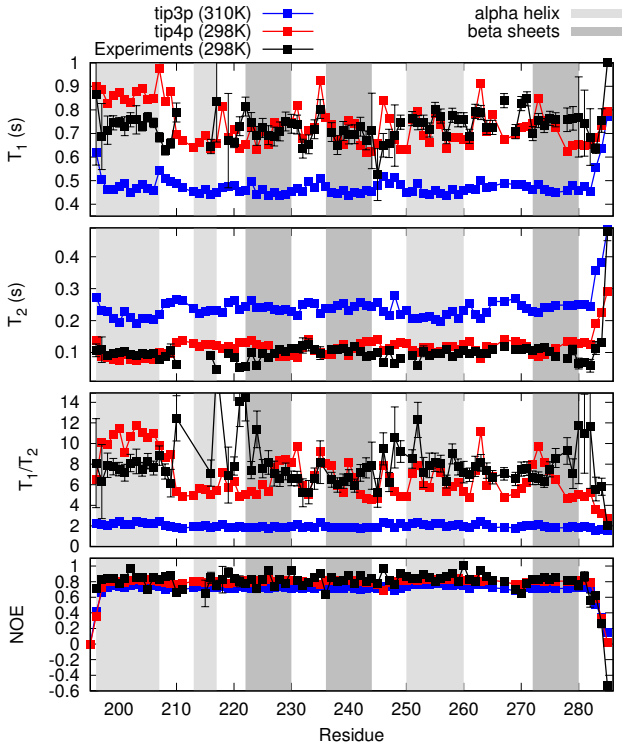

FIG. 3. Spin relaxation times for HpTonB-92 from experiments [**?**] and simulations with different water models.

diffusion of these proteins from the measured NMR data. The scaled rotational diffusion coefficients from MD models giving the best agreement with experimental spin relaxation data are shown in Table I. The scaled diffusion coefficients from tip3p simulations were chosen for HpTonB-92, because they give slightly better overall agreement with experiments than simulations with tip4p water model. The results are in line with values for more isotropic proteins, which are previously determined by different methods to interpret the spin relaxation data [45].

C. Interpretation of protein internal relaxation from MD simulations

Experimental spin relaxation times can be reproduced by MD simulations when overall rotational diffusion is corrected by scaling with a constant factor, as seen in Figs. 4 and 6. These simulations can be used to interpret internal relaxation processes observed in spin relaxation experiments for proteins.

TABLE II. Rotational diffusion coefficients scaled with constant factor which gives a good agreement for spin relaxation data, 2.9 for tip3p simulation of HpTonB and by 1.2 for tip4p simulation of PsTonB.

	HpTonB-92	PsTonB
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)	5.7 ± 0.1	7.2 ± 0.1

NMR experiments and the most realistic MD simulation based model for HpTonB-92 show very little variation between spin relaxation times for different residues in as seen in Fig. 3. This indicates a rigid protein structure, which is also seen in MD simulation snapshots overlayed in Fig. 3 A). Enhanced conformational sampling is seen only for few residues in terminal ends, which are also visible in spin relaxation data. Some deviation from average spin relaxation times are observed also between residues 210-222, which probably arises from sampling between two orientations of the α-helix (see discussion below and in Ref. [**?**]). Exceptionally low order parameters and long effetive correlation times are also observed for residues 245-250 in simulations and the same region gives small T_1 times in experiments. However, the interpretation of these observations is not straightforward, because low T_1 in this region are not reproduced by simulations. More

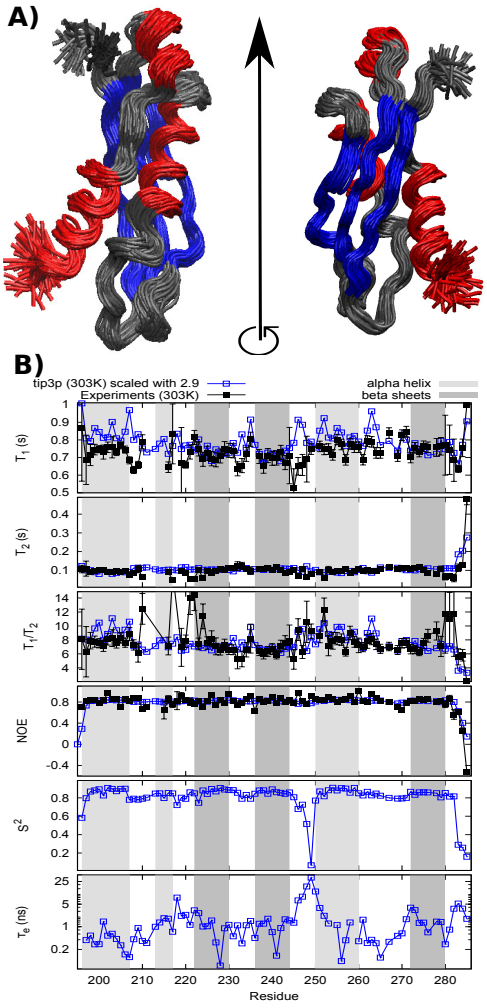

FIG. 4. A) Structures sampled by HpTonB-92 from MD simulations with tip3p at 303 K (100 structures from 400ns long trajectory). Secondary structures are colour labelled with Visual Molecular dynamics [47, 48]; α-helices are red and β-sheets are blue. B) Spin relaxation times from experiments and tip3p simulations 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

detailed discussion together with longer HpTonB construct is presented elsewhere [**?**].

More variety in internal dynamics between residues is observed for PsTonB protein, as seen in Fig. 6. Segments with enhanced conformational sampling are labelled with yellow colour in Fig. 6 A). The terminal ends show significantly enhanced conformational sampling in MD simulation snapshots, which is also observed in spin relaxation times from simulations and experiments. The terminal ends are also characterized by low order parameters and long effective internal correlation times arising from larger amount of sampled

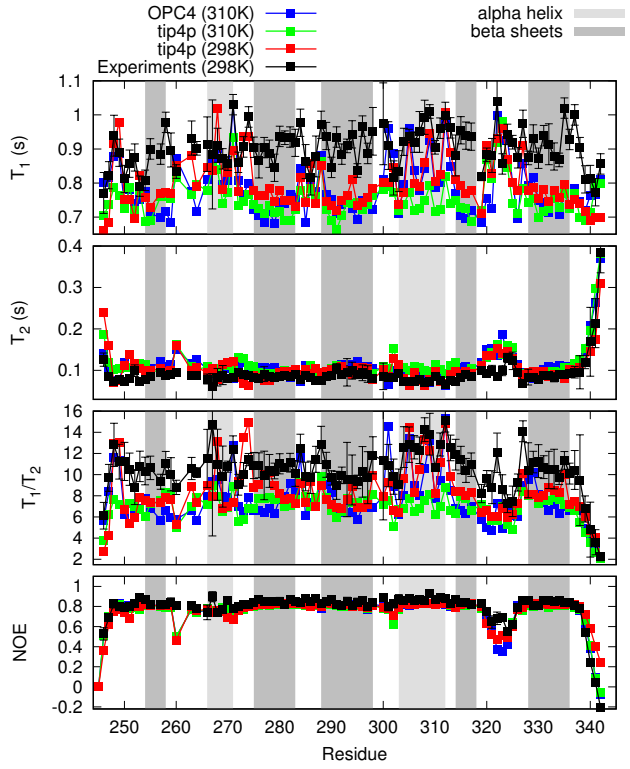

FIG. 5. Spin relaxation times for PsTonB from experiments [**?**] and simulations with different water models.

conformations. Enhanced conformational sampling is also observed for residues between 320-326, which is a loop between two β-sheets. MD simulations predict low order parameters and long internal effective correlation times are also for residues between 260-274, which similar region as residues 210-222 in HpTonB-92. In this case the simulations reveal two different orientations sampled by the α-helix in this region (colour labelled with pink in Fig. 5 A)), which also explains the lower resolution in NMR spectra for this region observed for PsTonB and HpTonB constructs [**?**].

Relaxation processes can be further quantified by analysing the timescales, which lead to spin relaxation times in agreement with experiments. This is exemplified in Fig. 7 by plotting the prefactors corresponding each timescale after fitting Eq. 11 to MD simulation data. The same residues for PsTonB as in Fig. 2 are used. Rotational relaxation of residue 331 in β-sheet is dominated by overall rotational timescales of $\sim$5.5 ns and $\sim$8 ns, while contribution from fast internal motions is much smaller, as expected for rigid structure with large order parameter value. Rotational relaxation of residue 322 in flexible loop is also dominated by overall rotation timescales around $\sim$8 ns, but fast motions from internal dynamics are more significant than for the β-sheet residue in agreement with lower order parameter value in the loop. On the other hand, rotational dynamics of residue 341 in N-terminal is dominated by timescales below 3 ns related to the internal protein relaxation. The contribution from timescales close to $\sim$13 ns is probably related to slow conformational

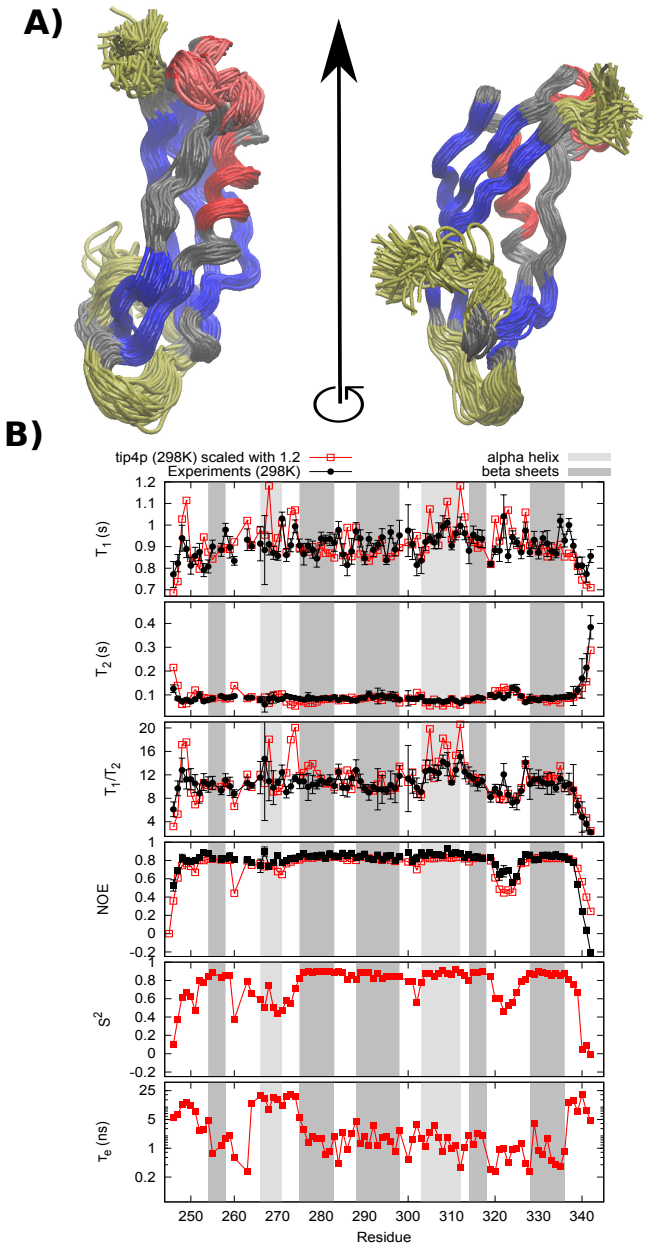

FIG. 6. A) Structures sampled by PsTonB from MD simulations with tip4p at 298 K (100 structures from 400ns long trajectory). Secondary structures are colour labelled with Visual Molecular dynamics [47, 48]; α-helixces are red and β-sheets are blue. Residues 246-251, 320-326 and 338-342 with increased internal dynamics are yellow and α-helix sampling between two orientations (residues 266-270) is pink in the left column. B) Spin relaxation times from experiments and tip4p simulations 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.

sampling of N-terminus, rather than overall rotational dynamics. This is in agreement with Fig. 5, suggesting that sampling of large amount of conformations leads to small order parameters and large effective correlation times.

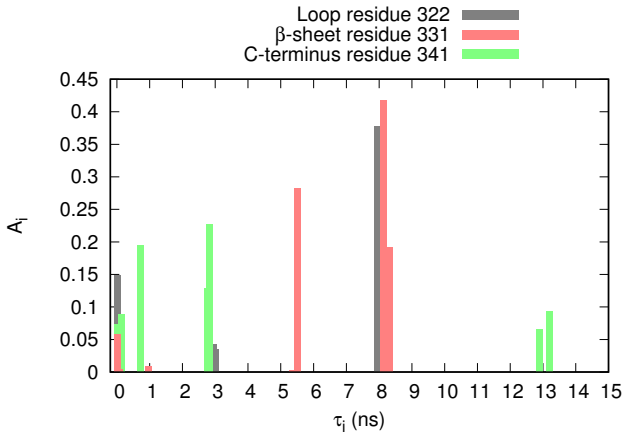

FIG. 7. Prefactors A_i corresponding different timescales τ_i in Eq.11 from correlation functions giving a good agreement with experimental spin relaxation times for PsTonB at 298K.

IV. DISCUSSION AND CONCLUSIONS

Protein rotational dynamics was separated to the overall brownian tumbling around protein intertia axes and internal conformational sampling. The overall rotational diffusion constants were calculated from the slope of mean square angle deviations (Eq. 9), which were found to be essentially linear for all inertia axes. The results indicate that monomeric proteins in dilute solution experience rotational brownian tumbling in agreement with previous MD simulation study [19]. Only a small subdiffusive behaviour was found with short timescales below 0.12 ns. This result can be used as a constrast for crowded environment, where anomalous diffusion is expected to be more significant [49].

General form of rotational correlation function for anisotropic rigid body (Eq. 8) with timescales from rotational diffusion constants was successfully fit to the overall rotational correlation functions from MD simulations for all individual N-H bonds. Furthermore, the new correlation functions calculated from Eq. 10 were in good agreement with the total correlation functions calculated from original trajectories. The results suggest that the intertia axes can be used to describe the overall rotational diffusion as a good approximation and that the assumption of separation of internal and overall rotational relaxation (Eq. 6) is a good approximation for the studied proteins, as observed previously also for other protein [7, 19].

Usage of the new correlation functions from Eq. 10 in spin relaxation time calculations reduces the statistical fluctuations with long lag times in correlation function calculation, which allows the higher accuracy with less simulation data. The essential reason is that the rotational diffusion constants can be determined by fitting a single parameter (linear slope) which is more robust than a direct fit of multiexponential correlation function to MD simulation result. After determining the diffusion constants and using Eq. 10, the contribution from overall rotation do not have statistical fluctuations in correlation func-

tions used in spin relaxation time calculations.

Comparison to experimental results revealed that overall rotational diffusion coefficients were overestimated by a factor of $\sim$3 in simulations with tip3p water, in agreement with previous studies [18–20]. Simulations with tip4p and opc4 water models gave more realistic diffusion coefficients, which overestimated diffusion only with factors $\sim$1-1.2. The realistic overall rotational dynamics was determined by scaling the diffusion coefficients to all directions post-simulationally by the same constant factors such that the calculated T_1/T_2 was in optimal agreement with experiments. Similar correction has been done previously for proteins with isotropic rotational diffusion with single exponential rotational correlation function [7, 10–12, 50]. Alternatively only order parameters are used in the comparison with experiments [9, 12, 13, 50]. Overall rotational diffusion is previously corrected for anisotropic molecules by using isotropic reorientational eigenmode dynamics (iRED) [18] or quaternions [20]. The advantage of these approaches is that the separation between internal and overall rotational dynamics is not required, thus they would be applicable also for intrinsically disoredered proteins. However, the correction of overall rotational diffusion is based on the assumption that incorrect diffusion arises from incorrect water viscosity. Thus, a water model reproducing correct overall rotational diffusion is required to compare intrinsically disordered molecule simulations to experimental relaxation data.

A good agreement between spin relaxation times from experiments and simulations is found after overall rotational diffusion constants are scaled with a constant factor. This allows the interpretation of internal protein dynamics from spin relaxation times by using MD simulations even with tip3p water model. Here we use the approach to indentify dynamical regions in two proteins, HpTonB-92 and PsTonB with relatively rigid structures. However, the approach could be useful for intepretation of spin relaxation data from proteins with more regions containing complicated local dynamics, like Calmodulin [51].

ACKNOWLEDGMENTS

We acknowledge CSC-IT center for science for computational resources

[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] V. den Bedem H and F. JS., Nat. Methods **12**, 307 (2015).

[4] J. R. Lewandowski, M. E. Halse, M. Blackledge, and L. Emsley, Science **348**, 578 (2015).

[5] J. M. Lamley, M. J. Lougher, H. J. Sass, M. Rogowski, S. Grzesiek, and J. R. Lewandowski, Phys. Chem. Chem. Phys. **17**, 21997 (2015).

[6] V. Kasinath, K. A. Sharp, and A. J. Wand, Journal of the American Chemical Society **135**, 15092 (2013).

[7] O. Allnr, N. Foloppe, and L. Nilsson, The Journal of Physical Chemistry B **119**, 1114 (2015).

[8] M. Akke, R. Brueschweiler, and A. G. Palmer, Journal of the American Chemical Society **115**, 9832 (1993).

[9] R. B. Best and M. Vendruscolo, Journal of the American Chemical Society **126**, 8090 (2004).

[10] S. A. Showalter and R. Brschweiler, Journal of Chemical Theory and Computation **3**, 961 (2007).

[11] S. A. Showalter, E. Johnson, M. Rance, and R. Brschweiler, Journal of the American Chemical Society **129**, 14146 (2007).

[12] 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).

[13] N. Trbovic, B. Kim, R. A. Friesner, and A. G. Palmer, Proteins: Structure, Function, and Bioinformatics **71**, 684 (2008).

[14] H. Wennerstroem, B. Lindman, O. Soederman, T. Drakenberg, and J. B. Rosenholm, Journal of the American Chemical Society **101**, 6860 (1979).

[15] G. Lipari and A. Szabo, J. Am. Chem. Soc. **104**, 4546 (1982).

[16] P. Luginbhl, K. V. Pervushin, H. Iwai, and K. Wthrich, Biochemistry **36**, 7305 (1997).

[17] J. Blake-Hall, O. Walker, and D. Fushman, "Characterization of the overall rotational diffusion of a protein from 15n relaxation measurements and hydrodynamic calculations," in *Protein NMR Techniques*, edited by A. K. Downing (Humana Press, Totowa, NJ, 2004) pp. 139–159.

[18] J. J. Prompers and R. Brschweiler, Journal of the American Chemical Society **124**, 4522 (2002).

[19] V. Wong and D. A. Case, The Journal of Physical Chemistry B **112**, 6013 (2008).

[20] J. S. Anderson and D. M. LeMaster, Biophysical Chemistry **168**, 28 (2012).

[21] C.-Y. Lu and D. A. V. Bout, The Journal of Chemical Physics **125**, 124701 (2006).

[22] A. Ciragan, A. S. Aranko, I. Tascon, and H. Iwa, Journal of Molecular Biology **428**, 4573 (2016).

[23] A. Abragam, *The Principles of Nuclear Magnetism* (Oxford University Press, 1961).

[24] L. E. Kay, D. A. Torchia, and A. Bax, Biochemistry **28**, 8972 (1989).

[25] Y. Hiyama, C. H. Niu, J. V. Silverton, A. Bavoso, and D. A. Torchia, Journal of the American Chemical Society **110**, 2378 (1988).

[26] D. E. Woessner, The Journal of Chemical Physics **37**, 647 (1962).

[27] $\tau_1 = (4D_{xx} + D_{yy} + D_{zz})^{-1}$, $\tau_2 = (D_{xx} + 4D_{yy} + D_{zz})^{-1}$, $\tau_3 = (D_{xx}+D_{yy}+4D_{zz})^{-1}$, $\tau_4 = [6(D+(D^2-L^2)^{-1/2}]^{-1}$, $\tau_5 = [6(D-(D^2-L^2)^{-1/2}]^{-1}$, $D = \frac{1}{3}(D_{xx} + D_{yy} + D_{zz})$ and $L^2 = \frac{1}{3}(D_{xx}D_{yy} + D_{xx}D_{zz} + D_{yy}D_{zz})$.

[28] P. Dosset, J.-C. Hus, M. Blackledge, and D. Marion, Journal of Biomolecular NMR **16**, 23 (2000).

[29] J. G. de la Torre, M. Huertas, and B. Carrasco, Journal of Magnetic Resonance **147**, 138 (2000).

[30] O. Fisette, P. Lage, S. Gagn, and S. Morin, Journal of

Biomedicine and Biotechnology **2012**, 254208 (2012).
[31] M. J. Abraham, T. Murtola, R. Schulz, S. Pll, J. C. Smith, B. Hess, and E. Lindahl, SoftwareX **12**, 19 (2015).
[32] 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).
[33] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).
[34] S. Izadi, R. Anandakrishnan, and A. V. Onufriev, The Journal of Physical Chemistry Letters **5**, 3863 (2014).
[35] G. Bussi, D. Donadio, and M. Parrinello, J. Chem. Phys. **126** (2007).
[36] M. Parrinello and A. Rahman, J. Appl. Phys. **52**, 7182 (1981).
[37] T. Darden, D. York, and L. Pedersen, J. Chem. Phys. **98**, 10089 (1993).
[38] U. L. Essman, M. L. Perera, M. L. Berkowitz, T. Larden, H. Lee, and L. G. Pedersen, J. Chem. Phys. **103**, 8577 (1995).
[39] B. Hess, J. Chem. Theory Comput. **4**, 116 (2008).
[40] M. Abraham, D. van der Spoel, E. Lindahl, B. Hess, and the GROMACS development team, *GROMACS user manual version 5.0.7* (2015).
[41] 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).
[42] "Matlab, r2016a, the mathworks, inc., natick, massachusetts, united states.".
[43] A. Nowacka, N. Bongartz, O. Ollila, T. Nylander, and D. Topgaard, J. Magn. Res. **230**, 165 (2013).
[44] T. M. Ferreira, O. H. S. Ollila, R. Pigliapochi, A. P. Dabkowska, and D. Topgaard, J. Chem. Phys. **142**, 044905 (2015).
[45] V. Krishnan and M. Cosman, Journal of Biomolecular NMR **12**, 177 (1998).
[46] W. R. Carper and C. E. Keller, The Journal of Physical Chemistry A **101**, 3246 (1997).
[47] D. Frishman and P. Argos, Proteins: Structure, Function, and Bioinformatics **23**, 566 (1995).
[48] W. Humphrey, A. Dalke, and K. Schulten, Journal of Molecular Graphics **14**, 33 (1996).
[49] F. Hfling and T. Franosch, Reports on Progress in Physics **76**, 046602 (2013).
[50] Y. Gu, D.-W. Li, and R. Brschweiler, Journal of Chemical Theory and Computation **10**, 2599 (2014).
[51] G. Barbato, M. Ikura, L. E. Kay, R. W. Pastor, and A. Bax, Biochemistry **31**, 5269 (1992).

SUPPLEMENTARY INFORMATION

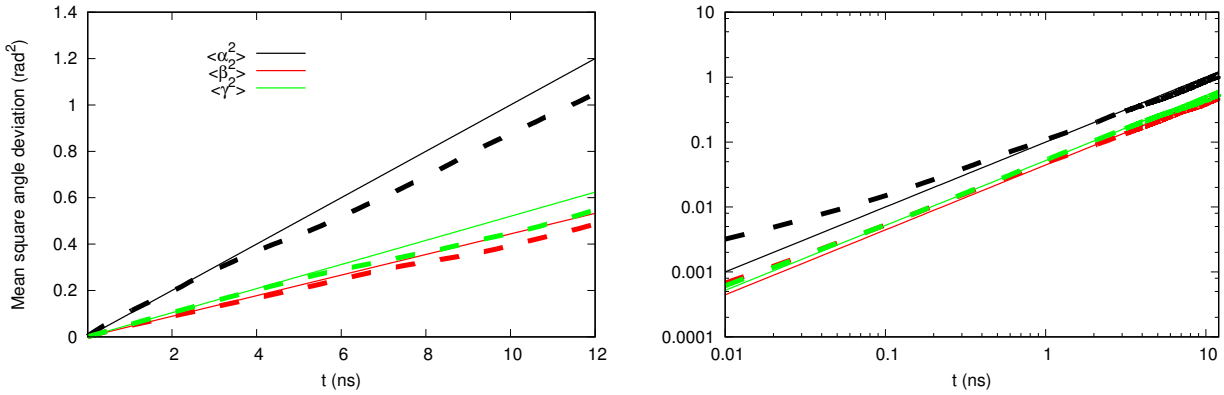

FIG. 8. The intertia tensor angles as a function of time and mean square angular deviations for PsTonB simulation with tip4p water model at 310K.

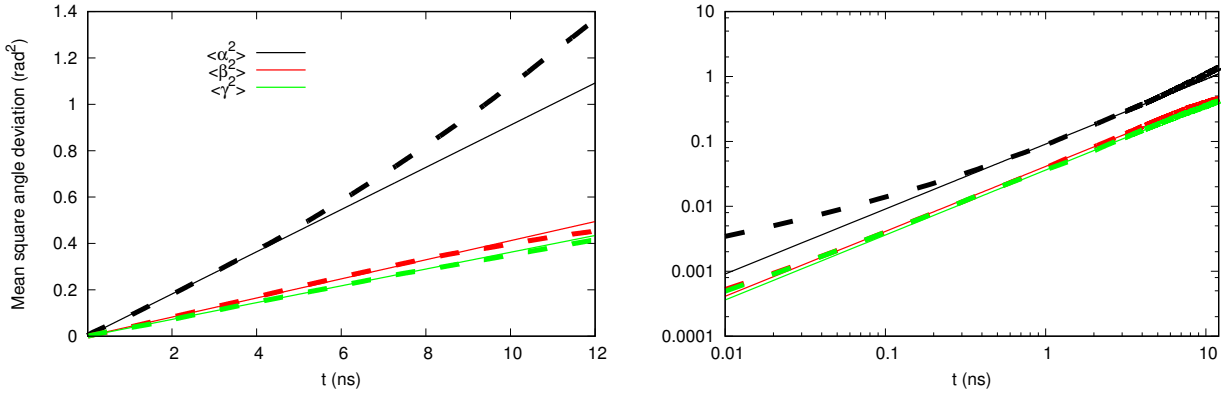

FIG. 9. The intertia tensor angles as a function of time and mean square angular deviations for PsTonB simulation with tip4p water model at 298K.