Predicting NMR Relaxation of Proteins from Molecular Dynamics Simulations with Accurate Methyl Rotation Barriers

Falk Hoffmann,[†] Frans A. A. Mulder,[‡] and Lars V. Schäfer[*,†]

[†]*Theoretical Chemistry, Ruhr University Bochum, D-44780 Bochum, Germany*

[‡]*Interdisciplinary Nanoscience Center (iNANO) and Department of Chemistry, Aarhus University, DK-8000 Aarhus, Denmark*

E-mail: lars.schaefer@ruhr-uni-bochum.de

Abstract

The internal dynamics of proteins occurring on time scales from picoseconds to nanoseconds can be sensitively probed by nuclear magnetic resonance (NMR) spin relaxation experiments, as well as by molecular dynamics (MD) simulations. This complementarity offers unique opportunities, provided that the two methods are compared at a suitable level. Recently, several groups have used MD simulations to compute the spectral density of backbone and side-chain molecular motions, and to predict NMR relaxation rates from these. Unfortunately, in the case of methyl groups in protein side-chains, inaccurate energy barriers to methyl rotation were responsible for a systematic discrepancy in the computed relaxation rates, as demonstrated for the AMBER ff99SB*-ILDN force field (and related parameter sets), impairing quantitative agreement between simulations and experiments. However, correspondence could be regained by emending the MD force field with accurate coupled cluster quantum chemical calculations. Spurred by this positive result, we tested whether this approach could be generally applicable, in spite of the fact that different MD force fields employ different water models. Improved methyl group rotation barriers for the CHARMM36 and AMBER ff15ipq protein force fields were derived, such that the NMR relaxation data obtained from the MD simulations now also display very good agreement with experiment. Results herein showcase the performance of present-day MD force fields, and manifest their refined ability to accurately describe internal protein dynamics.

Introduction

Nuclear magnetic resonance (NMR) spectroscopy is a powerful technique to study the dynamics of proteins on different time scales in a site-specific manner. In solution, fast dynamics on the ps–ns time scale can be probed with NMR relaxation.[1–5] In relaxation experiments the dynamics of a bond is measured in terms of relaxation rates, which characterize the relaxation in time of the ensemble of involved nuclear spins after a perturbation. At the same time, molecular dynamics (MD) simulations can computationally study the dynamics

of proteins at the atomic scale. The combination of NMR spectroscopy and MD simulations is therefore very powerful and has been used to scrutinize and improve on both methods, see *e.g.* Refs. 6–19.

The interpretation of the dynamics of a bond in a protein from NMR relaxation is typically done in terms of the square of the generalized order parameter, S^2, and its associated correlation time, τ_f. Lipari and Szabo[20,21] formulated such a simplified motional model based on the assumption that the global tumbling motion of the protein and the internal motion of the bond are statistically independent. While this model is usually appropriate for most of the backbone bonds in more structured regions of a protein, slow internal motions on time scales that are similar to that of overall tumbling challenge the validity of this model. In order to establish a quantitative connection between MD simulations and NMR relaxation experiments, it is therefore desirable to compare the dynamics of protein bonds directly at the level of the time correlation function (TCF), $C(t)$, and its Fourier transform, the spectral density $J(\omega)$, because these describe the dynamics without the need to invoke a motional model or a particular functional form of $J(\omega)$. While the full $C(t)$ can be obtained directly from MD simulations, NMR relaxation experiments sample specific points of the spectral density function at distinct combinations of the Larmor frequencies of the involved spins. The descriptive power of spectral densities from NMR relaxation experiments can be increased by measuring at multiple field strengths, which yields additional spectral density points, or by field shuttling.[22–24] However, these approaches are cumbersome and necessarily constrained to a limited number of discrete points.

On the computational side, the predictive power of MD simulations depends on sufficient sampling of the relevant phase space and the accuracy of the underlying potential energy function. The first describes the conformational space explored within the simulation time t_{sim} and limits the usability of MD simulations to accurately probe processes that occur on time scales that are shorter than t_{sim}. The latter aspect concerns the force field. Recent improved protein force fields[25–31] capture the dynamics of protein backbone amide bonds,

as *e.g.* probed by NMR through order parameters S^2_{NH} or relaxation rates,[32] with high accuracy. However, a comparably reliable prediction of dynamical parameters of side-chains by MD is more demanding.[7,10,15,33–35] This is especially the case for the motions of methyl groups in side-chains because their internal dynamics results from the convolution of fast and slow dynamics on time scales that range from ps to many ns and sometimes beyond. For example, measured NMR deuterium relaxation rates of $^{13}C\,H_2\,^2H$ methyl groups in ^{13}C-, fractionally ^2H-labeled proteins report on reorientational motions of ^{13}C–^2H bond vectors on time scales faster than the overall tumbling of the molecule, including bond librations and jumps between rotamer states of the respective side-chain χ dihedral angles that connect the methyl group to the backbone. The presence of such additional motions on slower time scales is well-documented and confirmed by scalar coupling measurements.[36,37]

Although considered difficult,[9,38,39] quantitatively capturing the time scales of methyl rotation by MD simulation is an area where comparison to experiment has been very fruitful. For example, for alanine residues in the CHARMM22 force field, improvements were made by reducing the methyl rotation barriers.[38,39] Similarly, rotation of the methyl groups around their 3-fold axis in the AMBER ff99SB*-ILDN[27–29] force field was improved by adjusting the methyl dihedral angle parameters against high-level coupled cluster CCSD(T) electronic structure calculations,[14] which are considered a gold standard in quantum chemistry.[40,41] Here, we extend our previous work[14] and reparametrize the energy barriers of methyl group rotations in the AMBER ff15ipq[30] and CHARMM36[25] protein force fields. To check and validate the revised force field parameters, we apply these two reparametrized force fields and compare them to our previous results[14] obtained with the AMBER ff99SB*-ILDN force field with revised methyl rotation barriers. For this purpose, the small globular protein ubiquitin was used as a model system, since NMR relaxation data are available for both backbone and side-chains. Our simulations demonstrate that the dynamics of backbone amide bonds and of side-chain methyl groups can be nearly quantitatively predicted by the MD simulations.

Theoretical Background

Relaxation rates depend on the spectral density function at specific Larmor frequencies ω_N, ω_H and ω_D of ^{15}N, ^{1}H and ^{2}H, respectively. For amide ^{15}N–^{1}H bonds in the protein backbone, the relaxation rates R_1 and R_2, which are the inverse of the longitudinal (or spin-lattice) relaxation time T_1 and of the transverse (or spin-spin) relaxation time T_2, respectively, and the ^{15}N{^{1}H} steady-state heteronuclear NOE are given by[42]

$$R_1 = f_{\text{DD}} \left[J(\omega_\text{H} - \omega_\text{N}) + 3J(\omega_\text{N}) + 6J(\omega_\text{H} + \omega_\text{N}) \right] + f_{\text{CSA}} J(\omega_\text{N}) \tag{1}$$

$$R_2 = \frac{1}{2} f_{\text{DD}} \left[4J(0) + J(\omega_\text{H} - \omega_\text{N}) + 3J(\omega_\text{N}) + 6J(\omega_\text{H}) + 6J(\omega_\text{H} + \omega_\text{N}) \right]$$
$$+ \frac{1}{6} f_{\text{CSA}} \left[4J(0) + 3J(\omega_\text{N}) \right] + R_{\text{ex}} \tag{2}$$

$$\text{NOE} = 1 + \frac{f_{\text{DD}}}{R_1} \frac{\gamma_\text{H}}{\gamma_\text{N}} \left[6J(\omega_\text{H} + \omega_\text{N}) - J(\omega_\text{H} - \omega_\text{N}) \right] \tag{3}$$

In these equations, f_{DD} and f_{CSA} are prefactors from dipole–dipole interactions of the involved spins and from the chemical shift anisotropy (CSA) tensor

$$f_{\text{DD}} = \frac{1}{4} \left(\frac{\mu_0 \hbar \gamma_\text{N} \gamma_\text{H}}{4\pi r_{\text{NH}}^3} \right)^2 \tag{4}$$

$$f_{\text{CSA}} = \frac{1}{3} \omega_\text{N}^2 \Delta_{\text{CSA}}^2 \tag{5}$$

with vacuum permeability μ_0, reduced Planck constant $\hbar$, r_{NH} the amide bond length, γ_N and γ_H the gyromagnetic ratios of ^{15}N and ^{1}H, respectively, and $\Delta_{\text{CSA}} = -170$ ppm the ^{15}N chemical shift anisotropy. R_{ex} in eq 2 is the exchange contribution to R_2. Note that unlike Chen and coworkers[32] we included the prefactor of 2/5 in the $J(\omega)$ definition,[43] rather than into f_{DD} and f_{CSA}. Although the CSA is often assumed to be uniform along the protein backbone, several reports have shown that variations exist, and in some instances these may be large.[44,45] This fact is often not considered. In addition, R_2 may contain contributions from microsecond to millisecond time scale dynamics that modulate the isotropic shielding

constant. Their presence needs confirmation or separation through additional experiments.[46]

While spin relaxation of the ^{15}N nucleus is mainly driven by its dipolar coupling to the bonded ^{1}H and through ^{15}N chemical shift anisotropy, deuterium relaxation is solely due to interactions between the electric field gradient eq and the ^{2}H nuclear quadrupole moment eQ, and this interaction varies by less than 0.5 % between methyl groups in proteins.[47] The longitudinal and in-phase transverse relaxation rates of ^{13}C–^{2}H bonds in methyl groups are given by

$$R(D_z) = \frac{3}{16} \left(\frac{qQe^2}{\hbar} \right)^2 [J(\omega_D) + 4J(2\omega_D)] \text{ and} \tag{6}$$

$$R(D_y) = \frac{1}{32} \left(\frac{qQe^2}{\hbar} \right)^2 [9J(0) + 15J(\omega_D) + 6J(2\omega_D)] , \tag{7}$$

respectively. The spectral density $J(\omega)$ is the Fourier transform (power spectrum) of the time correlation function (TCF)

$$J(\omega) = \int_0^\infty C(t)\, e^{-i\omega t}\, dt \tag{8}$$

The total TCF includes the dynamics of all motions of the bond relative to the external magnetic field $\vec{B}_0$. In general, three types of motions can be extracted: the global motion of the protein represented by an overall TCF $C_O(t)$, the internal motion of the bond represented by an internal TCF $C_{\text{int}}(t)$, and quantum mechanical zero-point vibrations represented by a constant factor ζ. The latter describes motions from quantum effects that cannot be captured in a classical MD simulation. The zero-point vibration affects the TCF of amide bonds and was captured by setting ζ to $(1.02/1.041)^6 \approx 0.885$.[48] For ^{13}C–^{2}H bonds of methyl groups $\zeta = 1$. If all three motions are statistically independent, the total TCF factorizes as

$$C(t) = \zeta \cdot C_O(t) \cdot C_{\text{int}}(t) \tag{9}$$

The TCF of overall motion $C_O(t)$ describes that the bond vector moves together with the

entire protein. If the protein has a non-spherical shape, diffusion has to be described by an anisotropic diffusion tensor. For most proteins, an axially symmetric tensor is sufficient.[49] In this model, the diffusion tensor is represented by the values along directions parallel $(D_\parallel)$ and perpendicular $(D_\perp)$ to the principal axes of the diffusion tensor. These values can be extracted from MD simulations by calculating the time scales of exponential fits of the eigenvalues of the rotation matrix with respect to a reference structure. The values can be converted to components of the overall TCF

$$C_{\mathrm{O}}(t) = \sum_{i=1}^{3} d_{\mathrm{i}}\, e^{-D_{\mathrm{i}} t} \tag{10}$$

with $D_1 = 6D_\perp$, $D_2 = 5D_\perp + D_\parallel$, and $D_3 = 2D_\perp + 4D_\parallel$. The prefactors d_{i} include the component $\mu_\parallel$ of the N–H or C–C$_{\mathrm{methyl}}$ (S—C$_{\mathrm{methyl}}$ for MET) bond vector $\vec{\mu}$ in the direction of the parallel axis of the principal axis frame (PAF) of the protein, $d_1 = (3\mu_\parallel^2 - 1)^2/4$, $d_2 = 3\mu_\parallel^2(1 - \mu_\parallel^2)$, and $d_3 = 3(1 - \mu_\parallel^2)^2/4$. If the protein has a spherical shape, the overall time correlation function is described by a single exponential decay

$$C_O(t) = \frac{1}{5} e^{-t/\tau_{\mathrm{R}}} \tag{11}$$

with $\tau_R = 1/(6D_{\mathrm{iso}})$. The internal TCF $C_{\mathrm{int}}(t)$ results from a superposition of several motions. For many amide bonds in structured regions, internal motions are dominated by fast librational motions. Bonds in disordered regions can have motions on a broader range of time scales. The dynamics of methyl bonds in long side-chains are often complex, as they include fast librational motions, rotation of the methyl group around its symmetry axis, and rotamer jumps around side-chain χ dihedral angles. Assuming that these motions are uncoupled, they can be described by exponential functions in the internal TCF

$$C_{\mathrm{int}}(t) = S^2 + \sum_{i=1}^{N} A_{\mathrm{i}}\, e^{-t/\tau_{\mathrm{i}}} \tag{12}$$

The two-parameter Lipari-Szabo (LS2) model[20,21] is commonly used to fit NMR relaxation data. In its simplest form, the internal TCF is fitted with two parameters, S^2 and the correlation time τ_f. This corresponds to $N = 1$, $A_1 = (1 - S^2)$, $\tau_1 = \tau_f$ in eq 12. For the determination of the methyl axis order parameter S^2_{axis}, one replaces S^2 by $S^2_{\text{axis}}/9$, which is the theoretical value obtained from complete motional averaging due to rotation around the 3-fold axis, assuming ideal tetrahedrality. Eq 12 shows the general definition of the order parameter as the long-time limit of the internal TCF

$$S^2 = \lim_{t \to \infty} C_{\text{int}}(t) \tag{13}$$

In practice, NMR relaxation experiments can only detect motions on time scales up to approximately τ_R. Furthermore, it is not possible to separate overall and internal motions in a solution NMR relaxation experiment, re-emphasizing the necessity of a simplified motional model such as the Lipari-Szabo formalism. MD simulations, however, are able to separate these motions. In MD, the TCFs can be calculated from the orientation of the vector at different times via

$$C(t) = \langle P_2 \left[\vec{\mu}(0) \cdot \vec{\mu}(t) \right] \rangle \tag{14}$$

In this formula, $P_2[x] = (3x^2 - 1)/2$ is the second Legendre polynomial, $\vec{\mu}(t)$ the unit vector of the bond at time t and $\langle \ldots \rangle$ indicates the average over all time step differences t. The total TCF describes the reorientation motion of the bond in the laboratory frame. Internal motions are obtained by superimposing the trajectory frames to a reference structure. Combining eqs 8, 9, 10 and 12 yields the spectral density

$$J(\omega) = \zeta \sum_{i=1}^{3} d_i \left[\frac{S^2 D_i}{D_i^2 + \omega^2} + \sum_{j=1}^{N} \frac{A_j(D_i + 1/\tau_i)}{(D_i + 1/\tau_i)^2 + \omega^2} \right] \tag{15}$$

Computational Methods

Reparametrization of methyl rotation energy barriers

We adjusted the force constants (amplitudes) of the dihedral angle potential energy terms describing the rotation of methyl groups of all methyl-containing amino acid side-chains in the AMBER ff15ipq[30] and CHARMM36[25] protein force fields. The same procedure was adopted as in our previous work,[14] in which we reparametrized the methyl rotation energy barriers in the AMBER ff99SB*-ILDN force field. In brief, relaxed potential energy surface (PES) scans around the dihedral angles describing methyl rotation were performed for blocked dipeptides in vacuo, ACE-X-NME, of the methyl-containing amino acids (X = ALA, MET, THR, LEU, ILE, VAL) to obtain the adiabatic energy profiles. These relaxed PES scans were performed at the M06-2X/pVDZ density functional level of theory.[50–52] For selected structures of the minima and several structures close to the maxima, single-point energies were obtained from CCSD(T)/cc-pVTZ coupled cluster calculations, yielding the reference energy barriers that were matched as closely as possible during force field reparametrization. To obtain the corresponding PES for methyl rotation with the force field, energy minimizations of the blocked dipeptides in vacuo were carried out using GROMACS version 5.0.6[53] compiled in double precision.

MD simulations

MD simulations were carried out with GROMACS version 5.0.6.[53] The setup was identical to our previous ubiquitin simulations with the ff99SB*-ILDN/TIP4P-2005 force field.[14] The X-ray crystal structure of ubiquitin (PDB 1UBQ)[54] was centered in a cubic periodic simulation cell with box vector lengths of ca. 7.94 nm. We note in passing that the finite box-size correction to the rotational diffusion constant, $k_B T/(6\eta V_{\text{box}})$ with solvent viscosity η and V_{box} the volume of the periodic simulation box,[55] is only $0.16 \cdot 10^7 \text{s}^{-1}$ in our case due to the large volume of our cubic simulation box (ca. 500 nm^3). This finite size correction

was neglected because it is smaller than the statistical uncertainties from our simulations. H-atoms were added to the protein and crystal waters, and the system was solvated with 16,235 water molecules. For the simulations with the AMBER ff15ipq force field, the $SPCE_b$ water model[56] was used, whereas the TIP3P water model[57] was used together with the CHARMM36 protein force field. These water models were chosen because the respective protein force fields were developed together with these water models. Sodium and chloride ions were added at a concentration of 0.15 mol/l such that the overall charge of the simulation box was zero. Prior to the production simulations, the system was energy-minimized and subsequently equilibrated in the NPT ensemble for 200 ps with harmonic position restraints on the protein. To keep the temperature constant at 300 K, the thermostat of Bussi and coworkers[58] was applied, with coupling time constants of $\tau_T = 0.1$ ps. Isotropic Parrinello-Rahman pressure coupling was used to maintain the pressure constant at 1 bar. The SETTLE[59] and LINCS[60] constraint algorithms were applied to constrain the internal degrees of freedom of water molecules and the bonds in the protein, respectively, allowing for integrating the equations of motion with 2 fs time steps. Lennard-Jones (6,12) interactions were smoothly shifted to zero at a cut-off distance of 1.0 nm; this cut-off distance was also used for distinguishing between short-range and long-range Coulomb interactions, which were treated with the particle mesh Ewald (PME) method[61] with a 0.12 nm grid spacing. Finally, for both ff15ipq and CHARMM36, ten 100 ns NPT production MD simulations were initiated from different random seeds for generating the initial velocities at 300 K. Coordinates were saved to disk every 1 ps.

Computation of NMR relaxation data

For the overall TCFs, $C_O(t)$, the rotation matrix between the starting structure of the MD simulations and the instantaneous structure at each trajectory frame was calculated. Rotations in the reference frame between time differences Δt were calculated in quaternion representation as described by Chen and coworkers.[62] Exponential fits to the auto correla-

tion functions of the diagonal elements of this matrix lead to time constants τ whose inverse are the components D_i (i=1,2,3) in eq 10. For analyses, the ten 100 ns trajectories were divided into parts of 10 ns. For each of these chunks, the TCFs were calculated up to a lag time of 5 ns to yield $C_O(t)$. For the calculations of the prefactors d_i, the axial component of the N–H and C–C_{methyl} bonds (S–C_{methyl} for MET) in the PAF of the protein was calculated.

Internal TCFs, $C_{int}(t)$, of the backbone amide N–H bonds and for one of the three C_{methyl}–H_{methyl} bonds of every methyl group were calculated up to a maximum lag time of 5 ns after aligning all trajectory frames with the initial structure. We used a spectral density approach for calculating the relaxation rates. In this approach, the internal TCFs of the amide or methyl vectors were fitted with eq 12. To automatically determine the number of parameters N in eq 12, N was incremented based on the difference between the raw TCF C_{int}^{raw} and the fitted TCF C_{int}^{fit}

$$\chi = \frac{C_{int}^{fit} - C_{int}^{raw}}{\Delta C_{fit}} \tag{16}$$

N was increased until χ did not decrease further or an upper bound of $N = 4$ was reached. The error ΔC_{fit} was calculated from the standard deviation of the 100 trajectory chunks that were used to fit C_{int}.

Total TCFs were calculated with eq 9 using $\zeta = (1.02/1.041)^6$ for amide bonds[48] and $\zeta = 1$ for methyl bonds. Spectral densities were calculated with eq 15. Relaxation rates and NOEs were calculated from these spectral densities using eqs 1 to 3 for amide bonds and eqs 6 and 7 for methyl bonds.

Results and Discussion

Reparametrization of methyl group rotation barriers

The height of the potential energy barrier for methyl group rotation heavily affects the dynamics of the C–H methyl bonds and, as a consequence, the time correlation functions,

spectral densities, and NMR relaxation rates. We reparametrized the methyl group rotation barriers in the AMBER ff15ipq and CHARMM36 force fields by adjusting the force constants k_{dih} of the dihedral angle potential energy terms that govern the rotation of the methyl group around its 3-fold symmetry axis. As described in our previous work[14] on the AMBER ff99SB*-ILDN force field, the force constants k_{dih} were adjusted to closely match the potential energy barriers obtained from coupled cluster CCSD(T) quantum chemical calculations of isolated blocked dipeptides.

Table 1 compares the potential energy barriers of methyl group rotation from the force fields to the reference values obtained from the coupled cluster calculations (the potential energy profiles for all methyls are shown in Fig. S1 in Supporting Information). Generally, the energy barriers in the force fields are too high, with the exception of the MET methyl group in the AMBER ff15ipq force field. The energy barriers in the original force fields deviate from the reference values to different extents for the different methyl groups, and the barriers were thus adjusted in a methyl group-specific manner (Table 2). The force field reparametrization reduces the mean unsigned error with respect to the CCSD(T) energy barriers from about 4 kJ/mol to 0.2 kJ/mol for both force fields.

Table 1: Potential energy barriers (in kJ/mol) of methyl group rotation in blocked dipeptides.

	CCSD(T)	ff15ipq (orig.)	ff15ipq (repar.)	CHARMM36 (orig.)	CHARMM36 (repar.)
ALA	14.2	22.0	14.3	18.4	14.3
MET	7.1	7.0	7.1	7.7	7.1
THR	11.4	15.4	11.4	17.9	11.5
VAL[a]	14.0/11.5	17.7/16.7	13.4/12.3	18.8/17.1	13.9/12.0
LEU[a]	14.1/12.9	16.4/15.5	14.0/13.0	16.7/16.8	13.5/13.6
ILE[b]	12.2/10.7	16.3/14.5	12.4/10.9	17.6/12.8	12.4/10.6
MUE[c]	–	3.8	0.2	3.9	0.2

[a] The values before and after the slash are for the trans/gauche$^-$ methyl groups, respectively. [b] The values before and after the slash are for the C^γ/C^δ methyl groups, respectively. [c] Mean unsigned error with respect to the CCSD(T)/cc-pVTZ energies, averaged over all methyl groups.

In our previous study,[14] we did not reparametrize rotation of THR methyl groups because

Table 2: Difference of the force constants, Δk_{dih}, of the dihedral energy terms for methyl group rotation between the reparametrized and the original force fields (in kJ/mol).

Methyl	ff15ipq	CHARMM36
ALA C^{β}	-0.42987	-0.23278
MET C^{ϵ}	+0.02965	-0.09508
VAL C^{γ}	-0.24593	-0.28615
LEU C^{δ}	-0.13638	-0.18106
ILE C^{γ}	-0.22447	-0.30111
ILE C^{δ}	-0.21158	-0.11778
THR C^{γ}	-0.33333	-0.54596

they are reasonably well described by the original ff99SB*-ILDN force field. As a larger error in the THR energy barrier for AMBER ff15ipq and CHARMM36 was observed, we also adjusted k_{dih} for THR in these force fields. To this end, the force constants of the C_{α}-C_{β}-C_{methyl}-$H_{\mathrm{methyl}}^{1,2,3}$ and the H_{β}-C_{β}-C_{methyl}-$H_{\mathrm{methyl}}^{1,2,3}$ dihedral terms were reduced, see Table 2. The dihedral angle terms involving the O_{γ}-atom of THR were left untouched.

NMR relaxation from MD simulations

The reparametrized force fields were systematically investigated and compared in terms of their dynamics, as described by NMR relaxation data. To that end, ten 100 ns MD simulations of ubiquitin were performed with the AMBER ff15ipq and CHARMM36 force fields, including the adjusted methyl rotation barriers described above. In the ff15ipq simulations the $SPCE_b$ water model[56] was used, whereas the TIP3P water model[57] was used in the CHARMM36 simulations. We compare our results to our previous ubiquitin simulations,[14] in which we demonstrated for the AMBER ff99SB*-ILDN protein force field in combination with the TIP4P-2005[63] water model that accurate methyl rotation energy barriers in the protein force field are essential for an accurate prediction of deuterium relaxation from MD simulations. For this reason, we did not carry out the MD simulations with the original ff15ipq and CHARMM36 force fields (*i.e.*, with too high methyl rotation barriers), but instead focus on the comparison of the different force fields with revised methyl rotation

13

barriers.

Rotational diffusion

The time scale of rotational diffusion in MD simulations strongly depends on the water model. [15,56,62,64–67] While the water model usually does not have a strong effect on the dynamics of backbone N–H bonds, it changes the time scale of overall rotational tumbling and therefore the spectral densities. Table 3 summarizes the rotational diffusion constants D_{iso} obtained from our simulations with the different water models. The isotropic diffusion constant of $3.92 \cdot 10^7 \mathrm{s}^{-1}$ from the MD simulations in TIP4P-2005 water is close to the diffusion constants obtained from NMR relaxation experiments, which yielded values of $4.76 \cdot 10^7 \mathrm{s}^{-1}$, [68] $4.14 \cdot 10^7 \mathrm{s}^{-1}$, [69] $4.07 \cdot 10^7 \mathrm{s}^{-1}$, [70] and $4.01 \cdot 10^7 \mathrm{s}^{-1}$. [71] Thus, the overall protein tumbling is described at the correct time scale in TIP4P-2005 water, a basic requirement for a quantitative description of $C_{\mathrm{O}}(t)$ and the prediction of NMR relaxation rates from MD simulations.

The isotropic diffusion constant from the simulations in $SPCE_{\mathrm{b}}$ water is $3.56 \cdot 10^7 \mathrm{s}^{-1}$, i.e., the protein tumbles a little bit slower than in simulations with the TIP4P-2005 water model. Protein tumbling in TIP3P water is much faster $(10.71 \cdot 10^7 \mathrm{s}^{-1})$, in line with the much higher translational self-diffusion constant of TIP3P water. The corresponding tumbling correlation times $\tau_{\mathrm{R}} = 1/(6D_{\mathrm{iso}})$ are 4.25 ns, 4.68 ns, and 1.56 ns in TIP4P-2005, $SPCE_{\mathrm{b}}$, and TIP3P water, respectively.

Table 3: Rotational diffusion of ubiquitin in MD simulations.

	$D_{\mathrm{iso}} \ [10^7 \mathrm{s}^{-1}]$
ff99SB*-ILDN/TIP4P-2005	3.92 ± 0.18
ff15ipq/SPCE$_{\mathrm{b}}$	3.56 ± 0.42
CHARMM36/TIP3P	10.71 ± 0.87

Overall tumbling of the protein is slow compared to most of its internal dynamics and mainly affects spectral densities at low frequencies. This has implications for the interpretation of NMR relaxation data. The computed relaxation rates that depend on $J(0)$, such as R_2 for backbone amides or $R(D_{\mathrm{y}})$ for side-chain methyls, are expected to be affected

strongly by the water model used in the simulations, as this determines the time scale of protein rotational diffusion.

Backbone relaxation

We now turn to the computation of ^{15}N relaxation rates and ^{15}N{^{1}H} NOEs of backbone N–H bonds, as shown in Figure 1. The three force fields compared here correctly predict the site-to-site variation. For example, all relaxation rates are higher in the β-sheet regions comprising residues 2-7, 11-15, and 41-45, and especially also in the α-helix formed by residues 23-34. Significantly lower values are observed at the turn residues 7-10 and in the unstructured region around GLY35 (with an exception for the CHARMM36 force field in R_2). Taken together, all three force fields yield a consistent picture that is in agreement with the experimental data, see Table 4. We conclude that the backbone dynamics of ubiquitin is accurately described by all force fields used.

Table 4: Agreement between computational and experimental backbone relaxation data.

RMSD	ff15ipq/SPCE$_b$	ff99SB*-ILDN/TIP4P-2005	CHARMM36/TIP3P^a
^{15}N R_1 [s^{-1}]	0.28	0.17	0.14/0.17
^{15}N R_2 [s^{-1}]	0.47	0.54	3.03/0.46
^{15}N{^{1}H} NOE	0.07	0.06	0.32/0.09
Pearson coefficient R_P			
^{15}N R_1	0.88	0.93	0.93/0.94
^{15}N R_2	0.89	0.90	0.91/0.92
^{15}N{^{1}H} NOE	0.99	0.98	0.99/0.99

^{a} The values before and after the slash correspond to the unscaled and scaled rotational diffusion times, respectively.

Averaged over all amide bonds, the longitudinal relaxation rate R_1 in the AMBER ff15ipq force field of 1.82 s^{-1} is slightly lower than the values of 1.92 s^{-1} and 1.97 s^{-1} for the AMBER ff99SB*-ILDN and the CHARMM36 force field, respectively. Figure 2 shows the separation of the relaxation rates into the different contributions according to eq 1. The deviation in the relaxation rates between AMBER ff15ipq and the other force fields is mainly due to $J(\omega_N)$, which makes a dominant contribution to R_1. The frequency ω_N is the lowest frequency
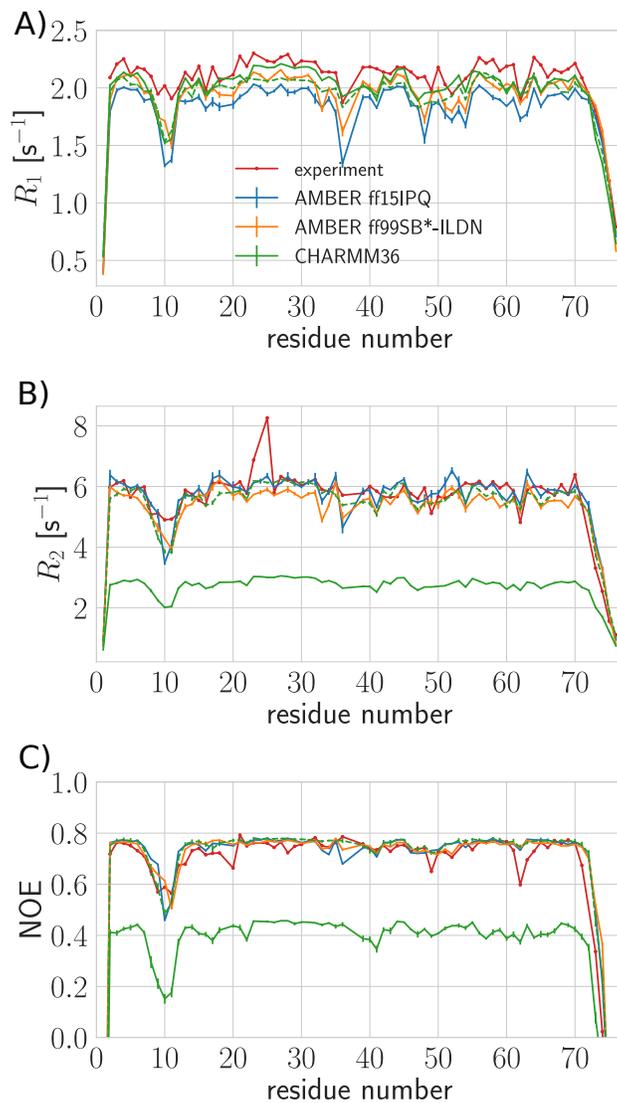
Figure 1: ^{15}N relaxation rates R_1 (A) and R_2 (B), and ^{15}N{^{1}H} NOE (C) from MD simulations with the AMBER ff99SB*-ILDN force field (orange), AMBER ff15ipq force field (blue), and CHARMM36 force field (solid and dashed green lines for unscaled and scaled rotational diffusion, respectively) are compared to the experimental values (red) from Tjandra and coworkers.[70]

that influences R_1, pointing towards faster internal dynamics, at least at this frequency, in the ff15ipq force field than in the other two force fields. The average R_1 from ^{15}N NMR relaxation of ubiquitin is even slightly higher, 2.09 s^{-1}.[70]

For the transverse relaxation rate R_2, the bond-averaged relaxation rates are 5.58 s^{-1} and 5.29 s^{-1} for the AMBER ff15ipq and ff99SB*-ILDN force fields, which are in good agreement with the experimental value of 5.65 s^{-1}.[70] The discrepancy between simulations and experiment for ILE23 and ASN25 are due to chemical exchange on slow μs–ms time scales,[72] because the R_{ex} contribution to R_2 (see eq 2) is not captured in the MD simulations. For CHARMM36, the average R_2 is much lower, 2.68 s^{-1}. The underestimation of this rate can be attributed to the TIP3P water model used in the CHARMM36 simulations, in which protein tumbling is too fast (see above). Tumbling is much slower than the internal dynamics of backbone vectors in structured protein regions and therefore primarily influences slow dynamics at low frequencies, which affect R_2 via $J(0)$, see eq 2. Faster tumbling leads to lower values of $J(0)$ and R_2. One possible way to overcome too fast rotational protein tumbling would be to perform the CHARMM36 simulations in combination with another water model that yields a more realistic rotational diffusion time, such as TIP4P-2005 or SPCE$_b$, as done for the AMBER force fields. However, this was not done in this case because the CHARMM36 protein force field was developed together with TIP3P, and thus the protein–water interactions would have to be carefully tested for a different water model, which is beyond the scope of this work. Alternatively, rotational velocity rescaling can be applied, as e.g. done by Anderson and coworkers.[65,66] In the same vein, to investigate the influence of the overall tumbling in the CHARMM36/TIP3P simulations, we rescaled the rotational diffusion time for computing $C_O(t)$ via eq 11 by a factor of 2.73, which corresponds to the ratio of the rotational diffusion constants obtained in TIP3P water and TIP4P-2005 water (Table 3). The results for the CHARMM36 simulations with this scaled overall rotational diffusion time are shown as dashed green lines in Figure 1. They are in equally good agreement with the experimental values as the simulations with the two AMBER force fields, as can also be seen

from the improved RMSD values in Table 4.

The bond-averaged ^{15}N$\{^1$H$\}$ NOEs are 0.67, 0.66, and 0.33 for ff15ipq, ff99SB*-ILDN, and unscaled CHARMM36, respectively. The experimental average NOE is 0.67, close to the simulations with the AMBER force fields. The NOE is due to the difference between $6J(\omega_H + \omega_N)$ and $J(\omega_H - \omega_N)$, with $6J(\omega_H + \omega_N)$ as the dominant contribution, see eq 3. The low NOE from the CHARMM36 simulations is again attributed to the too fast tumbling in TIP3P water, as scaling the rotational protein diffusion time improves the agreement with experiments (Figure 1 and Table 4), similar to R_2. The high correlation coefficients of about 0.99 show that all force fields correctly describe the site-to-site variation of the ^{15}N$\{^1$H$\}$ NOE.

Methyl relaxation rates

The ^{2}H relaxation rates $R(D_z)$ and $R(D_y)$ of methyl groups in protein side-chains are reported in Figure 3; the comparison with the experimental data[73] is also summarized in Table 5 and shown in Fig. S2 in Supporting Information. Overall, the agreement between MD simulations and NMR experiments is good, especially for $R(D_y)$. The scaling of the overall rotational diffusion time, as applied in the CHARMM36/TIP3P simulations, substantially improves the agreement for $R(D_y)$. Faster rotational diffusion reduces $R(D_y)$, which depends on slow dynamics via $J(0)$. The good correlation with experimental $R(D_y)$ values is reflected in the high correlation coefficients of 0.86, 0.83 and 0.83/0.90 for ff15ipq, ff99SB*-ILDN, and CHARMM36, respectively (Table 5). Taken together, our results show that slow dynamics of methyl groups on the ns time scale is well captured by the force fields with properly adjusted methyl rotation barriers.

For fast dynamics, as probed by $R(D_z)$, the correlation with experiment is rather poor, with correlation coefficients of only about 0.3 for the three force fields (Table 5). $R(D_z)$ is influenced by the spectral density at higher frequencies (ω_D, $2\omega_D$), and the simulated values fall into a narrow range between 10 s^{-1} and 40 s^{-1}, similar to experiment.[73] Differences

18
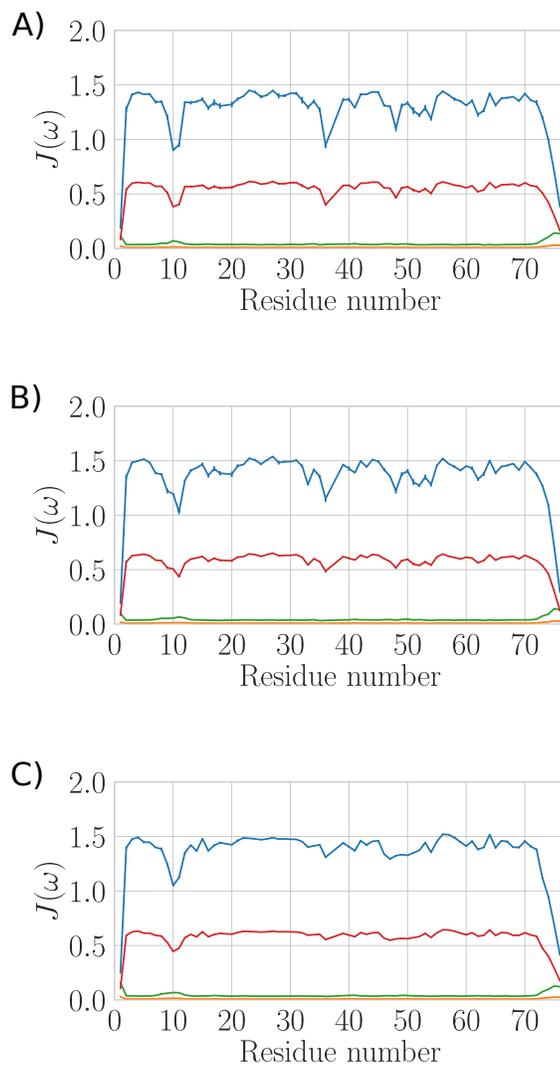
Figure 2: Different contributions to ^{15}N R_1. The plots show the values of $f_{\mathrm{DD}}3J(\omega_{\mathrm{N}})$ (blue), $f_{\mathrm{CSA}}J(\omega_{\mathrm{N}})$ (red), $f_{\mathrm{DD}}6J(\omega_{\mathrm{H}}+\omega_{\mathrm{N}})$ (green) and $f_{\mathrm{DD}}J(\omega_{\mathrm{H}}-\omega_{\mathrm{N}})$ (orange) from the simulations with AMBER ff15ipq/SPCE$_{\mathrm{b}}$ (A), AMBER ff99SB*-ILDN/TIP4P-2005 (B) and CHARMM36/TIP3P (C) force fields. For CHARMM36/TIP3P, the results obtained with scaled overall rotational diffusion are shown.
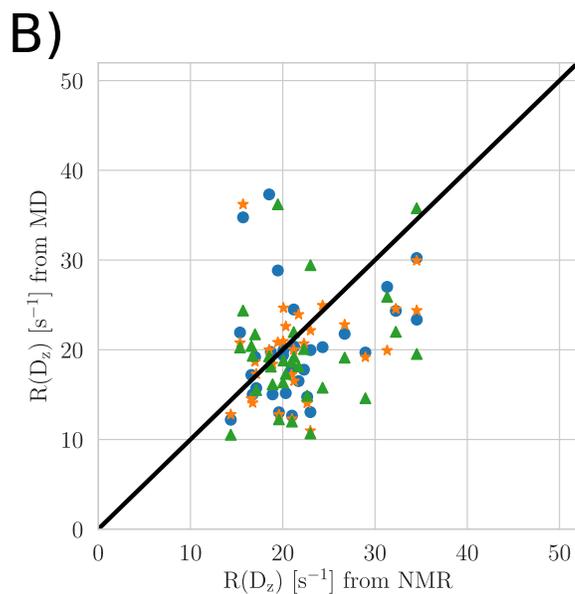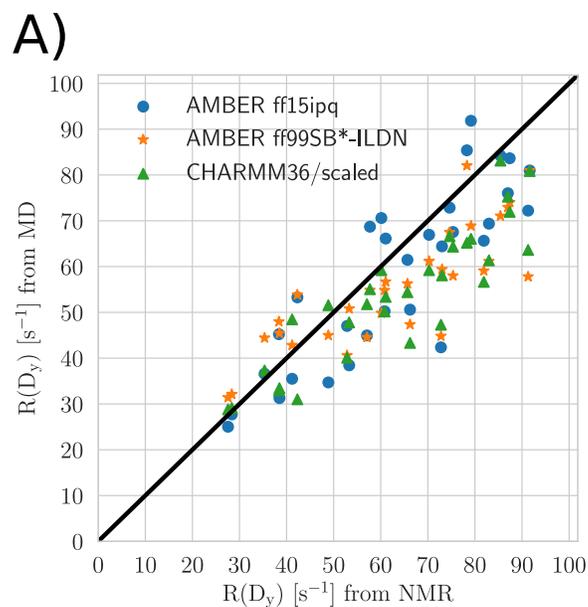
Figure 3: Methyl deuterium relaxation rates $R(D_y)$ (A) and $R(D_z)$ (B) from MD simulations with the AMBER ff99SB*-ILDN force field (orange stars), AMBER ff15ipq force field (blue circles) and the CHARMM36 force field with scaled overall rotational diffusion (green triangles) are correlated to the experimental values from Liao and coworkers.[73]

Table 5: Agreement between computational and experimental deuterium relaxation data of side-chain methyl groups.

RMSD	ff15ipq/SPCE$_b$	ff99SB*-ILDN/TIP4P-2005	CHARMM36/TIP3P[a]
^{2}H $R(D_y)$ [s^{-1}]	11.1	13.5	28.9/12.9
^{2}H $R(D_z)$ [s^{-1}]	7.2	6.5	7.5/7.2
S^2_{axis} (from LS2 model)	0.13	0.12	0.10/0.10
Pearson coefficient R_P			
^{2}H $R(D_y)$	0.86	0.83	0.83/0.90
^{2}H $R(D_z)$	0.26	0.32	0.27/0.29
S^2_{axis} (from LS2 model)	0.85	0.89	0.93/0.93

[a] The values before and after the slash correspond to the unscaled and scaled rotational diffusion times, respectively.

between the fast dynamics on the ps time scale of these methyl groups are thus very challenging to capture, considering the complicated dynamics especially of longer side-chains. We recently reported a much higher correlation coefficient between MD and NMR of 0.72 for $R(D_z)$ of methyl groups in T4 lysozyme,[15] showing that also fast dynamics as probed by $R(D_z)$ can in principle be described reliably in MD simulations. Notably, the RMSD between $R(D_z)$ from MD and NMR of about 7 s^{-1} for ubiquitin (Table 5) is small compared to the absolute values and in a similar range as the RMSD obtained for T4 lysozyme in our previous work, due to the reparametrization of the methyl rotation barriers.[14,15] Our results emphasize that still further improvements of biomolecular force fields are needed in order to achieve a similar correlation between experimental and computational relaxation rates for methyl groups as is routinely obtained for backbone amide bonds.

Methyl order parameters

Next, we turn to methyl order parameters and their prediction from MD simulations. Figure 4 shows the comparison of the methyl axis order parameters obtained from our MD simulations with the experimental values from Liao and coworkers.[73] As was done in the experimental study,[73] we used the LS2 model to obtain the methyl order parameters. In general, the agreement between experimental and computational S^2_{axis} is good, with a Pearson correlation coefficient and RMSD of approximately 0.9 and 0.1, respectively (Table 5).

These results are comparable to previous MD simulation studies of ubiquitin, which reported RMSDs for S_axis^2 of 0.125,[74] 0.15,[75] and 0.104.[76]
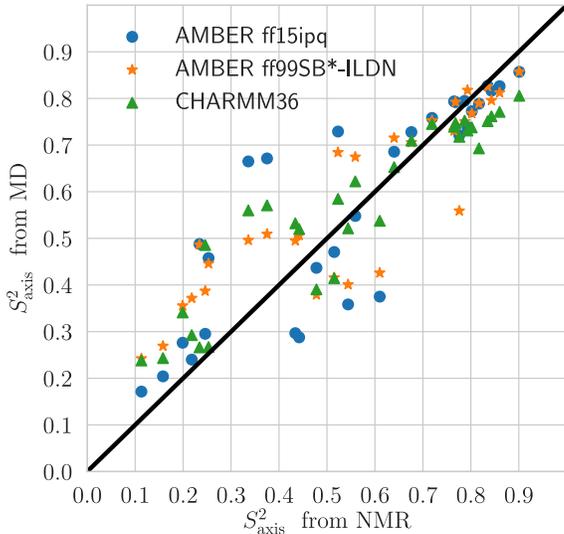

Figure 4: Methyl order parameter S_axis^2 from MD simulations with the ff99SB*-ILDN (orange stars), ff15ipq (blue circles), and CHARMM36 force field (green triangles) are correlated to the experimental values from Liao and coworkers.[73] The first residue in the chain (MET1) was excluded.

Taken together, given the different approximations involved on both the simulation side (force field accuracy, limited simulation time scales) and the NMR side (limited number of frequencies at which the spectral density function can be sampled), our results demonstrate the level of accuracy that can be expected from modern biomolecular force fields in terms of the internal dynamics of methyl groups.

Conclusions

Quantitative agreement between experimental and computational NMR relaxation data of proteins requires an accurate description of the dynamics in the simulations. Thus, both the minima and also the barriers on the energy landscape have to be accurately described. Here, we adjusted the energy barriers for the rotation of methyl groups in protein side-

chains in the AMBER ff15ipq and CHARMM36 force fields. Using ubiquitin as a model system, we systematically tested and validated these two reparametrized force fields, together with the recent AMBER ff99SB*-ILDN force field with reparametrized methyl rotations, [14] against experimental NMR relaxation data of both backbone amides and side-chain methyl groups. The site-to-site variation of amide relaxation rates is captured correctly by all force fields. However, the absolute value depends strongly on the time scale of rotational diffusion of the protein, which is determined by the water model. Furthermore, the dynamics of methyl groups in side-chains, as represented by $R(D_\mathrm{y})$ or S^2_axis, is predicted quite well by all force fields. However, predicting site-to-site variation in the exact time scale of rotation of individual methyl groups, which dominates $R(D_\mathrm{z})$, requires further improvements.

Supporting Information Available

Potential energy curves for methyl rotation in all side-chains. Computed relaxation rates of all methyl groups in comparison to experiment. Python scripts to modify the methyl group rotation barriers in Gromacs topology files (.itp or .top files) have been uploaded to our website `www.molecular-simulation.org`.

Acknowledgement

References

(1) Peng, J. W. Exposing the Moving Parts of Proteins with NMR Spectroscopy. *J. Phys. Chem. Lett.* **2012**, *3*, 1039–1051.

(2) Sapienza, P. J.; Lee, A. L. Using NMR to Study Fast Dynamics in Proteins: Methods and Applications. *Curr. Opin. Pharmacol.* **2010**, *10*, 723–730.

(3) Mittermaier, A.; Kay, L. E. New Tools Provide New Insights in NMR Studies of Protein Dynamics. *Science* **2006**, *312*, 224–228.

(4) Palmer, A. G. NMR Characterization of the Dynamics of Biomacromolecules. *Chem. Rev.* **2004**, *104*, 3623–3640.

(5) Brüschweiler, R. New Approaches to the Dynamic Interpretation and Prediction of NMR Relaxation Data from Proteins. *Curr. Opin. Struct. Biol.* **2003**, *13*, 175–183.

(6) Case, D. A. Molecular Dynamics and NMR Spin Relaxation in Proteins. *Acc. Chem. Res.* **2002**, *35*, 325–331.

(7) Skrynnikov, N. R.; Millet, O.; Kay, L. E. Deuterium Spin Probes of Side-Chain Dynamics in Proteins. 2. Spectral Density Mapping and Identification of Nanosecond Time-Scale Side-Chain Motions. *J. Am. Chem. Soc.* **2002**, *124*, 6449–6460.

(8) Lindorff-Larsen, K.; Best, R. B.; DePristo, M. A.; Dobson, C. M.; Vendruscolo, M. Simultaneous Determination of Protein Structure and Dynamics. *Nature* **2005**, *433*, 128–132.

(9) Xue, Y.; Pavlova, M. S.; Ryabov, Y. E.; Reif, B.; Skrynnikov, N. R. Methyl Rotation Barriers in Proteins from ^{2}H Relaxation Data. Implications for Protein Structure. *J. Am. Chem. Soc.* **2007**, *129*, 6827–6838.

(10) Showalter, S. A.; Johnson, E.; Rance, M.; Brüschweiler, R. Toward Quantitative Interpretation of Methyl Side-Chain Dynamics from NMR by Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **2007**, *129*, 14146–14147.

(11) Rajeshwar T., R.; Smith, J. C.; Krishnan, M. Hidden Regularity and Universal Classification of Fast Side Chain Motions in Proteins. *J. Am. Chem. Soc.* **2014**, *136*, 8590–8605.

(12) Salvi, N.; Abyzov, A.; Blackledge, M. Multi-Timescale Dynamics in Intrinsically Disordered Proteins from NMR Relaxation and Molecular Simulation. *J. Phys. Chem. Lett.* **2016**, *7*, 2483–2489.

(13) Salvi, N.; Abyzov, A.; Blackledge, M. Solvent-Dependent Segmental Dynamics in Intrinsically Disordered Proteins. *Sci. Adv.* **2019**, *5*, eaax2348.

(14) Hoffmann, F.; Mulder, F. A. A.; Schäfer, L. V. Accurate Methyl Group Dynamics in Protein Simulations with AMBER Force Fields. *J. Phys. Chem. B* **2018**, *122*, 5038–5048.

(15) Hoffmann, F.; Xue, M.; Schäfer, L. V.; Mulder, F. A. A. Narrowing the Gap Between Experimental and Computational Determination of Methyl Group Dynamics in Proteins. *Phys. Chem. Chem. Phys.* **2018**, *20*, 24577–24590.

(16) Möckel, C.; Kubiak, J.; Schillinger, O.; Kühnemuth, R.; Della Corte, D.; Schröder, G. F.; Willbold, D.; Strodel, B.; Seidel, C. A. M.; Neudecker, P. Integrated NMR, Fluorescence, and Molecular Dynamics Benchmark Study of Protein Mechanics and Hydrodynamics. *J. Phys. Chem. B* **2019**, *123*, 1453–1480.

(17) Rashid, S.; Lee, B. L.; Wajda, B.; Spyracopoulos, L. Side-Chain Dynamics of the Trifluoroacetone Cysteine Derivative Characterized by ^{19}F NMR Relaxation and Molecular Dynamics Simulations. *J. Phys. Chem. B* **2019**, *123*, 3665–3671.

(18) Tchaicheeyan, O.; Mendelman, N.; Zerbetto, M.; Meirovitch, E. Local Ordering at Mobile Sites in Proteins: Combining Perspectives from NMR Relaxation and Molecular Dynamics. *J. Phys. Chem. B* **2019**, *123*, 2745–2755.

(19) Smith, A. A.; Ernst, M.; Riniker, S.; Meier, B. H. Localized and Collective Motions in HET-s(218-289) Fibrils from Combined NMR Relaxation and MD Simulation. *Angew. Chem. Int. Ed.* **2019**, *58*, 9383–9388.

(20) Lipari, G.; Szabo, A. Model-Free Approach to the Interpretation of Nuclear Magnetic Resonance Relaxation in Macromolecules. 1. Theory and Range of Validity. *J. Am. Chem. Soc.* **1982**, *104*, 4546–4559.

(21) Lipari, G.; Szabo, A. Model-Free Approach to the Interpretation of Nuclear Magnetic Resonance Relaxation in Macromolecules. 2. Analysis of Experimental Results. *J. Am. Chem. Soc.* **1982**, *104*, 4559–4570.

(22) Millet, O.; Muhandiram, D. R.; Skrynnikov, N. R.; Kay, L. E. Deuterium Spin Probes of Side-Chain Dynamics in Proteins. 1. Measurement of Five Relaxation Rates per Deuteron in ^{13}C-Labeled and Fractionally ^{2}H-Enriched Proteins in Solution. *J. Am. Chem. Soc.* **2002**, *124*, 6439–6448.

(23) Redfield, A. G. Shuttling Device for High-Resolution Measurements of Relaxation and Related Phenomena in Solution at Low Field, using a Shared Commercial 500 MHz NMR Instrument. *Magn. Reson. Chem.* **2003**, *41*, 753–768.

(24) Cousin, S. F.; Kaderávek, P.; Bolik-Coulon, N.; Gu, Y.; Charlier, C.; Carlier, L.; Bruschweiler-Li, L.; Marquardsen, T.; Tyburn, J.-M.; Brüschweiler, R.,; Ferrage, F. Time-Resolved Protein Side-Chain Motions Unraveled by High-Resolution Relaxometry and Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **2018**, *140*, 13456–13465.

(25) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, A. D. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Im-

proved Sampling of the Backbone Ψ, Φ and Side-Chain $\chi 1$ and $\chi 2$ Dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8*, 3257–3273.

(26) Huang, J.; MacKerell, A. D. CHARMM36 All-atom Additive Protein Force Field: Validation Based on Comparison to NMR Data. *J. Comput. Chem.* **2013**, *34*, 2135–2145.

(27) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins* **2006**, *65*, 712–725.

(28) Best, R. B.; Hummer, G. Optimized Molecular Dynamics Force Fields Applied to the Helix–Coil Transition of Polypeptides. *J. Phys. Chem. B* **2009**, *113*, 9004–9015.

(29) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved Side-Chain Torsion Potentials for the Amber ff99SB Protein Force Field. *Proteins* **2010**, *78*, 1950–1958.

(30) Debiec, K. T.; Cerutti, D. S.; Baker, L. R.; Gronenborn, A. M.; Case, D. A.; Chong, L. T. Further Along the Road Less Traveled: AMBER ff15ipq, an Original Protein Force Field Built on a Self-Consistent Physical Model. *J. Chem. Theory Comput.* **2016**, *12*, 3926–3947.

(31) Wang, L.-P.; McKiernan, K. A.; Gomes, J.; Beauchamp, K. A.; Head-Gordon, T.; Rice, J. E.; Swope, W. C.; Martinez, T. J.; Pande, V. S. Building a More Predictive Protein Force Field: A Systematic and Reproducible Route to AMBER-FB15. *J. Phys. Chem. B* **2017**, *121*, 4023–4039.

(32) Chen, P.-c.; Hologne, M.; Walker, O.; Hennig, J. Ab Initio Prediction of NMR Spin Relaxation Parameters from Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2018**, *14*, 1009–1019.

(33) Bremi, T.; Brüschweiler, R.; Ernst, R. R. A Protocol for the Interpretation of Side-Chain Dynamics Based on NMR Relaxation: Application to Phenylalanines in Antamanide. *J. Am. Chem. Soc.* **1997**, *119*, 4272–4284.

(34) Best, R. B.; Clarke, J.; Karplus, M. What Contributions to Protein Side-Chain Dynamics are Probed by NMR Experiments? A Molecular Dynamics Simulation Analysis. *J. Mol. Biol.* **2005**, *349*, 185–203.

(35) O´Brien, E. S.; Wand, A. J.; Sharp, K. A. On the Ability of Molecular Dynamics Force Fields to Recapitulate NMR Derived Protein Side Chain NMR Order Parameters. *Protein Sci.* **2016**, *25*, 1156–1160.

(36) Chou, J. J.; Case, D. A.; Bax, A. Insights into the Mobility of Methyl-Bearing Side Chains in Proteins from $^3J_{CC}$ and $^3J_{CN}$ couplings. *J. Am. Chem. Soc.* **2003**, *125*, 8959–8966.

(37) Lindorff-Larsen, K.; Best, R. B.; Vendruscolo, M. Interpreting Dynamically-Averaged Scalar Couplings in Proteins. *J. Biomol. NMR* **2005**, *32*, 273–280.

(38) Chatfield, D. C.; Szabo, A.; Brooks, B. R. Molecular Dynamics of Staphylococcal Nuclease: Comparison of Simulation with ^{15}N and ^{13}C NMR Relaxation Data. *J. Am. Chem. Soc.* **1998**, *120*, 5301–5311.

(39) Chatfield, D. C.; Wong, S. E. Methyl Motional Parameters in Crystalline L-Alanine: Molecular Dynamics Simulation and NMR. *J. Phys. Chem. B* **2000**, *104*, 11342–11348.

(40) Bartlett, R. J.; Musiał, M. Coupled-Cluster Theory in Quantum Chemistry. *Rev. Mod. Phys.* **2007**, *79*, 291–352.

(41) Hättig, C.; Klopper, W.; Köhn, A.; Tew, D. P. Explicitly Correlated Electrons in Molecules. *Chem. Rev.* **2012**, *112*, 4–74.

(42) Lee, L. K.; Rance, M.; Chazin, W. J.; Palmer, A. Rotational Diffusion Anisotropy of Proteins from Simultaneous Analysis of ^{15}N and ^{13}C$^\alpha$ Nuclear Spin Relaxation. *J. Biomol. NMR* **1997**, *9*, 287–298.

(43) Ghose, R.; Fushman, D.; Cowburn, D. Determination of the Rotational Diffusion Tensor of Macromolecules in Solution from NMR Relaxation Data with a Combination of Exact and Approximate Methods – Application to the Determination of Interdomain Orientation in Multidomain Proteins. *J. Magn. Reson.* **2001**, *149*, 204–217.

(44) Hall, J. B.; Fushman, D. Variability of the ^{15}N Chemical Shielding Tensors in the B3 Domain of Protein G from ^{15}N Relaxation Measurements at Several Fields. Implications for Backbone Order Parameters. *J. Am. Chem. Soc.* **2006**, *128*, 7855–7870.

(45) Yao, L.; Grishaev, A.; Cornilescu, G.; Bax, A. Site-Specific Backbone Amide ^{15}N Chemical Shift Anisotropy Tensors in a Small Protein from Liquid Crystal and Cross-Correlated Relaxation Measurements. *J. Am. Chem. Soc.* **2010**, *132*, 4295–4309.

(46) Palmer, A. G. Chemical Exchange in Biomacromolecules: Past, Present, and Future. *J. Magn. Reson.* **2014**, *241*, 3–17.

(47) Mittermaier, A.; Kay, L. E. Measurement of Methyl ^{2}H Quadrupolar Couplings in Oriented Proteins. How Uniform Is the Quadrupolar Coupling Constant? *J. Am. Chem. Soc.* **1999**, *121*, 10608–10613.

(48) Case, D. A. Calculations of NMR Dipolar Coupling Strengths in Model Peptides. *J. Biomol. NMR* **1999**, *15*, 95–102.

(49) Ryabov, Y. E.; Geraghty, C.; Varshney, A.; Fushman, D. An Efficient Computational Method for Predicting Rotational Diffusion Tensors of Globular Proteins Using an Ellipsoid Representation. *J. Am. Chem. Soc.* **2006**, *128*, 15432–15444.

(50) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals. *Theor. Chem. Acc.* **2008**, *120*, 215–241.

(51) Dunning Jr., T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron through Neon and Hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

(52) Kendall, R. A.; Dunning Jr., T. H.; Harrison, R. J. Electron Affinities of the First-row Atoms Revisited. Systematic Basis Sets and Wave Functions. *J. Chem. Phys.* **1992**, *96*, 6796–6806.

(53) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1–2*, 19–25.

(54) Vijay-Kumar, S.; Bugg, C. E.; Cook, W. J. Structure of Ubiquitin Refined at 1.8 Å Resolution. *J. Mol. Biol.* **1987**, *194*, 531–544.

(55) Linke, M.; Köfinger, J.; Hummer, G. Rotational Diffusion Depends on Box Size in Molecular Dynamics Simulations. *J. Phys. Chem. Lett.* **2018**, *9*, 2874–2878.

(56) Takemura, K.; Kitao, A. Water Model Tuning for Improved Reproduction of Rotational Diffusion and NMR Spectral Density. *J. Phys. Chem. B* **2012**, *116*, 6279–6287.

(57) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.

(58) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.

(59) Miyamoto, S.; Kollman, P. A. Settle: An Analytical Version of the SHAKE and RAT-TLE Algorithm for Rigid Water Models. *J. Comput. Chem.* **1992**, *13*, 952–962.

(60) Hess, B. P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 116–122.

(61) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577–8593.

(62) Chen, P.-c.; Hologne, M.; Walker, O. Computing the Rotational Diffusion of Biomolecules via Molecular Dynamics Simulation and Quaternion Orientations. *J. Phys. Chem. B* **2017**, *121*, 1812–1823.

(63) Abascal, J. L. F.; Vega, C. A General Purpose Model for the Condensed Phases of Water: TIP4P/2005. *J. Chem. Phys.* **2005**, *123*, 234505.

(64) Wong, V.; Case, D. A. Evaluating Rotational Diffusion from Protein MD Simulations. *J. Phys. Chem. B* **2008**, *112*, 6013–6024.

(65) Anderson, J. S.; LeMaster, D. M. Rotational Velocity Rescaling of Molecular Dynamics Trajectories for Direct Prediction of Protein NMR Relaxation. *Biophys. Chem.* **2012**, *168–169*, 28–39.

(66) Anderson, J. S.; Hernández, G.; LeMaster, D. M. Prediction of Bond Vector Autocorrelation Functions from Larmor Frequency-Selective Order Parameter Analysis of NMR Relaxation Data. *J. Chem. Theory Comput.* **2017**, *13*, 3276–3289.

(67) Ollila, O. H. S.; Heikkinen, H. A.; Iwai, H. Rotational Dynamics of Proteins from Spin Relaxation Times and Molecular Dynamics Simulations. *J. Phys. Chem. B* **2018**, *122*, 6559–6569.

(68) Lee, A. L.; Flynn, P. F.; Wand, A. J. Comparison of ^{2}H and ^{13}C NMR Relaxation

Techniques for the Study of Protein Methyl Group Dynamics in Solution. *J. Am. Chem. Soc.* **1999**, *121*, 2891–2902.

(69) Lienin, S. F.; Bremi, T.; Brutscher, B.; Brüschweiler, R.; Ernst, R. R. Anisotropic Intramolecular Backbone Dynamics of Ubiquitin Characterized by NMR Relaxation and MD Computer Simulation. *J. Am. Chem. Soc.* **1998**, *120*, 9870–9879.

(70) Tjandra, N.; Feller, S. E.; Pastor, R. W.; Bax, A. Rotational Diffusion Anisotropy of Human Ubiquitin from ^{15}N NMR Relaxation. *J. Am. Chem. Soc.* **1995**, *117*, 12562–12566.

(71) Sheppard, D.; Li, D.-W.; Brüschweiler, R.; Tugarinov, V. Deuterium Spin Probes of Backbone Order in Proteins: ^{2}H NMR Relaxation Study of Deuterated Carbon α Sites. *J. Am. Chem. Soc.* **2009**, *131*, 15853–15865.

(72) Massi, F.; Grey, M. J.; Palmer, A. G. Microsecond timescale backbone conformational dynamics in ubiquitin studied with NMR R1ρ relaxation experiments. *Prot. Sci.* **2005**, *14*, 735–742.

(73) Liao, X.; Long, D.; Li, D.-W.; Brüschweiler, R.; Tugarinov, V. Probing Side-Chain Dynamics in Proteins by the Measurement of Nine Deuterium Relaxation Rates per Methyl Group. *J. Phys. Chem. B* **2012**, *116*, 606–620.

(74) Long, D.; Li, D.-W.; Walter, K. F. A.; Griesinger, C.; Brüschweiler, R. Toward a Predictive Understanding of Slow Methyl Group Dynamics in Proteins. *Biophys. J.* **2011**, *101*, 910–915.

(75) Bowman, G. R. Accurately Modeling Nanosecond Protein Dynamics Requires at Least Microseconds of Simulation. *J. Comput. Chem.* **2016**, *37*, 558–566.

(76) Kasinath, V.; Sharp, K. A.; Wand, A. J. Microscopic Insights into the NMR Relaxation-

Based Protein Conformational Entropy Meter. *J. Am. Chem. Soc.* **2013**, *135*, 15092–15100.

Supporting information for:

Predicting NMR Relaxation of Proteins from

Molecular Dynamics Simulations with

Accurate Methyl Rotation Barriers

Falk Hoffmann,[†] Frans A. A. Mulder,[‡] and Lars V. Schäfer*,[†]

†Theoretical Chemistry, Ruhr University Bochum, D-44780 Bochum, Germany

‡Interdisciplinary Nanoscience Center (iNANO) and Department of Chemistry, Aarhus
University, DK-8000 Aarhus, Denmark

E-mail: lars.schaefer@ruhr-uni-bochum.de
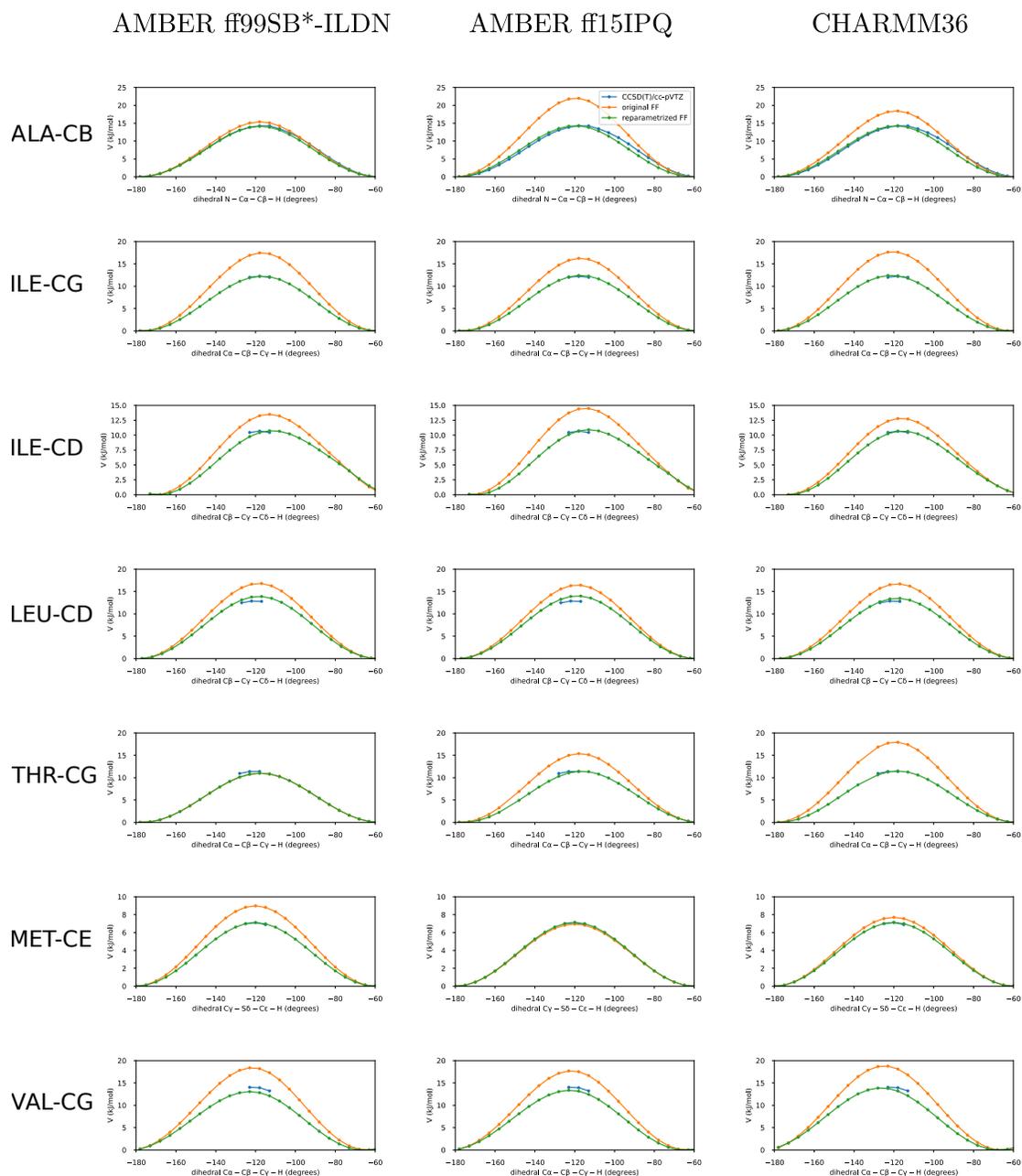
Figure S1: Potential energy curves of methyl group rotation in blocked dipeptides. Single-point energies of selected points (close to the minima and maxima) were calculated with CCSD(T)/cc-pVTZ (blue). Energies from force field calculations with the original and reparametrized force field parameter sets of AMBER ff99SB*-ILDN (left), AMBER ff15ipq (center) and CHARMM36 (right) are plotted in orange and green, respectively.
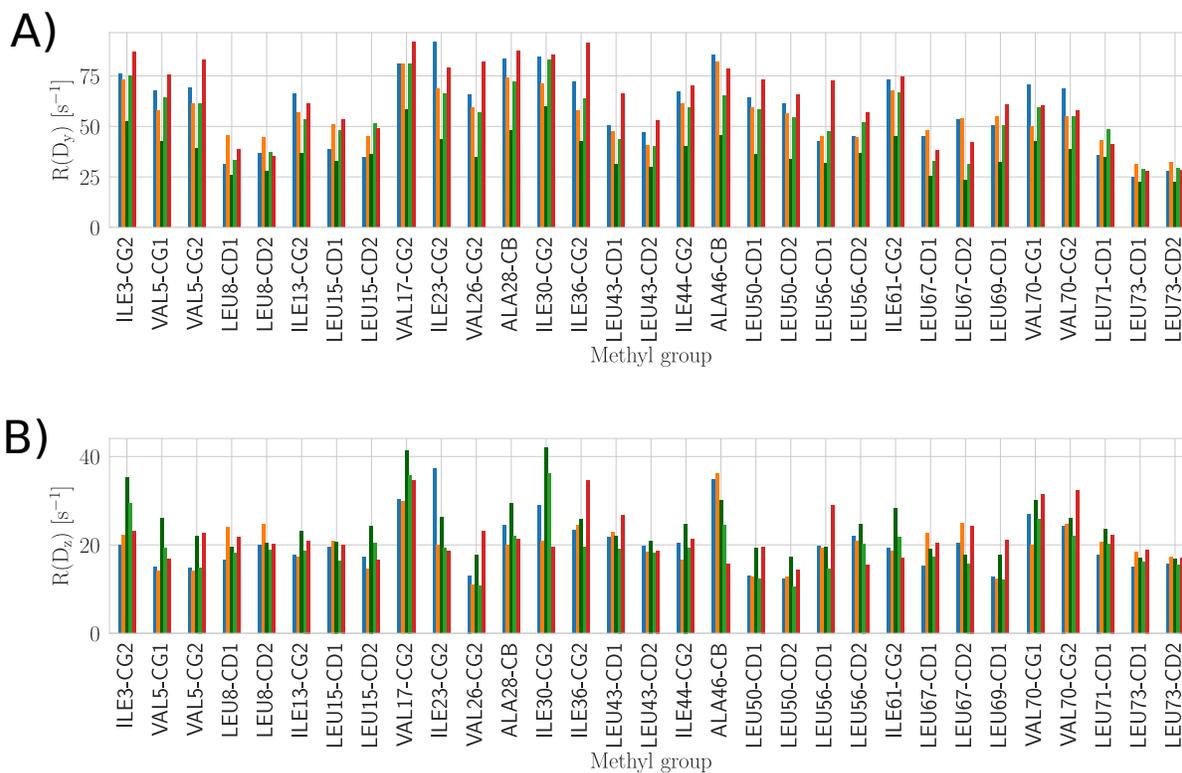
Figure S2: Methyl relaxation rates $R(D_y)$ (A) and $R(D_z)$ (B) from MD simulations with the AMBER ff99SB*-ILDN force field (orange), AMBER ff15ipq force field (blue), and the CHARMM36 force field (dark and light green for unscaled and scaled overal rotational diffusion time, respectively) are compared to the experimental values from Liao and coworkers.