

Combining Molecular and Spin Dynamics Simulations with Solid-State NMR: A Case Study of Amphiphilic Lysine-Leucine Repeat Peptide Aggregates

Prashant S. Emani^{a,‡,†}; Yeneneh Y. Yimer^{b,‡,††}; Stephen K. Davidowski^{a,†††}; Rachel N. Gebhart^a; Helen E. Ferreira^a; Ilya Kuprov^c; Jim Pfaendtner^{b,*}; Gary P. Drobny^{a,*}

^a Department of Chemistry, University of Washington, Box 351700, Seattle, WA 98195-1700, U.S.A.

^b Department of Chemical Engineering, University of Washington, 105 Benson Hall, Box 351750, Seattle, WA 98195-1750, U.S.A.

^c Department of Chemistry, University of Southampton, Highfield, Southampton SO17 1BJ, U.K.

KEYWORDS Molecular dynamics, Solid-state NMR, Uniformly-labeled residues, Spin networks, LK α 14.

ABSTRACT: Interpreting dynamics in solid-state molecular systems requires characterization of the potentially heterogeneous environmental contexts of molecules. In particular, the analysis of solid-state NMR (ssNMR) data to elucidate molecular dynamics involves modeling the restriction to overall tumbling by neighbors, as well as the concentrations of water and buffer. In this exploration of the factors that influence motion, we utilize atomistic molecular dynamics (MD) trajectories of peptide aggregates with varying hydration to mimic an amorphous solid-state environment, and predict ssNMR relaxation rates. We also account for spin diffusion in multiply spin-labeled (up to 19 nuclei) residues, with several models of dipolar-coupling networks. The framework serves as a general approach to determine essential spin couplings affecting relaxation, benchmark MD force fields, and reveal the hydration-dependence of dynamics in a crowded environment. We demonstrate the methodology on a previously characterized amphiphilic 14-residue lysine-leucine repeat peptide, LK α 14 (Ac-LKKLLKLLKLLKL-c), which has an α -helical secondary structure and putatively forms leucine-burying tetramers in the solid state. We measure R_1 relaxation rates of uniformly ^{13}C -labeled, and site-specific ^2H -labeled leucines in the hydrophobic core of LK α 14 at multiple hydration levels. Studies of 9 and 18 tetramer bundles reveal that: (a) for the incoherent component of ^{13}C relaxation, nearest-neighbor spin interactions dominate, while ^1H - ^1H interactions have minimal impact; (b) AMBER ff14SB dihedral barriers for the leucine bond (“methyl rotation barriers”) must be lowered by a factor of 0.7 to better match the ^2H data; (c) proton-driven spin diffusion (PDSD) explains some of the discrepancy between experimental and simulated rates for the C_β and C_α nuclei; and (d) ^{13}C relaxation rates are mostly underestimated in the MD simulations at all hydrations, and the discrepancies identify likely motions missing in the 50 ns MD trajectories.

1. INTRODUCTION

Nuclear magnetic resonance (NMR) spectroscopic data have long been used to benchmark and constrain molecular dynamics (MD) trajectories, and to thus extract atomic-detail structural and dynamic models^{1–13} consistent with the NMR observables. Such an approach necessarily involves significant challenges. One source of complexity relates to the diversity of observables, used either as constraints or as fitting targets, such as chemical shifts^{6,8}, RDCs⁷, NOEs¹⁰, and relaxation times and order parameters^{1,2,4,5,9,11,12,14–16}, which collectively cover a wide range of motional time scales, and inter- and intra-molecular length scales. Accounting for these experimental outputs requires a detailed understanding of the pulse sequences, spin interactions, ranges of time- and length-scale applicability, and assumptions demarcating the relevant molecular degrees of freedom. A second challenge, on the MD side, relates to the need for a careful assessment and reproduction of the sample conditions, including but not restricted to

temperature⁹, concentration of the solute(s)^{17,18}, and concentration of buffer^{9,18}. In the current investigation, the primary aim is to model dynamics in a solid-state sample by using MD trajectories to simulate relaxation rates for uniformly ^{13}C - and ^{15}N -labeled leucines, and for methyl- ^2H leucines, measured at different hydration levels. This involves addressing both the NMR-associated challenge of accounting for the combined contributions of spin diffusion in uniformly labeled residues and of motions to the relaxation rates, as well as the MD challenge of modelling packed solid sample conditions with a potentially variable amount of associated water.

The impact of spin diffusion is of experimental relevance in situations where site-specific labeling is either difficult or expensive, and where the relaxation contributions of neighboring spins cannot be suppressed using pulse techniques, due, for example, to an overlap of the resonances of two or more spins. In fact, spin diffusion among natural abundance or labeled samples has even been exploited in NMR experiments to ex-

quires a full characterization of the spin-coupling networks and their impacts. While general theoretical approaches have been available for many years^{25,26}, the advent of sophisticated computational methods^{27,28} have now made quantifying spin diffusion in specific molecular systems more tractable. In this work, we use the framework for spin dynamics calculations available in the SPINACH software package²⁷, while modifying the relaxation algorithm to allow for intra-molecular and restricted global dynamics. We construct three models of dipolar coupling between the relevant spins, and assess the impacts of the inclusion or exclusion of different couplings.

The simulation of molecular dynamics in the case of sample conditions other than dilute solutions has been addressed for aggregate (for example, references²⁹⁻³⁴) and multi-component systems^{12,35,36}. In order to approach solid-state conditions, defined here as the range of conditions under which the solute particles experience restricted overall rotation, the concentration of the solute needs to be significantly increased to properly account for packing interactions. Moreover, depending on the sample and the method of preparation, there may be a need to include buffer, water or an interacting partner in the simulations. Our approach is to pack multiple monomers of the molecule together with a variable number of water molecules into a given volume and assess the relaxation rates as a function of the molecule-to-water ratio. In particular, it is to be emphasized that the MD simulations discussed herein are directed towards amorphous solid-state conditions.

We apply our methodology to the analysis of the structure and dynamics of an amphiphilic lysine-leucine repeat peptide, LK α 14 (Ac-LKKLLKLLKLLKL-c) (Figure 1), so designated due to its α -helical secondary structure in concentrated solutions and at air-water interfaces^{34,37}, and when bound to a variety of surfaces^{38,39}. We have chosen this as our model system based on the availability of substantial prior experimental study and model-building: several studies^{32,37,38,40-42} have indicated that LK α 14 forms an α -helix concurrently with pre-aggregation in solution remaining an aggregate upon surface adsorption), with the most likely aggregation state being a tetramer (Figures 1A and 1B); recent work²² has also shown that this peptide has the ability to precipitate silica nanospheres out of silicic acid solutions, and retains its α -helical secondary structure in both the neat (non-buffered) solid-state and when co-precipitated with silica. These results were interpreted as evidence of tetramerization in the neat and silica co-precipitate conditions, with the tetramer potentially serving as a primary scaffold for the patterning of silica. Given these prior constraints on the aggregate-length-scale ($\geq nm$) structure, we are able to bias MD simulations in a principled manner and assess the degree to which NMR-based dynamics can be captured by our integrative framework. The trajectories are set up with a 3×3 grid of tetramer bundles arranged in one (9 bundles) or two (18 bundles) layers, with four levels of hydration (Figure 1C): approximately 10, 15, 20, or 30 waters (abbreviated as “w”) per amino acid of the peptide. In addition to using LK α 14 as a well-characterized system for methodological assessment, this study also aims to lay the methodological groundwork for a future MD-NMR analysis of silica co-precipitation by LK α 14 by revealing the self-aggregation mechanisms in the neat state.

2. METHODS

Detailed derivations of all theory and computational subsections are provided in SI text.

2.1 Experimental methods

2.1.1 Sample preparation: Protected amino acids were purchased from P3 Biosystems (Louisville, KY), Fmoc-Leu-Wang resin from EMD Millipore (Darmstadt, Germany), and protected Leu(d_3)-OH was purchased from Cambridge Isotope Laboratories (Tewksbury, MA). Acetic Anhydride was purchased from Fisher Scientific (Waltham, MA). Trifluoroacetic acid (TFA) and triisopropyl silane (TIS) were purchased from Sigma Aldrich (St. Louis, MS). All materials were used as received, without further purification.

LK peptides (Ac-LKKLLKLLKLLKL-c) were synthesized de novo on a CEM Liberty Blue Peptide Synthesizer using standard Fmoc chemistry. Site specific 2H enrichment was achieved at the L5, L8 and L11 positions of 3 separate LK peptides using Fmoc-Leu-OH-5,5,5- d_3 . Site specific ^{13}C enrichment was achieved at the L7 and L11 positions of 2 separate LK peptides using uniformly ^{13}C and ^{15}N labeled Fmoc-Leu-OH. After synthesis, the n-terminus was deprotected and acylated using neat acetic anhydride. The peptide was subsequently cleaved from the resin by addition to a mixture of 95% TFA, 2.5% TIS, and 2.5% water at a ratio of 10 mL of cocktail for each gram of resin. The reaction was mechanically agitated for 2 hours. Dropwise addition of the liquid phase to 40 mL of cold t-butyl methyl ether resulted in the precipitation of the peptide. The peptide was separated from the mixture via centrifugation (3,000 RPM for 20 min). Cold cleaved peptide was then dissolved in water and purified through a 0.45 μm Nylon syringe filter before further lyophilization.

Hydration of the samples was accomplished by pipetting a set amount of deuterium-depleted water into the peptide sample within a 600 μL microcentrifuge tube: for $w = 15$, 43.5 μL of water was added for every 20 mg of peptide, while for $w = 20$ and $w = 30$, 58 and 87 μL of water were added per 20 mg of peptide, respectively. The samples were mixed, and then pipetted into glass capillaries which were inserted into glass NMR tubes prior to being run on the spectrometer.

2.1.2 NMR Experiments: 2H Solid-state NMR experiments were performed on a ‘home-built’ 400 MHz wide bore spectrometer equipped with a 4 mm Varian HXY MAS probe. The spectrometer was configured to operate at a 2H resonant frequency of 61.4 MHz. Static deuterium T_1 measurements were made using an inversion-recovery pulse sequence which included a quadrupolar echo for detection⁴³. The data were collected using a recycle delay of 0.5 seconds, 65536 scans. Pulse widths used for $\pi/2$ and π pulses were 4.5 and 9.0 μs respectively. Ten delay increments between 0.0005 and 0.5 s were used.

^{13}C solid-state NMR experiments were performed on a Bruker 500 MHz spectrometer equipped with a 3.2 mm HXY MAS probe operating at a ^{13}C resonant frequency of 125.29 MHz. T_1 values were obtained by utilizing the cross polarization (CP) based pulse sequence developed by Torchia⁴⁴. Cross polarization was optimized by setting the ^{13}C power of the contact pulse to approximately 50 kHz and tuning the 1H power during the contact pulse to the +1 SSB of the Hartman Hahn match condition (60kHz). The data were collected using a recycle delay of 1 second, 1024 scans, MAS speed of 10 kHz, a 1H $\pi/2$ pulse of 3.0 μs , and two pulse phase modulated (TPPM) 1H decoupling of 80 kHz during acquisition. A total of 14 τ delays were chosen between 0.050 and 6 seconds for the T_1 measurements.

2.1.3 Data Processing: Post-processing of the 2H FID data were carried out in ‘home-built’ software and exported for further processing in MATLAB. All spectra were baseline corrected before the powder pattern was integrated and plotted against the delay time. The resulting curve was then fit to an exponential. Post processing and baseline correcting of the ^{13}C data was carried out in Topspin 3.2. The T_1 was then calculate by fitting the spectra using home-built R code. The ^{13}C rates for the C_γ and $C_{\delta_{1,2}}$ were obtained by deconvoluting three overlapping peaks in the spectra. As such, the peak of smallest intensity was taken to be associated with the C_γ , with half the intensity of the two other peaks combined. The C_{δ_1} is taken consistently to be the peak with the higher magnitude of the chemical shift out of the two remaining peaks.

2.2 Theory of ^{13}C Relaxation Times

The relaxation superoperator is calculated using SPINACH based on expressions in Kuprov⁴⁵, modified to include intramolecular motions in the powder-averaged solid-state sample undergoing MAS rotation. Bloch-Redfield-Wangsness (BRW) theory is used to express the superoperator as the sum over integrals of two-time correlation functions $C_{mm}(L, S, L', S', \tau)$ of the orientation-dependent spin interactions and functions of the 2nd rank tensor spin operators $\hat{T}_k^{(2)}(L, S)$:

$$\hat{R} = - \sum_{km} \int_0^\infty d\tau \frac{1}{5} \left(\sum_u [d_{ku}^{(2)}(\beta_{MAS})]^2 \right) \cdot \sum_{\substack{\text{spins } L, S \\ L', S'}} C_{mm}(L, S, L', S', \tau) \left[\hat{T}_k^{(2)}(L, S) e^{-i\hat{H}_0\tau} \hat{T}_k^{(2)\dagger}(L', S') e^{i\hat{H}_0\tau} \right] \quad (1)$$

Here, $d_{ku}^{(2)}(\beta_{MAS})$ is the 2nd rank Wigner small-d matrix (defined in SI text according to conventions discussed in refs. 46,47); $\hat{T}_k^{(2)}(L, S)$ is the 2nd rank tensor spin operator of projection k for the interaction between the spins L and S ; \hat{H}_0 is the Zeeman Hamiltonian; and the two-time correlation function is given by

$$C_{mq}(L, S, L', S', \tau) = \langle (\Phi_m(L, S, t) - \langle \Phi_m(L, S, t) \rangle) (\Phi_q^*(L', S', t + \tau) - \langle \Phi_q^*(L', S', t + \tau) \rangle) \rangle \quad (2)$$

where $\langle \dots \rangle$ represents ensemble-averaging and $\Phi_m(L, S, t)$ encapsulates information about the orientation dependence of the spin interaction between spins L and S at time t . Note that the correlation functions are calculated as correlations of the deviations from the ensemble average. This is necessary in the case of solid-state NMR to ensure that the correlation functions decay to 0 at longer times. Expressions for $\Phi_m(L, S, t)$ are provided for the dipolar and CSA interactions:

$$\Phi_m^{Dip}(L, S, t) = -\sqrt{6} \frac{\mu_0 \gamma_I \gamma_S \hbar}{4\pi r^3} d_{m,0}^{(2)}(\Omega_{PAS \rightarrow C}(L, S, t)) \quad (3a)$$

$$\Phi_m^{CSA}(S, t) = \gamma_S B_0 \frac{3\sigma_{ZZ} - Tr(\vec{\sigma})}{\sqrt{6}} \mathcal{D}_{m,0}^{(2)}(\Omega_{PAS \rightarrow C}(S, t)) + \gamma_S B_0 \frac{\sigma_{XX} - \sigma_{YY}}{2} \left(\mathcal{D}_{m,2}^{(2)}(\Omega_{PAS \rightarrow C}(S, t)) + \mathcal{D}_{m,-2}^{(2)}(\Omega_{PAS \rightarrow C}(S, t)) \right) \quad (3b)$$

where the $\Omega_{PAS \rightarrow C} = \{\alpha_{PAS \rightarrow C}, \beta_{PAS \rightarrow C}, \gamma_{PAS \rightarrow C}\}$ is the set of Euler angles transforming the principal axis system (PAS) of a given spin interaction to the crystal director (C) frame; $\mathcal{D}_{m,n}^{(2)}(\Omega_{PAS \rightarrow C}) = e^{-im\alpha_{PAS \rightarrow C}} d_{mn}(\beta_{PAS \rightarrow C}) e^{-in\gamma_{PAS \rightarrow C}}$ is the 2nd order Wigner-D rotation matrix for a passive transformation from C to PAS; $\vec{\sigma}$ is the chemical shift tensor, and all the other terms have their usual meanings. Note that the quantification of the correlation function in terms of $\Omega_{PAS \rightarrow C}$ implies that the time variation of $C_{mq}(L, S, L', S', \tau)$ includes all intra-molecular motions as well as all motions of the entire molecule relative to the crystal frame (i.e. rotations relative to the tetramer bundle and rotations of the bundle as a whole relative to the MD simulation 'crystal' frame).

Finally, to account for the decay of the relaxation superoperator to the equilibrium state, \hat{R} in Eq. 1 is multiplied by an exponential factor to yield $\hat{R} \rightarrow \hat{R} \cdot \exp\left(\frac{\hat{H}_0}{k_B T}\right)$, in accordance with the DiBari-Levitt⁴⁸ method of equilibration.

2.3 Expressing the Relaxation Superoperator in an Observable Basis

To properly account for the potentially non-monoexponential behavior of the longitudinal relaxation of nuclear spins, the cross-relaxation of all contributing spin observables must be quantified. This implies constructing a basis set of spin observables that are involved in the evolution of the longitudinal spins of all the nuclei of interest, and subsequently expressing \hat{R} in this basis. For the basis, we will choose three 1st order tensor terms $\hat{T}_{0,\pm 1}^{(1)}(I) = \{I_-, I_Z, I_+\}$ for each spin in the system (total number of spins = 18 in the case of multiply labeled leucine), five 2nd order terms $\hat{T}_{0,\pm 1,\pm 2}^{(2)}(I, S) = \left\{ \frac{1}{2} I_- S_-, \frac{1}{2} (I_Z S_- + I_- S_Z), \frac{1}{\sqrt{6}} (3I_Z S_Z - \vec{I} \cdot \vec{S}), -\frac{1}{2} (I_Z S_+ + I_+ S_Z), \frac{1}{2} I_+ S_+ \right\}$ for each pair of spins in the system (total number of pairs = $^{18}C_2 = 153$ for leucine)

and only one 3rd order term $\hat{T}_0^{(3)}(I, S^{(i)}, S^{(j)}) = \{I_Z S_Z^{(i)} S_Z^{(j)}\}$ for every triplet of spins (total number of triplets = $^{18}C_3 = 816$ for leucine). We restrict the number of third order terms to reduce the dimensionality of the problem, while at the same time keeping terms such as $I_Z S_Z^{(i)} S_Z^{(j)}$ that have been shown to have an influence on the longitudinal relaxation.²⁶ Then the superoperator expressed in this observable basis is given by

$$\hat{R}_{obs} = \begin{pmatrix} \langle \hat{I}_Z \rangle^\dagger \cdot \hat{R} \cdot \langle \hat{I}_Z \rangle & \langle \hat{I}_Z \rangle^\dagger \cdot \hat{R} \cdot \langle S_Z^{(i)} \rangle & \dots \\ \langle S_Z^{(i)} \rangle^\dagger \cdot \hat{R} \cdot \langle \hat{I}_Z \rangle & \ddots & \vdots \\ \vdots & \dots & \langle I_Z S_Z^{(i)} S_Z^{(j)} \rangle^\dagger \cdot \hat{R} \cdot \langle I_Z S_Z^{(i)} S_Z^{(j)} \rangle \end{pmatrix} \quad (4)$$

Then, if the superoperator is diagonalized as $\hat{R}_{obs} = \vec{V} \cdot \vec{\Lambda} \cdot \vec{V}^{-1}$, the result of a non-selective inversion recovery experiment to observe the time-dependence of a given ^{13}C nucleus $\langle \hat{I}_Z(Obs) \rangle(t)$ is

$$\frac{\langle \hat{I}_Z(Obs) \rangle(t)}{\langle \hat{I}_Z(Obs) \rangle(0)} = \sum_{j=1}^{N_{^{13}C}} \left(\vec{V} \cdot e^{-\vec{\Lambda}t} \cdot \vec{V}^{-1} \right)_{\langle \hat{I}_Z(Obs) \rangle, \langle I_{Z,j} \rangle} \quad (5)$$

where the summation is over all the ^{13}C nuclei in the system that are inverted by the non-selective 180° pulse, and the subscripted $\{\langle \hat{I}_Z(Obs) \rangle, \langle I_{Z,j} \rangle\}$ mark the observable-associated indices of the matrix $\vec{V} \cdot e^{-\vec{\Lambda}t} \cdot \vec{V}^{-1}$. The resulting spin relaxation time is obtained as an exponential fit to Equation 5, i.e. by fitting an approximate exponential function to the multi-exponential decay, as would be done for the experimental value.

To account for the effects of coherent proton-driven spin diffusion (PDS) effects on the ^{13}C relaxation rates, we made additional changes (see SI text for further details), the most important of which was the inclusion of all unique longitudinal- ^{13}C -spin-difference observable pairs $\langle (\hat{I}_Z^{(i)} - \hat{I}_Z^{(j)}) \rangle$, for all $\{i, j < i\}$ in the observable basis (see for example, 49–51). Importantly, these observables were projected onto the relaxation superoperator without DiBari-Levitt equilibration to capture the coherent component of spin diffusion. These projections on to the coherent component of the superoperator were then readily incorporated into \hat{R}_{obs} , and the relaxation rates were recalculated.

2.4 Theory of 2H Relaxation Times

The static deuterium relaxation T_{1Z} times are calculated using expressions derived by Torchia and Szabo.⁵² The T_{1Z} relaxation times for a dominant quadrupolar interaction in 2H are given by

$$\frac{1}{T_{1Z}} = \frac{\omega_Q^2}{3} (J_1(\omega_D) + 4J_2(2\omega_D)) \quad (6)$$

where $\omega_Q = \frac{3e^2qQ}{4h} = 2\pi \left(\frac{3e^2qQ}{4h} \right) = 2\pi \left(\frac{3}{4} \nu_{QCC} \right)$ (ν_{QCC} is the quadrupolar coupling constant (QCC)) and $J_m(\omega)$ is the spectral density of order m with the definition

$$J_m(\omega) = 2 \int_0^\infty dt C_m(t) \cos(\omega t) \quad (7)$$

In our samples, the leucine residues are triply deuterium-labeled at one of the two δ -methyl positions. The orientations of the electric field gradient (EFG) tensor, which defines the spatial dependence of the interaction and thus the correlation functions, are calculated in the lab frame by first rotating the principal axis system (PAS) of the EFG tensor into the crystal frame (C) and then rotating the C frame into the lab frame. The second frame transformation is simplified in a powder averaged sample by integrating over all crystallite orientations. The PAS is defined as being axially symmetric about the C- 2H bond for the methyl groups with a QCC = 167 kHz^{53,54} and an asymmetry parameter $\eta = 0$ ⁵⁵⁻⁵⁷. Therefore, to calculate the correlation functions, the orientations of the methyl C- 2H bond orientations need to be extracted from each MD frame. The powder-averaged correlation function is then calculated as

$$\bar{C}_m(t) = \frac{1}{5} \sum_{a=-2}^2 \left(\left\langle D_{0,a}^{(2)*} \left(\vec{\Omega}_{PAS \rightarrow C}(0) \right) - \langle D_{0,a}^{(2)*} \left(\vec{\Omega}_{PAS \rightarrow C}(0) \right) \right\rangle \right), \left(\left\langle D_{0,a}^{(2)} \left(\vec{\Omega}_{PAS \rightarrow C}(t) \right) - \langle D_{0,a}^{(2)} \left(\vec{\Omega}_{PAS \rightarrow C}(t) \right) \right\rangle \right) \quad (8)$$

where $\Omega_{PAS \rightarrow C} = \{\alpha_{PAS \rightarrow C}, \beta_{PAS \rightarrow C}, \gamma_{PAS \rightarrow C}\}$ and the Wigner rotation matrices were defined previously.

2.5 Molecular dynamics trajectories

The starting point of the MD simulations was taken to be a set of 9 or 18 tetramers of α -helical LK α 14 arrayed in a grid with 3×3 or $2 \times (3 \times 3)$ (i.e. two layers of 3×3 grids) bundles, respectively, with the axes of symmetry of all the bundles aligned. The MD runs were carried out in GROMACS version 4.6.5^{58,59}, and the force-field parameters used for each component were as follows: for H_2O , a TIP3P rigid water model⁶⁰; for LK α 14, the AMBER ff14SB force field⁶¹; and for Cl^- (added for charge neutralization), GAFF⁶². The Lennard-Jones potential was cut off at 10 Å and shifted to zero. A cutoff distance of 11 Å was used for the electrostatic interactions and, beyond the cutoff, particle-mesh Ewald (PME) summations were used to calculate the long-range electrostatic interactions⁶³. The LINCS algorithm⁶⁴ was applied to hydrogen-heavy atom bonds. A sequence of equilibration steps (described in SI section S2.1) was followed by the production run carried out under NPT for 50 ns at 300 K and 1 atm (velocity-rescaling thermostat and Parrinello-Rahman barostat). Periodic boundary conditions were applied in all directions. As the final production runs were carried out under NPT conditions, the volume of the box varied during the course of each run. For estimation of error bars, three MD replicas were run for each set of physical conditions and number of bundles.

The LK α 14 monomers were arranged in an anti-parallel configuration within each tetramer bundle, i.e. looking from the top down along the axis of symmetry and going clockwise, each monomer was placed alternately in a C-terminus to N-terminus or N-terminus to C-terminus configuration. For both the 9 and 18 bundles case, we carried out simulations with four different numbers of water molecules and the dihedral potential about the $C_\gamma - C_\delta$ parameters set at two different values: the standard AMBER ff14SB force field value, and at a value scaled by a factor of 0.7. The hydration levels were set at $\sim 10, 15, 20$ and 30 water molecules per residue (“w”) implying a total of $\sim 5000, 7500, 10000$, and 15000 waters and $\sim 10000, 15000, 20000$, and 31000 waters in the simulation box for the 9- and 18-bundle cases, respectively, according to

$$\frac{\text{No. of waters per residue} \times \text{Total no. of water molecules in volume}}{\text{No. of bundles} \times 14 (= \text{No. of residues in LK}\alpha 14)} \quad (9)$$

The parameter sets are summarized in Table S1.

The fact that we simulated three replicates for each set of conditions allowed us to take into account different initial conditions in the MD trajectories. Moreover, given that each of the simulations is run on 9 or 18 bundles, we have 36 or 72 copies of the residue considered, respectively. This allows for a reasonable sampling of the diversity in supramolecular arrangements of the tetramer, as well as of the individual residues. This is very relevant given our choice to orient the tetramers in a particular, antiparallel configuration. However, we noticed a significant diversity in the relative orientation of the monomers within the tetramer bundles and in the distance of individual monomers from the tetramer centroid during the course of the trajectories. This gave us confidence in the simulated sampling of the supramolecular configurations.

2.6 Correlation function calculations

Given our inclusion of only two types of spin interactions, dipolar coupling and CSA interactions, we calculate three types of correlation functions: dipole-dipole, $C_{mq}^{Dip-Dip}(\{I, S\}, \{I', S'\}, \tau)$; CSA-CSA, $C_{mq}^{CSA-CSA}(\{S\}, \{S'\}, \tau)$; and a pair of cross-correlations, $C_{mq}^{CSA-Dip}(\{S\}, \{I', S'\}, \tau)$ and $C_{mq}^{Dip-CSA}(\{I, S\}, \{S'\}, \tau)$. In our simulations, the experimental labeling scheme was reflected in our consideration of only nuclei within single residues for each separate calculation of couplings, though there is no methodological restriction on the number of nuclei and/or the number of residues. For every residue of interest, the internuclear vectors for the dipolar coupling and the bonds needed to calculate the CSA tensor were extracted from the trajectories. For a given trajectory, ensemble averages of all correlation functions $C_{mq} = \langle (\langle \Phi_m(I, S, t) - \langle \Phi_m(I, S, t) \rangle) (\langle \Phi_q^*(I', S', t + \tau) - \langle \Phi_q^*(I', S', t + \tau) \rangle) \rangle$ are calculated as the average, for a given set

of spin pairs $\{(I, S), (I', S')\}$ and indices $\{m, q\}$, over every pair of time points that are separated by the same time difference τ . The 50 ns trajectories were sampled at 2500 time points, leading to an integration time step of 20 ps (see SI text, equation 25). To account for limited statistics at values of τ approaching the length of the trajectory, we imposed an upper integration limit of 80% of the trajectory length for the integral in Equation 1. For a 50 ns simulations, this amounts to an upper limit of 40 ns on the maximum time difference τ considered. Additionally, the integration for each pair of spins and each pair of indices was stopped when the function first crossed the x-axis. This was necessary as limited statistics often caused large fluctuations in the correlation functions after they were expected to have converged. This does necessarily contribute a certain amount to the error, which is accounted for by analyses involving an extension of the trajectory for better statistics (see SI text under “Error Analysis”). However, it is important to state that motions on time scales $\geq \sim 10$ ns may not be sufficiently sampled in these simulations and thus may contribute to discrepancies between simulation and experiment. The correlation function calculations were carried out using in-house C++ code.

The correlation functions reported in this paper are averaged over all four residues within each bundle. This is done as a compromise between simultaneously accounting for the inherent diversity in the behavior of the residues in a sample, on the one hand, while still capturing the necessary ensemble-average.

2.7 CSA tensor calculations

The CSA tensors were determined using published empirical tensors and geometries for the three nuclei deemed to have non-negligible CSA tensor values: the amide ^{15}N , the carbonyl ^{13}C and the alpha ^{13}C . The bond orientations used in the calculation of these tensor orientations with respect to each MD frame are extracted from the trajectories. The tensor orientations were calculated using in-house Mathematica code. Details of the geometries are provided in the SI text.

2.8 Relaxation rate calculation in SPINACH

The “redfield” relaxation module in Matlab-based SPINACH²⁷ package (version 1.5) was modified to allow for the integration of \hat{R} with the inclusion of discretized, time-varying correlation functions representing all motions (internal or entire-molecule) relative to the crystal frame. The standard SPINACH parameter set used for the calculations is provided in the SI text. The connectivity maps defining the three types of spin networks described in the results are set up in SPINACH by setting the “bas.manual” option equal to a logical matrix with 1’s at the indices corresponding to coupled spins, and 0’s otherwise. The modifications are briefly summarized here (details in SI text):

- Define the spin operators $\hat{I}_k^{(2)}(I, S)$ needed for the calculation in Equation 1 using two new modules T_Operator.m (for the dipolar couplings) and CS_Operator.m (for the CSA interactions).
- Define three separate relaxation modules for the dipole-dipole, CSA-CSA and cross-correlation components and combine the superoperators from each at the end.
- Modify the number of spin indices needed in the integration. The reduction in the number of the indices occurs due to the powder-averaging and MAS rotation (assumed to be much faster than the relaxation rates, thereby leading to simplified expressions).
- Extract the correlation function elements from text files. The order of the elements follows a particular convention (see SI text), used to ensure proper multiplication of the spatial and spin components in Equation 1.
- Carry out discrete Boole integration to calculate \hat{R} .
- A new module ObservableBasis.m is defined to project \hat{R} on to the observable basis of choice, calculate the associated eigensystem of \hat{R}_{obs} , and fit the simulated time dependence to a single exponential for matching to the experimental T_1 IR values.
- For computational efficiency, we used the basis set approximation with all product states up to and including 3 spins. We made

some tests of this approximation as described in the SI text (section S3.3).

2.9 Error analysis

Details are provided in the SI text (section S4).

2.10 Density functional theory (DFT) calculations

We calculated the dihedral barriers for the $C_\gamma - C_\delta$ bond. The leucine molecule is fully optimized at B3LYP/6-31+G(d,p) level using the Gaussian 09 package⁶⁵. The energy calculations were performed by varying the torsion angle from ϕ to $\phi + 360^\circ$ in steps of 10° , where ϕ is the torsion angle for the initial optimized structure. At each torsion angle all atoms in the residue, except for the four atoms which form the torsion angle, are allowed to relax. The resulting potential profile is shown in the SI figure S5.

2.11 Calculation of the methyl rotation barriers using the potential of mean force

We calculated the barrier for the potential of mean force (ΔV_{pmf}), defined following the procedure of Chatfield and Wong⁶⁶. The potential of mean force is calculated by binning the dihedral angles $C_\beta - C_\gamma - C_{\delta 1} - H_{\delta 11}$ and $C_\beta - C_\gamma - C_{\delta 2} - H_{\delta 21}$ extracted from the MD frames, and by fitting the Boltzmann-inverted probability distribution.

Table 1. Experimental ^{13}C R_1 relaxation rates for all carbon nuclei in L7 and L11²².

Nucleus	R_1 for L7 (in Hz): Lyophilized	R_1 for L7 (in Hz): w = 15	R_1 for L7 (in Hz): w = 20	R_1 for L11 (in Hz): Lyophilized	R_1 for L11 (in Hz): w = 15	R_1 for L11 (in Hz): w = 20
CO	0.78 ± 0.01	1.04 ± 0.06	0.92 ± 0.06	0.87 ± 0.01	1.12 ± 0.05	0.90 ± 0.05
C_α	0.82 ± 0.01	1.47 ± 0.09	1.56 ± 0.08	0.91 ± 0.02	1.43 ± 0.07	1.59 ± 0.10
C_β	0.88 ± 0.01	2.93 ± 0.34	3.75 ± 0.45	1.01 ± 0.03	2.48 ± 0.19	4.46 ± 0.46
C_γ	0.92 ± 0.02	2.63 ± 0.17	2.56 ± 0.16	1.17 ± 0.15	2.65 ± 0.17	2.60 ± 0.13
$C_{\delta 1}$	0.89 ± 0.02^a	2.77 ± 0.18^a	2.87 ± 0.19^a	1.03 ± 0.03^a	2.85 ± 0.18^a	2.94 ± 0.20
$C_{\delta 2}$	0.95 ± 0.01^a	2.77 ± 0.18^a	2.87 ± 0.19^a	0.90 ± 0.13^a	2.82 ± 0.21^a	2.53 ± 0.15

^a The association of the experimental $C_{\delta 1,2}$ rates with those of the MD simulations is arbitrary, as there is no way of unambiguously assigning the labels “1, 2” in the experiments. The order was chosen based on matching the trends in the MD simulation with those in the experimental data.

3. RESULTS AND DISCUSSION

The experimental data obtained for the ^{13}C - and ^2H -labeled samples are shown in Tables 1 and 2.

Experiments on the ^{13}C -labeled samples were carried out at three different hydration levels chosen to enable comparison to the MD simulations: lyophilized, 15w and 20w. Experiments on the ^2H -labeled samples were done at four different hydration levels: lyophilized, 15w, 20w and 30w. As described in the Methods, the water content was set by pipetting in the appropriate amount of water into lyophilized sample. For the sake of comparison, recasting these hydrations in terms of the ratio of the weight of water in grams to the weight of protein in grams (the so-called ‘h’ ratio⁶⁷) yields $h \approx 0, 2.2, 2.9, 4.4$ based on the pipetted weights discussed in the Methods section. The MD simulations were not carried out under lyophilized conditions, however, as drastically lowering the water content raises the question of proper equilibration of the dynamics of the molecule. It is possible to remove the water in a stepwise manner and equilibrate the

ΔV_{pmf} is the difference in the free energies of the peak and trough of the periodic best-fit curve. Details are provided in the SI text (section S2.5).

system at each step, but properly simulating the lyophilized (or low-hydration) state would involve an understanding of the rate of water removal in experiments and the associated drop in ambient pressure at each step. This is an active area of research.

Table 2. Experimental methyl ^2H R_1 relaxation rates for L5, L8 and L11

Hydration state	R_1 for L5 (in Hz)	R_1 for L8 (in Hz)	R_1 for L11 (in Hz)
Lyophilized ^a	23.1 ± 1.6	23.0 ± 0.5	23.2 ± 0.5
w = 15	21.9 ± 1.0	24.1 ± 1.5	23.8 ± 0.8
w = 20	24.3 ± 1.5	24.3 ± 1.0	24.6 ± 0.6
w = 30	21.3 ± 1.0	25.0 ± 1.1	23.9 ± 0.6

^a Error bars are calculated based on variation from reasonably changing the integration cut-offs for spectra.

In Table 1, it can be seen that the experimental relaxation rates generally increase from those in the lyophilized state to the hydrated states, but that the differences between the 15w

and 20w states are not significantly large. The exception is the C_β nucleus, which seems to exhibit the greatest sensitivity to hydration changes. In Table 2, the methyl ^2H relaxation rates show very little variation with respect to changes in hydration, similar to other such studies (see ref. 67 and references therein). This observation is mostly consistent with the ^{13}C rates for the methyl carbons, except for the fact that the lyophilized rates are significantly lower for the methyl carbons. In the following sections, we endeavor to interpret these results by first accounting for the spin diffusion in the

multiply labeled samples, and then considering the dynamics of the molecule as probed by the MD simulations.

The experiments were also designed to explore different potential dynamic environments with respect to the tetramer bundles: as seen in the model in Figure 1A and as discussed in refs. 22,68, L5 and L7 are believed to be on the outer side of the hydrophobic core, L8 inside the core, with L11 intermediate between the two. However, no consistent discrepancies or trends can be observed in either the ^{13}C or the ^2H relaxation rates.

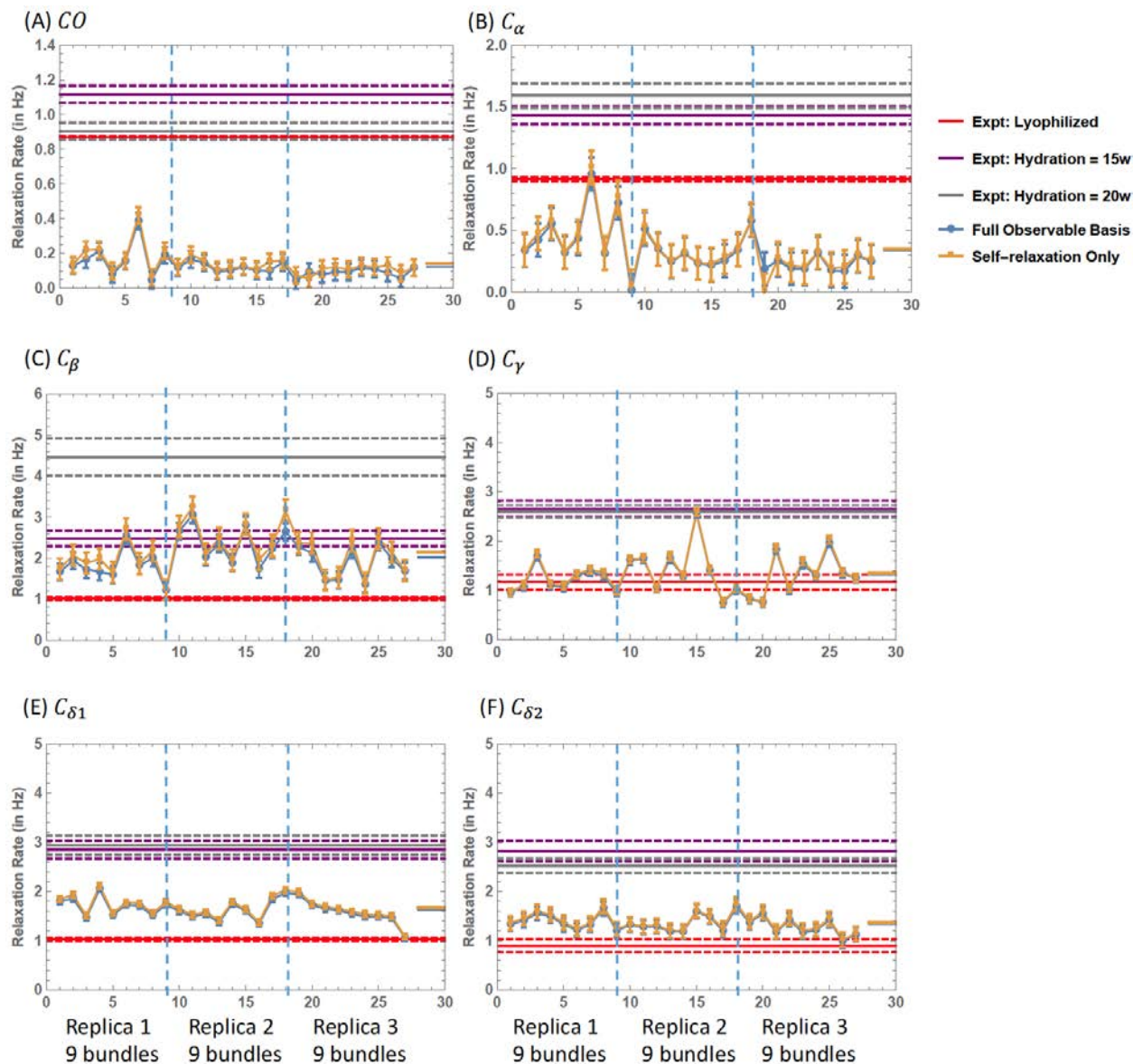


Figure 2. Comparison of ^{13}C relaxation rates for each bundle between the “Nearest-neighbor + Proton” model using the “Full Observable Basis” and using only “Self-relaxation”. The trajectories considered are the three replicas for the L11 residue, 9-bundle, 10w simulation of 50 ns length. Backbone: (A) CO ; (B) C_α . Side chain: (C) C_β ; (D) C_γ ; (E) C_{δ_1} ; and (F) C_{δ_2} . Note that the y-axes have different scales for each of the nuclei. Results for the three replicas were concatenated (separated by vertical blue, dashed lines) to generate the graphs (leading to 27 total bundles). Error bars are generated as the standard deviation of the results for the three replicas. The experimental data is shown as solid horizontal lines (red for the lyophilized data, and purple for the hydrated data), with the error bars shown as parallel, dashed lines. As a visual guide, averages of each of the trend lines are presented as colored dashes in the right part of the frame.

In the following sections, we present the simulated relaxation rates as averages over the rates of the four monomers within the tetramer bundles. As explained in the Methods and SI text (section S2.2), this was done to balance the two requirements of ensemble-averaging in the calculation of the correlation functions, and the preservation of sample heterogeneity. Thus, each tetramer bundle has its own specific set of relaxation rates.

3.1 Effect of using a full observable basis and of varying the size of the spin network for ^{13}C simulations

The relaxation of a nucleus within a network of other magnetically active spins is, in general, not mono-exponential. This means that the longitudinal relaxation of a particular carbon nucleus ($\langle \hat{I}_Z \rangle(t)$) depends not only upon $\langle \hat{I}_Z \rangle(t)$ itself, but also on the relaxation rates of other spin observables of that carbon, and those of the other nuclei. These cross relaxation terms must be accounted for in a complete analysis of a complex, multi-spin system. The resulting equations can be solved by using a matrix equation involving the expectation values of a basis set of observables (details in Methods and SI text). We have tested the impact of including such a basis set of observables by including three 1st order tensor terms $\hat{T}_{0,\pm 1}^{(1)}(I) = \{I_-, I_Z, I_+\}$ for each spin in the system, five 2nd order terms $\hat{T}_{0,\pm 1,\pm 2}^{(2)}(I, S) = \left\{ \frac{1}{2}I_-S_-, \frac{1}{2}(I_ZS_- + I_-S_Z), \frac{1}{\sqrt{6}}(3I_ZS_Z - \vec{I} \cdot \vec{S}), -\frac{1}{2}(I_ZS_+ + I_+S_Z), \frac{1}{2}I_+S_+ \right\}$ for each pair of spins, and one 3rd order term $\hat{T}_0^{(3)}(I, S^{(i)}, S^{(j)}) = \{I_ZS_Z^{(i)}S_Z^{(j)}\}$ for every triplet of spins. Additionally, given our ability to modify the size of the spin network, we explicitly considered three cases involving the 18 nuclei (6 ^{13}C , 1 ^{15}N , and 11 ^1H nuclei) in a leucine residue differing in the connectivity map assumed between the nuclei:

- Nearest-neighbor + Proton interactions: Nearest-neighbor (bonded) interactions involving the ^{13}C and ^{15}N nuclei, and with all ^1H nuclei interacting with each other (17 bonded interactions + 55 ^1H - ^1H interaction pairs = 72 interaction pairs)
- Nearest-neighbor-only interactions: Nearest-neighbor (bonded) interactions only, involving the ^{13}C , ^1H and ^{15}N nuclei, without the ^1H - ^1H interaction pairs (17 bonded interactions)
- Carbon-network interactions: Interactions of the ^{13}C nuclei with all other nuclei (bonded and non-bonded), with no ^1H - ^1H interaction pairs (98 interaction pairs)

We first considered the impact on relaxation rates of using the full basis set of observables to assess the multi-exponential behavior of the longitudinal relaxation $\langle \hat{I}_Z \rangle(t)$ of the carbons. The connectivity map (a) (“Nearest-neighbor + Proton interactions”) was employed in this analysis. The rates calculated using the full observable basis were compared against those calculated only using the self-relaxation rates, i.e. the coefficients of $\langle \hat{I}_Z \rangle(t)$ on the right-hand side of the rate equation (see Section S1.2 in SI). The calculations were done on the trajectories for the L11 residue with 9 bundles at a hydration of 10w. The results are presented in Figure 2.

As can be seen from Figure 2, there are no significant differences in the relaxation rates for the carbon nuclei using the full observable basis as compared to the self-relaxation rates. This, however, is not true for the proton relaxation rates, which do vary to a larger extent between the two calculations (Table

S5). The beta and gamma protons are especially impacted by the inclusion of the full observable basis. The likely interpretation is as follows: given that the carbon energy splitting is much smaller than the proton splitting, and that in a ^{13}C inversion recovery experiment the proton populations are not initially perturbed from equilibrium, there is no significant change in the values of the proton observables, and so the cross-relaxation effects are small. Moreover, the ^{13}C - ^{13}C couplings are significantly weaker than the ^{13}C - ^1H couplings and so no significant exchange occurs between ^{13}C observables. The net result is the predominance of the self-relaxation mechanism. On the other hand, proton-proton couplings are much stronger, and there is significant exchange of spin order between all the coupled protons. Therefore, given our interest only in ^{13}C relaxation, for the rest of the paper we mainly discuss the self-relaxation rates, as these rates can be calculated significantly faster than those using the full observable basis. However, it is to be noted that the full basis must be used to simulate the ^1H inversion recovery rates.

Also, for the calculation of the self-relaxation rates of the ^{13}C nuclei, we found it unnecessary to include the ^1H - ^1H interaction pairs, as these interactions do not seem to contribute in 2nd-order BRW perturbation theory for carbon self-relaxation. We ran a comparative set of self-relaxation calculations for the “Nearest-neighbor + Proton” case and for the “Nearest-neighbor-only” case and found no difference between the two sets of ^{13}C results (Table S6). The ^1H nuclei, however, do understandably show a difference in rates between the two cases.

Next, we tested the “Carbon-network” model relative to the “Nearest-neighbor-only” model and found that the two methods produced very similar relaxation rates (Figure S3 in the SI). The rates from the “Carbon-network” model do produce uniformly higher relaxation rates, as should be expected from increasing the connectivity of the ^{13}C spins, but the differences between the two results are within the error bars. We conclude that, in the interests of computational efficiency, it is sufficient to consider the simplest “Nearest-neighbor-only” model without significant loss of information.

The final conclusions from our studies into the size and connectivity of the spin network are that (a) considering the full observable basis is important in the calculation of the ^1H relaxation rates, but not for the ^{13}C rates, where the self-relaxation rates almost entirely account for the relaxation; and (b) a simple model involving the nearest-neighbor bonded interactions for the ^{13}C nuclei, with no ^1H - ^1H interactions, is sufficient to explain almost all the relaxation behavior of the ^{13}C nuclei.

It is essential to note that the conclusions of this section come with an important caveat: comparisons have only been made between the incoherent relaxation components of the rates, while the effect of the coherent proton-driven spin diffusion (PDS) must be considered as well. In fact, in section 3.3 we show that PDS creates a discrepancy between the different spin network models, at least for certain nuclei (the C_β , and, to a lesser degree, the C_α nuclei) and for the lower hydration state tested. However, the results for the incoherent relaxation effects are expected to hold true for many nuclei and for significantly higher hydration levels.

3.2 Modification of the AMBER ff14SB dihedral potential energy for the $C_\gamma - C_\delta$ bond

We carried out all our simulations using the AMBER ff14SB force field⁶¹ (hereafter abbreviated as FF14). We calculated both the C_δ and H_δ relaxation rates for the L11 residue based on this force field, and compared them to the respective experimental values (Figures 3 and 4; the L11 residue is unique in our data set for having both ^{13}C and ^2H data available). It was found that the simulations led to a reasonable fit for one of the C_δ nuclei, but the simulated rates were significantly higher than their experimental counterparts for the remaining three comparisons.

Accordingly, recognizing the sensitivity of the methyl carbon relaxation rates primarily to the rate of methyl group rotation about the $C_\gamma - C_\delta$ bonds, we lowered the potential energy barrier for the methyl group rotation by a factor of 0.7. In the simulations for the 18-bundle case shown in Fig. 3, for both the $C_{\delta 1}$ (Fig. 3A) and $C_{\delta 2}$ (Fig. 3B) nuclei, the following features are prominent. Firstly, the simulations run with the unmodified FF14 (green data points) show a higher relaxation

rate than those with the barrier reduced by a factor of 0.7. While for $C_{\delta 2}$, the experimental data are better fit by the FF14 rates, the results for $C_{\delta 1}$ are unclear. The second interesting aspect relates to the absence of any significant effect of changing the water content on the simulated rates of the C_δ nuclei (this agrees with the ^2H data-based findings of ref. ⁶⁹). This is because the C_δ nuclei are primarily impacted by the methyl group rotation, which modulates the orientations of the three protons directly attached to the C_δ 's. However, the experimental rates increase from the lyophilized to the hydrated states. To reconcile the two, it may be noted that the experimental rates appear to stabilize once a certain hydration level is reached. It is possible that the simulated conditions, with the lowest hydration at 10w, are within the region of stabilized rates. Testing this hypothesis would require lowering the simulation hydration to much lower levels, which is difficult for the reasons mentioned earlier.

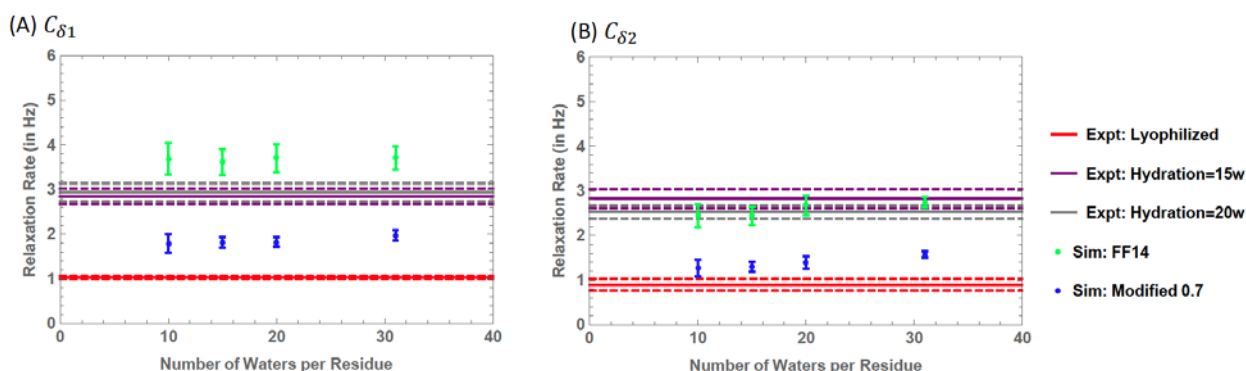


Figure 3. Median L11 $^{13}\text{C}_\delta$ relaxation rates for single tetramer bundles from the 18-bundle simulations as a function of water content and scaling of the $C_\gamma - C_\delta$ dihedral torsion parameters relative to the values in FF14: (A) $C_{\delta 1}$; and (B) $C_{\delta 2}$. “Xw” implies that there are X waters per residue of LK α 14; “FF14” refers to the unmodified potential energy scaling of FF14 (dashed trend lines) and “Modified 0.7” implies a scaling down of the potential energy (and thus of the barriers) by a factor of 0.7 relative to that of FF14 (solid trend lines). Three replicas were generated for each 50 ns simulation and the medians shown are for the combined results of all three replicas (54 bundles in total). The bundle-to-bundle variation is captured by the interquartile range shown as error bars for each data point. The experimental data is shown as solid horizontal lines (red for the lyophilized, purple for w=15, and grey for w=20 data), with the error bars shown as parallel, dashed lines.

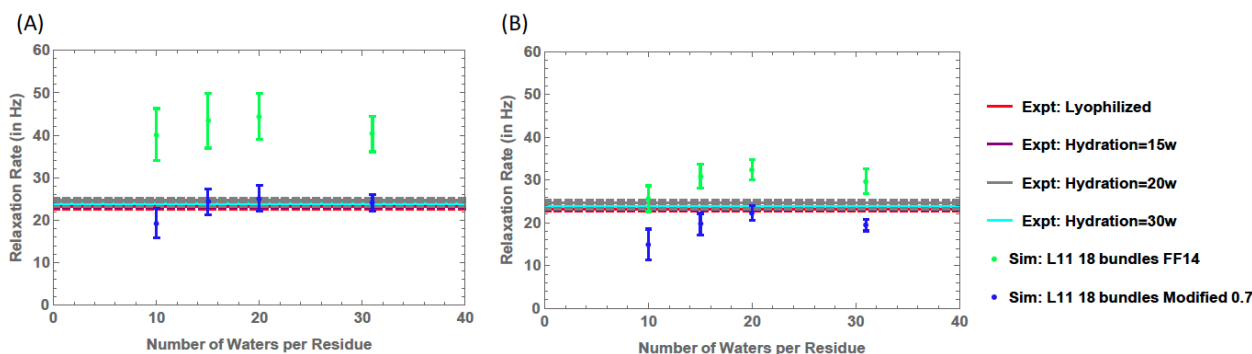


Figure 4. Median L11 $^2\text{H}_\delta$ relaxation rates from the 18-bundle simulations as a function of water content and scaling of the $C_\gamma - C_\delta$ dihedral torsion parameters relative to the values in FF14: (A) $H_{\delta 1}$ and (B) $H_{\delta 2}$ methyl group relaxation rates for the unmodified FF14 case and the modified 0.7 scaling case. “Xw” implies that there are X waters per residue of LK α 14; “FF14” refers to the unmodified potential energy scaling of FF14 (dashed trend lines) and “Modified 0.7” implies a scaling down of the potential energy (and thus of the barriers) by a factor of 0.7 relative to that of FF14 (solid trend lines). The relaxation rates reported are the averaged rates over the three deuterons in a given methyl group. Three replicas were generated for each 50 ns simulation and the medians shown are for the combined results of all three replicas (54 bundles in total). The bundle-to-bundle variation is captured by the interquartile range shown as error bars for each data

point. The experimental data is shown as solid horizontal lines (red for the lyophilized, purple for w=15, grey for w=20 data, and cyan for w=30 data), with the error bars shown as parallel, dashed lines.

To establish the better-fitting methyl rotation barrier of the two values, we examined the ^2H relaxation rates for L11 (Figure 4). The modified 0.7 scaling case shows better fits to the experimental rates for both methyl groups. Moreover, changing the water content seems to have little impact on both the experimental and simulation relaxation rates. The two possible contributions to the discrepancy between the ^{13}C and ^2H experiments are the different inherent sensitivities of the two measurements to motional rates, and the fact that the ^2H labeling scheme does not discriminate between the two methyl groups and so the experimental results are averages over the rates of both sets of deuterons. The averaging over the two methyl groups may wash out any observed differences.

To further investigate the impact of the barrier scaling, we also compared the ^2H data for the L5 and L8 residues (Figure S4). The results for these residues also show an excellent fit of the modified 0.7 simulations to the experimental data. We conclude that the scaling by a factor of 0.7 better matches the conditions across multiple residues and two different experiments. Accordingly, subsequent calculations are carried out with the barrier scaled down by a factor of 0.7.

Many interesting points can be noted in this study. It is clear that, with the lowered barrier, some of the simulated C_δ rates are lower than the experimental values. We place more confidence in the ability of the ^2H data to uniquely determine the methyl rotation rates, as ^{13}C data may be further influenced by slower motions closer to the Larmor frequency of ^{13}C . Thus, the lack of fit of the C_δ rates to experiments is likely due to the existence of slower (relative to the methyl rotation) sidechain motions not captured in the MD simulations. However, it remains unclear as to whether such a difference in contributing motions could also explain the fact that there is no difference between the experimental ^2H rates for the lyophilized and hydrated cases, while there exists a significant discrepancy between the two for the ^{13}C rates.

Additionally, in both the ^{13}C and ^2H data, there is a clear distinction between the two potential barrier values. We interpret this as suggesting that the methyl groups need to rotate faster about their axis of symmetry relative to the FF14 rates in order to match experimental conditions. Interestingly, this faster rotation translates into a lowered relaxation rate, arising from the fact that the fast methyl rotation falls into a range of motional rates where the dipolar coupling between the C_δ and H_δ nuclei is partially pre-averaged and thus reduced. In addition to the fact that we obtained a better fit with the scaling of 0.7, the particular choice of scaling is further corroborated by three independent pieces of evidence. Firstly, we carried out a density functional theory (DFT) simulation of the $C_\gamma - C_\delta$ dihedral potential energy surface for an isolated leucine and found a value for the barrier to be about 3.2 – 3.3 kcal/mol (Figure S5A). Secondly, work by Xue et al⁷⁰ found methyl barriers for the “fluid” hydrophobic core in α -spc SH3 averaging 2.8 ± 0.9 kcal/mol (experimental data), 3.1 kcal/mol (MD-based potential of mean force calculations), 3.6 kcal/mol (so-called “rigid barrier” calculation on an X-ray structure) and 4.1 kcal/mol (rigid barrier calculation on an NMR structure). The authors claim that NMR structures, for the most part, are obtained by first simulated annealing the structures down to 0 K, thus leading to close-packing of side-chains and increased methyl rotation barriers. Thus, the calculations from NMR structures are

less reliable for reporting room temperature behavior. In fact, NMR structures optimized using the GROMOS MD package treat the methyl groups as expanded spheres and evade the issue of close-packing, and correspondingly, calculations yield barriers in the range of 3.0 – 3.6 kcal/mol. We hypothesized that the four monomers in a tetramer bundle would allow for a much more dynamic hydrophobic core and, accordingly, we compared our results to the range of 2.8 – 3.6 kcal/mol. Scaling the potential by a factor of 0.7 brought the barrier to a comparable range, as calculated using the potential of mean force (the barrier is labeled as ΔV_{pmf} ; shown in Figure S5B). For all the 18-bundle simulations shown above, ΔV_{pmf} calculated using the dihedral angle distributions (from all 72 instances of L11 in a given simulation of 18 bundles) ranged from 4.0 – 4.3 kcal/mol ($C_{\delta 1}$) and 3.8 – 4.0 kcal/mol ($C_{\delta 2}$) in the unmodified FF14 case, to 3.3 – 3.4 kcal/mol ($C_{\delta 1}$) and 3.0 – 3.1 kcal/mol ($C_{\delta 2}$) for the 0.7 scaling case. It is therefore clear that the simulations that best fit the experimental data, those with a methyl barrier scaled by 0.7, correspond more closely to previously studied methyl rotation barriers in “fluid” hydrophobic cores. Thirdly, recent work by Hoffmann et al¹⁴ has also found the AMBER ff14SB*-ILDN methyl group potential energy barriers for several amino acids to be higher than both those suggested by CCSD(T) simulations on isolated blocked dipeptides (of the form ACE-X-NME), and NMR data in ubiquitin. They find that the AMBER methyl barriers need to be reduced by a factor of about 0.82-0.83 to match the CCSD(T) calculations. While there is a difference between their scaling and ours, we note that the barriers may be influenced by the particular molecular context of the side chain. Also, our ^{13}C experimental rates could still be fit by a barrier of 0.8 (by interpolating between the FF14 and 0.7 scaled values), so there is a range of values that could reasonably fit the data.

It is worth remarking on the fact that the two methyl groups seem to have slightly different ^{13}C and ^2H relaxation rates. The experimental data for the methyl carbons also corroborates the difference. This is especially true for the hydrated sample. The facts that the ΔV_{pmf} barriers are also different and that the lower ΔV_{pmf} corresponds universally to a lower relaxation rate indicates that this is a genuine discrepancy between the two methyl groups in the MD simulations, rather than a mere artifact of the relaxation rate calculation. The difference in ΔV_{pmf} is also observed in the side chains of α -spc SH3 reported by Xue et al⁷⁰. Currently, the exact nature of the causative physical conditions in the simulations is unclear. Some preliminary calculations involving the modification of other torsion angles along the side chain suggest that the difference may have to do with the conformations of the side chain asymmetrically affecting one methyl group and not the other. Further work is required to tease out the exact contributing causes.

3.3 Variation of relaxation rates with water content and the impact of PDS

The ^{13}C relaxation rates for the 18-bundle simulations with differing water contents are shown for each carbon nucleus of residues L11 (Figure 5) and L7 (Figure S6).

An increase in the water content almost universally causes an increase in the relaxation rate for the CO , C_α , C_β , and C_γ . While this is likely caused by an increase in mobility, in gen-

eral, the response of R_1 depends on the specific rates, amplitudes and degree of symmetry of the motion. In fact, a counterexample is provided above by the reduction in relaxation rates upon lowering the $C_\gamma - C_\delta$ dihedral barrier (thus speeding up the rotation). Thus, a combination of contributing factors must be considered in the interpretation. Upon increasing the ambient water, both the amplitude and rate of bulk motions of the molecule (including both tumbling of the entire tetramer bundle and of the monomer relative to the tetramer axis) increase. The motion will also become increasingly more isotropic. All of these modifications of the bulk motions would reduce the degree of autocorrelation of the bond orientation. Consequently, in order for R_1 to increase under these conditions, the bulk motional rates would need to be on the slower side of the window of frequency-sensitivity of R_1 . The local

motions of the atoms, however, may indeed be more complicated.

The agreement between experiment and simulations is different for different nuclei, albeit with a consistent underestimation of the experimental values at the same hydration. The simulated CO rates are significantly lower than the experimental rates. For C_α , the simulated rates increase with an increase in hydration, but at a faster rate than the experiments. The rate at the highest hydration approaches the experimental values. For C_β , the match to the experiments is the best out of all the nuclei. The C_γ data show a lower rate of increase with hydration, which, taken together with the previous discussion on the underestimation of C_{δ_1} and C_{δ_2} rates, implies missing sidechain motions involving the C_γ , C_{δ_1} and C_{δ_2} nuclei (for example, see ref. ⁶⁸).

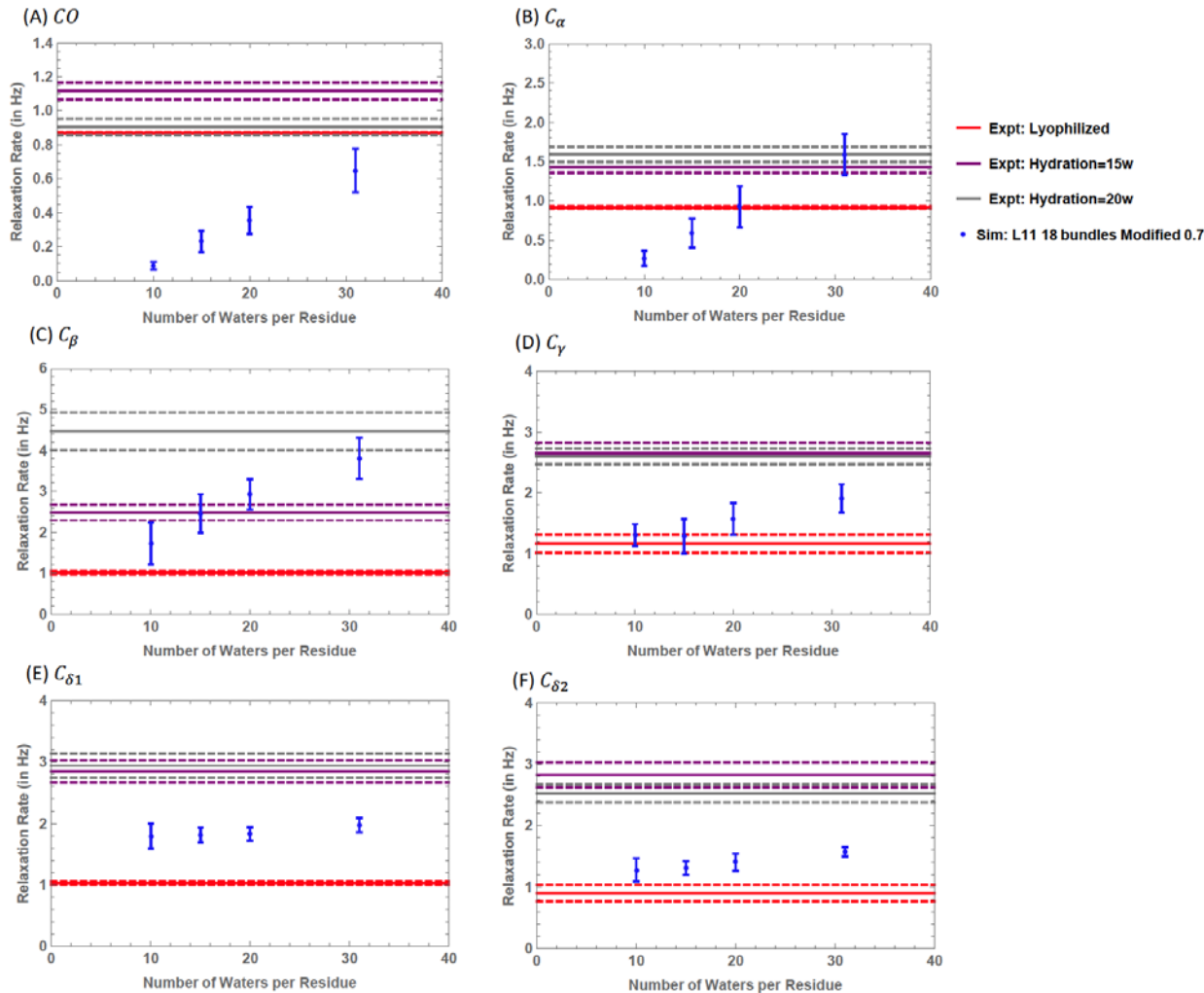


Figure 5. Median L11 ^{13}C relaxation rates for single tetramer bundles from the 18-bundle simulations as a function of water content: Backbone: (A) CO ; (B) C_α . Side chain: (C) C_β ; (D) C_γ ; (E) C_{δ_1} ; and (F) C_{δ_2} . Note that the y-axes have different scales for each of the nuclei. “Modified 0.7” implies a scaling down of the $C_\gamma - C_\delta$ dihedral potential energy (and thus of the barriers) by a factor of 0.7 relative to that of FF14. “Xw” implies that there are X waters per residue of LK α 14. Three replicas were generated for each 50 ns simulation and the medians shown are for the combined results of all three replicas (54 bundles in total). The bundle-to-bundle variation is captured by the interquartile range shown as error bars for each data point. The experimental data is shown as solid horizontal lines (red for the lyophilized, purple for w=15, and grey for w=20 data), with the error bars shown as parallel, dashed lines.

Before exploring the discrepancies further, we need to consider the effects of proton-driven spin diffusion (PDS) on the relaxation rates. Extensive previous literature^{24,50,51} has shown

the importance of coherent spin effects to relaxation in a solid-state system, in addition to the Bloch-Redfield-Wangsness solution-like “relaxation” component of spin evolution con-

considered so far. Conceptually, this corresponds to the coherent exchange of energy between carbon nuclei or any other set of equivalent spins that are connected to a shared proton network when not averaged by rapid motions. The effect of PDS is expected to be especially relevant in the case of slow MAS (where the coherent exchanges are incompletely averaged out). We therefore carried out a series of calculations of the effects of PDS for the lowest hydration MD simulations, and for both the “Nearest-neighbor + Proton” and “Carbon-network” interaction models.

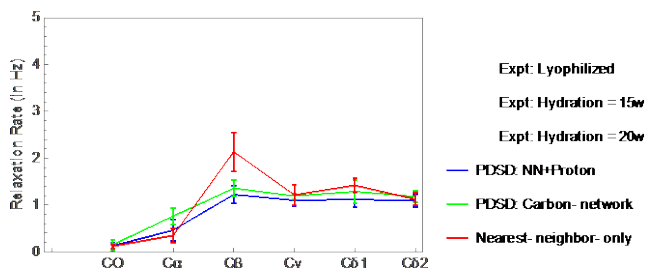


Figure 6. Effect of PDS on relaxation rates of the ^{13}C nuclei in L11 for two connectivity models: the Nearest-neighbor + Proton interaction model (“NN+Proton”) and the Carbon-network model. The simulations were run for the 9-bundles case at a hydration of $w=10$, with the methyl barrier reduced by a factor of 0.7 relative to that of FF14. The experimental rates are shown as dashed lines with the error regions filled out.

In Figure 6, calculations of relaxation rates with PDS included were carried out for the 9-bundles case at a hydration of $w = 10$, the lowest hydration considered in our MD simulations. We chose this hydration level under the assumption that the effects of PDS will be maximized when motional averaging is minimized. We also restricted the analysis to the 9-bundles case due to the time-consuming nature of the PDS simulations. The Nearest-neighbor + Proton interaction model allows for coherent exchanges through the inclusion of interactions of the form “Carbon(i)-Proton(i)-Proton(j)-Carbon(j)”, while the Carbon-network model includes more direct carbon-carbon interactions through terms of the form “Carbon(i)-Proton-Carbon(j)”. We thus are able to test the impacts of two different forms of spin connectivity on the observed PDS. These models were compared to the case of nearest-neighbor-only interactions with self-relaxation terms. As noted earlier, if PDS has no effect, these rates should be very similar.

Some interesting aspects may be noted:

(a) PDS causes the relaxation rates to be homogenized to a certain degree, with the effect on the C_β nucleus being the most pronounced. In fact, the relative homogeneity in the simulated side-chain relaxation rates approaches the similarity in the experimental rates for the side-chain nuclei for the lyophilized and $w=15$ cases. We also noticed that the relative pattern of the rates among the nuclei for the “Nearest-neighbor-only with only self-relaxation terms” case mirrors that of the $w=20$ experimental data (with, of course, different absolute magnitudes). We hypothesize that this may reflect the fact that at higher hydration the effect of PDS is averaged out more strongly by motion, leading to the incoherent self-relaxation terms dominating. Further studies at a higher hydration will be able to test this hypothesis.

(b) The C_β relaxation rate is reduced for the Nearest-neighbor + Proton and Carbon-network models, relative to the Nearest-neighbor-only case with only self-relaxation terms. The C_α nucleus also shows a slight increase in value, approaching closer to the experimental rate for the lyophilized case.

(c) The CO rates are not substantially affected, and neither are the rates for the C_γ , $C_{\delta 1}$ and $C_{\delta 2}$ nuclei. This implies that the results of the previous section on the lowering of the methyl barriers are still expected to hold true.

(d) The difference between the PDS results for the two connectivity models is not very significant. While there is a slight increase in the rate for the nucleus for the Carbon-network model, the magnitude of the change relative to the other model is small. It is therefore currently unclear how much the two different sets of interactions impact the PDS contribution.

We therefore conclude that some of the discrepancy in the rates between simulation and experiment for the C_β nucleus, and, to a lesser degree, for the C_α nucleus can be explained by the impact of PDS.

There are still clearly remaining differences between experiment and simulation for many of the nuclei. This is especially true for the backbone CO rates (and for the residual PDS-adjusted C_α rates). It is interesting to compare these results to those of Lewandowski et al²⁴, who find similar homogeneity in the dry crystalline $[\text{U-}^{13}\text{C}, ^{15}\text{N}]\text{-Ala } ^{13}\text{C}$ rates at a MAS spinning rate of 16 kHz. However, they note that increasing the spinning rate to 60 kHz leads to a reduction in the homogenizing effect of PDS and a recovery of the site-specific relaxation rates. The pattern of the site-specific rates shows a remarkable similarity to the MD-based simulated rates in our work (very low CO and C_α rates; very high C_β rates). The reconciliation of the two sets of results, theirs showing that PDS has a significant impact for the backbone, while our results indicating a more modest effect, could possibly be done through a consideration of the particular experimental conditions. Lewandowski et al²⁴ use crystalline samples of uniformly labeled alanine, while our MD simulations were carried out on single residues within hydrated peptides. In the crystalline state, the rigidity of the monomers, as well as the relatively close packing of the labeled monomers should lead to an increase in the impact of both intra- and inter-molecular PDS. We expect the labeled leucines in our studies, buried within the hydrophobic bundle centers, to be on average sufficiently distant from other labeled residues in other monomers. Therefore, in order to speculate on the deviation of the simulated backbone rates from experiment, it helps to first parse out the various possible spin contributions: (a) The backbone nuclei could dipolar-interact transiently with the ambient water molecules, especially those with longer protein-surface-residence times⁷¹, thereby increasing the number of dipolar-coupled partners relative to our calculation. We did not account for these coupling partners due to the difficulty of tracking the distance and position of all water molecules within a certain cut-off radius. However, given the fact that the simulated CO relaxation rates are much smaller than those for the extremely dry lyophilized samples, there seems to be a need for an additional contribution. (b) There could be a contribution from other spins in the system that are not included; in the lyophilized and low hydration samples, there could perhaps be relaxation arising from proximity of neighboring molecules if the

density is sufficiently high. This high-density scenario cannot be tested with the current MD set-up. (c) Finally, the CO nucleus might also be relaxed by the amide proton from the previous residue (i.e. the peptide-bonded amide group). Given the planar geometry the motion of the carbonyl C and amide H may have highly correlated motions and thus may be very efficient at mutual relaxation. In order to test this possibility, we ran simulations with the neighboring amide H included but found that this additional $CO - H$ interaction only has a contribution of about 5-10% to the simulated rate, which is insufficient to explain the overall difference.

Of course, the MD simulations may underestimate the degree of motion in the both the backbone and side chain. This is an important point to consider, as MD simulations are frequently used in the validation of and comparison to experimental data. It is possible that there are motions occurring on a slower time scale that are not captured in the 50 ns simulations considered here. However, given that the range of time scale sensitivity of the ^{13}C R_1 rates does not extend significantly beyond $(\omega_{^{13}C} \tau_{Larmor})^{-1} \approx (125 \text{ MHz})^{-1} \sim 10 \text{ ns}$, these motions may not be significant enough to explain the discrepancy. In fact, when we extended one set of simulations to a length of 150 ns (Figure S2), the rates were nearly identical to those obtained with the 50 ns simulations. This leaves the possibility of an inherent deficiency in the dynamics as described by MD (see also Best et al.⁵ where an order parameter analysis of backbone motions shows deviations in certain regions). Furthermore, there may be additional modes of dynamics arising due to contact between molecules at low hydration levels. This can be tested only by further reduction of the water content to near-lyophilized states.

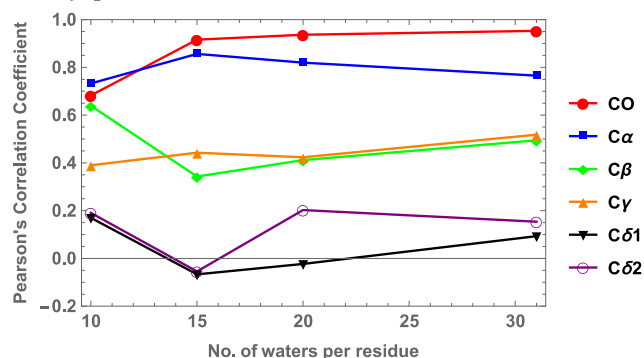


Figure 7. Pearson’s correlation coefficient for the MD simulation relaxation rates of the various ^{13}C nuclei in L11 with those of L7, as a function of the number of waters per residue in the simulation box. The simulations were run for the 18-bundles case, with the methyl barrier reduced by a factor of 0.7 relative to that of FF14. The correlation coefficients are calculated over all three replicas of the simulation conditions.

Considering the MD simulations alone, we observed that the backbone relaxation rates for L7 and L11 are highly correlated (Figure 7), calculated as the correlation of all the tetramer-bundle-averaged rates across the three replicas (=54 rates for the 18-bundle case). This implies that the motions relevant to the relaxation rates are similar for the backbones of both residues: these include both the overall motion of the tetramer bundles, as well as the local torsional modes. The degree of correlation between the L7 and L11 CO rates also seems to increase with increasing water content, suggesting the im-

portance of the water-induced motions to the relaxation rates. To confirm that the high correlations were not accidental, arising from a broad similarity of the backbone rates for all bundles, we calculated the average correlation of the L11 rates with 100 permutations of the L7 rates and found this value to be close to 0. Understandably, the correlation decreases from the backbone nuclei moving towards the sidechain nuclei. These effects arise due to a combination of two contributions: the different degrees of insulation of the hydrophobic side chains of L7 and L11 from the external environment due to disparate burial within the hydrophobic core, and the increase of the relaxation effects of local side chain motions relative to the background of motions experienced by the backbone.

In summary, it is apparent that there are likely deficiencies in the MD description of the dynamics of the backbone and sidechain nuclei (at least as described by AMBER ff14SB with a TIP3P water model). It is, however, interesting to note that the simulated rates begin to approach the experimental values for higher hydration levels. This could signal that restricted-tumbling solid-state conditions present unique challenges to standard MD approaches, and that the approach to the solution state recovers the applicability of the simulations. However, the observed discrepancies offer a powerful means of identifying degrees of freedom essential to the dynamics of the residues: for example, the fact that C_β rates match the data well, while the C_γ , C_{δ_1} and C_{δ_2} rates underestimate the experiments, indicates that there are likely $C_\beta - C_\gamma$ torsional conformations relevant to relaxation that are not captured by the MD (as seen in ref. ⁶⁸).

4. CONCLUSION

We have endeavored to place on firm footing the simulation of solid-state samples using a combination of two different types of NMR experiments at varying hydration levels with MD trajectories of arrays of peptide aggregates. In doing so, we have uncovered a deeper understanding of the spin interactions most relevant to the NMR relaxation rates, determining that the ^{13}C rates are largely determined by the self-relaxation terms in the rate matrix and the nearest-neighbor interactions. Additionally, methyl group rotation (i.e. $C_\gamma - C_\delta$ dihedral angle) barriers for leucines in the MD force fields need to be reduced by a factor of 0.7 relative to the AMBER ff14SB value to match the experimental data. We also found that the methyl relaxation rates were in general independent of the degree of hydration. Deficiencies in the MD description of the backbone atoms were identified, showing significantly less motion than expected on the basis of the experiments. We have also shown how the methods used here may serve as guides towards the identification of degrees of freedom relevant to relaxation mechanisms.

Having constructed this framework, the goals for future work in each of the three domains of interest are:

1. **Details of motions:** To study possible concerted dynamics in hydrophobic cores (for example, see ref. 72), the impact of water density on structure and dynamics in solids and high-concentration solutions, and the details of intra- and inter- aggregate dynamics.
2. **Spin dynamics:** To consider the role of water molecules in relaxation analysis, including as a function of changes in water density at interfaces relative to the bulk, as well as to employ in the analysis of multiply-labeled samples.

3. **MD simulations:** To benchmark the trajectories using several independent sets of data and, ultimately, find non-context-specific modifications to the force fields necessary to match a broad spectrum of experimental data.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.

Detailed theoretical and computational methods, tests of the code and convergence of the MD trajectories, error analysis, supplementary results (PDF)

AUTHOR INFORMATION

Corresponding Author

* Gary P. Drobny

Department of Chemistry

University of Washington

Box 351700

Seattle, WA 98195-1700, U.S.A.

Ph: 206-685-2052

E-mail: drobny@chem.washington.edu

* Jim Pfaendtner

Department of Chemical Engineering

University of Washington

Box 351750

Seattle, WA 98195-1750, U.S.A

Ph: 206-616-8128

E-mail: jpfaendt@uw.edu

Present Addresses

† Department of Molecular Biophysics and Biochemistry, Yale University, 266 Whitney Avenue, PO Box 208114, New Haven, CT 06520-8114.

†† Starbucks Coffee Company, 2401 Utah Avenue S, Seattle, WA 98134.

††† School of Molecular Sciences, Arizona State University, Physical Sciences Center PSD-D102, Tempe, AZ 85287.

Author Contributions

‡ These authors contributed equally.

ACKNOWLEDGMENT

G.P.D. acknowledges NSF grant MCB 1715123 and NIH grant R01-GM109417. J. P. acknowledges NSF grant CBET-1264459. The authors would like to thank Vance Jaeger and Kayla Sprenger for helpful discussions. We would also like to thank Ariel Zane for providing additional insight regarding published data.

ABBREVIATIONS

ssNMR, Solid-state NMR; MD, Molecular dynamics.

REFERENCES

- Brüschweiler, R.; Case, D. A. Collective NMR Relaxation Model Applied to Protein Dynamics. *Phys. Rev. Lett.* **1994**, *72* (6), 940–943.
- Lienin, S. F.; Brüschweiler, R. Characterization of Collective and Anisotropic Reorientational Protein Dynamics. *Phys. Rev. Lett.* **2000**, *84* (23), 5439–5442.
- Case, D. A. Molecular Dynamics and NMR Spin Relaxation in Proteins. *Acc. Chem. Res.* **2002**, *35* (6), 325–331.
- Best, R. B.; Rutherford, T. J.; Freund, S. M. V.; Clarke, J. Hydrophobic Core Fluidity of Homologous Protein Domains: Relation of Side-Chain Dynamics to Core Composition and Packing. *Biochemistry* **2004**, *43* (5), 1145–1155.
- 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* (1), 185–203.
- Robustelli, P.; Stafford, K. A.; Palmer, A. G. Interpreting Protein Structural Dynamics from NMR Chemical Shifts. *J Am Chem Soc* **2012**, *134* (14), 6365–6374.
- Salmon, L.; Pierce, L.; Grimm, A.; Ortega Roldan, J. L.; Mollica, L.; Jensen, M. R.; Van Nuland, N.; Markwick, P. R. L.; McCammon, J. A.; Blackledge, M. Multi-Timescale Conformational Dynamics of the SH3 Domain of CD2-Associated Protein Using NMR Spectroscopy and Accelerated Molecular Dynamics. *Angew. Chemie - Int. Ed.* **2012**, *51* (25), 6103–6106.
- Robustelli, P.; Trbovic, N.; Friesner, R. A.; Palmer, A. G. Conformational Dynamics of the Partially Disordered Yeast Transcription Factor GCN4. *J. Chem. Theory Comput.* **2013**, *9* (11), 5190–5200.
- Giambasu, G. M.; York, D. M.; Case, D. A. Structural Fidelity and NMR Relaxation Analysis in a Prototype RNA Hairpin. *Rna* **2015**, *21* (5), 963–974.
- Lindorff-Larsen, K.; Best, R. B.; Depristo, M. A.; Dobson, C. M.; Vendruscolo, M. Simultaneous Determination of Protein Structure and Dynamics. *Nature* **2005**, *433* (7022), 128–132.
- Salmon, L.; Giambasu, G. M.; Nikolova, E. N.; Petzold, K.; Bhattacharya, A.; Case, D. A.; Al-Hashimi, H. M. Modulating RNA Alignment Using Directional Dynamic Kinks: Application in Determining an Atomic-Resolution Ensemble for a Hairpin Using NMR Residual Dipolar Couplings. *J. Am. Chem. Soc.* **2015**, *137* (40), 12954–12965.
- Ingolfsson, H. I.; Li, Y.; Vostrikov, V. V.; Gu, H.; Hinton, J. F.; Koeppel, R. E.; Roux, B.; Andersen, O. S. Gramicidin A Backbone and Side Chain Dynamics Evaluated by Molecular Dynamics Simulations and Nuclear Magnetic Resonance Experiments. I: Molecular Dynamics Simulations. *J. Phys. Chem. B* **2011**, *115* (22), 7417–7426.
- Bottaro, S.; Bussi, G.; Kennedy, S. D.; Turner, D. H.; Lindorff-Larsen, K. Conformational Ensemble of RNA Oligonucleotides from Integrating NMR and Molecular Simulations. *Sci. Adv.* **2018**, *4*, eaar8521:1-8.
- 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* (19), 5038–5048.
- Chen, P.; Hologne, M.; Walker, O.; Hennig, J. Ab Initio Prediction of NMR Spin Relaxation Parameters from Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2018**, *14* (2), 1009–1019.
- Philips, A.; Marchenko, A.; Truffandier, L. A.; Autschbach, J. Quadrupolar NMR Relaxation from Ab Initio Molecular Dynamics: Improved Sampling and Cluster Models versus Periodic Calculations. *J. Chem. Theory Comput.* **2017**, *13* (9), 4397–4409.
- Nguyen, H. D.; Hall, C. K. Molecular Dynamics Simulations of Spontaneous Fibril Formation by Random-Coil Peptides. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (46), 16180–16185.
- Petrov, D.; Zagrovic, B. Are Current Atomistic Force Fields

- Accurate Enough to Study Proteins in Crowded Environments? *PLoS Comput. Biol.* **2014**, *10* (5).
- (19) Buffy, J. J.; Waring, A. J.; Hong, M. Determination of Peptide Oligomerization in Lipid Bilayers Using 19F Spin Diffusion NMR. *J. Am. Chem. Soc.* **2005**, *127* (12), 4477–4483.
- (20) Su, Y.; Doherty, T.; Waring, A. J.; Ruchala, P.; Hong, M. Roles of Arginine and Lysine Residues in the Translocation of a Cell-Penetrating Peptide from 13C, 31P, and 19F Solid-State NMR. *Biochemistry* **2009**, *48* (21), 4587–4595.
- (21) Armstrong, D.; Zidovetzki, R. Helical Wheel Projections <http://rslab.ucr.edu/scripts/wheel/wheel.cgi>.
- (22) Zane, A. C.; Michelet, C.; Roehrich, A.; Emani, P. S.; Drobny, G. P. Silica Morphogenesis by Lysine-Leucine Peptides with Hydrophobic Periodicity. *Langmuir* **2014**, *30* (24), 7152–7161.
- (23) Dick-Perez, M.; Wang, T.; Salazar, A.; Zabolina, O. A.; Hong, M. Multidimensional Solid-State NMR Studies of the Structure and Dynamics of Pectic Polysaccharides in Uniformly 13C-Labeled Arabidopsis Primary Cell Walls. *Magn. Reson. Chem.* **2012**, *50* (8), 539–550.
- (24) Lewandowski, J. R.; Sein, J.; Sass, H. J.; Grzesiek, S.; Blackledge, M.; Emsley, L. Measurement of Site-Specific 13C Spin-Lattice Relaxation in a Crystalline Protein. *J. Am. Chem. Soc.* **2010**, *132* (24), 8252–8254.
- (25) Cheung, T. T. P. Spin Diffusion in NMR in Solids. *Phys. Rev. B* **1981**, *23* (3), 1404–1418.
- (26) Kumar, A.; Christy Rani Grace, R.; Madhu, P. K. Cross-Correlations in NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* **2000**, *37* (3), 191–319.
- (27) Hogben, H. J.; Krzystyniak, M.; Charnock, G. T. P.; Hore, P. J.; Kuprov, I. Spinach - A Software Library for Simulation of Spin Dynamics in Large Spin Systems. *J. Magn. Reson.* **2011**, *208* (2), 179–194.
- (28) Tošner, Z.; Andersen, R.; Stevansson, B.; Edén, M.; Nielsen, N. C.; Vosegaard, T. Computer-Intensive Simulation of Solid-State NMR Experiments Using SIMPSON. *J. Magn. Reson.* **2014**, *246*, 79–93.
- (29) Gsponer, J.; Habertür, U.; Cafilisch, A. The Role of Side-Chain Interactions in the Early Steps of Aggregation: Molecular Dynamics Simulations of an Amyloid-Forming Peptide from the Yeast Prion Sup35. *Proc. Natl. Acad. Sci.* **2003**, *100* (9), 5154–5159.
- (30) Hwang, W.; Zhang, S.; Kamm, R. D.; Karplus, M. Kinetic Control of Dimer Structure Formation in Amyloid Fibrillogenesis. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (35), 12916–12921.
- (31) Ma, B.; Nussinov, R. Simulations as Analytical Tools to Understand Protein Aggregation and Predict Amyloid Conformation. *Curr. Opin. Chem. Biol.* **2006**, *10* (5), 445–452.
- (32) Baio, J. E.; Zane, A.; Jaeger, V.; Roehrich, A. M.; Lutz, H.; Pfaendner, J.; Drobny, G. P.; Weidner, T. Diatom Mimics: Directing the Formation of Biosilica Nanoparticles by Controlled Folding of Lysine-Leucine Peptides. *J. Am. Chem. Soc.* **2014**, *136* (43), 15134–15137.
- (33) Liao, C.; Esai Selvan, M.; Zhao, J.; Slimovitch, J. L.; Schneebeli, S. T.; Shelley, M.; Shelley, J. C.; Li, J. Melittin Aggregation in Aqueous Solutions: Insight from Molecular Dynamics Simulations. *J Phys Chem B* **2015**.
- (34) Dalgicdir, C.; Globisch, C.; Peter, C.; Sayar, M. Tipping the Scale from Disorder to Alpha-Helix: Folding of Amphiphilic Peptides in the Presence of Macroscopic and Molecular Interfaces. *PLoS Comput. Biol.* **2015**, *11* (8), 1–24.
- (35) Sansom, M. S.; Kerr, I. D.; Smith, G. R.; Son, H. S. The Influenza A Virus M2 Channel: A Molecular Modeling and Simulation Study. *Virology* **1997**, *233* (1), 163–173.
- (36) Wu, Y.; Voth, G. A. A Computational Study of the Closed and Open States of the Influenza a M2 Proton Channel. *Biophys. J.* **2005**, *89* (4), 2402–2411.
- (37) DeGrado, W. F.; Lear, J. D. Induction of Peptide Conformation at Apolar Water Interfaces. 1. A Study with Model Peptides of Defined Hydrophobic Periodicity. *J. Am. Chem. Soc.* **1985**, *107* (10), 7684–7689.
- (38) Long, J. R.; Oyler, N.; Drobny, G. P.; Stayton, P. S. Assembly of Alpha-Helical Peptide Coatings on Hydrophobic Surfaces. *J. Am. Chem. Soc.* **2002**, *124* (22), 6297–6303.
- (39) Bower, P. V.; Louie, E. A.; Long, J. R.; Stayton, P. S.; Drobny, G. P. Solid-State NMR Structural Studies of Peptides Immobilized on Gold Nanoparticles. *Langmuir* **2005**, *21* (7), 3002–3007.
- (40) Mermut, O.; Phillips, D. C.; York, R. L.; McCrea, K. R.; Ward, R. S.; Somorjai, G. A. In Situ Adsorption Studies of a 14-Amino Acid Leucine-Lysine Peptide onto Hydrophobic Polystyrene and Hydrophilic Silica Surfaces Using Quartz Crystal Microbalance, Atomic Force Microscopy, and Sum Frequency Generation Vibrational Spectroscopy. *J. Am. Chem. Soc.* **2006**, *128* (11), 3598–3607.
- (41) Phillips, D. C.; York, R. L.; Mermut, O.; McCrea, K. R.; Ward, R. S.; Somorjai, G. A. Side Chain, Chain Length, and Sequence Effects on Amphiphilic Peptide Adsorption at Hydrophobic and Hydrophilic Surfaces Studied by Sum-Frequency Generation Vibrational Spectroscopy and Quartz Crystal Microbalance. *J. Phys. Chem. C* **2007**, *111* (1), 255–261.
- (42) Buchko, G. W.; Jain, A.; Reback, M. L.; Shaw, W. J. Structural Characterization of the Model Amphipathic Peptide Ac-LKKLLKLLKLLK-NH2 in Aqueous Solution and with 2,2,2-Trifluoroethanol and 1,1,1,3,3,3-Hexafluoroisopropanol. *Can. J. Chem.* **2013**, *91* (6), 406–413.
- (43) Davis, J. H. Deuterium Magnetic Resonance Study of the Gel and Liquid Crystalline Phases of Dipalmitoyl Phosphatidylcholine. *Biophys. J.* **1979**, *27* (3), 339–358.
- (44) Torchia, D. A. The Measurement of Proton-Enhanced Carbon-13 T1 Values by a Method Which Suppresses Artifacts. *J. Magn. Reson.* **1978**, *30* (3), 613–616.
- (45) Kuprov, I. Diagonalization-Free Implementation of Spin Relaxation Theory for Large Spin Systems. *J. Magn. Reson.* **2011**, *209* (1), 31–38.
- (46) Bouten, M. On the Rotation Operators in Quantum Mechanics. *Physica* **1969**, *42* (4), 572–580.
- (47) Mueller, L. J. Tensors and Rotations in NMR. *Concepts Magn. Reson. Part A* **2011**, *38* (5), 221–235.
- (48) Levitt, M. H.; Di Bari, L. Steady State in Magnetic Resonance Pulse Experiments. *Phys. Rev. Lett.* **1992**, *69* (21), 3124–3127.
- (49) Kubo, A.; McDowell, C. A. Spectral Spin Diffusion in Polycrystalline Solids under Magic-Angle Spinning. *J. Chem. Soc., Faraday Trans.* **1988**, *84* (11), 3713–3730.
- (50) Veshtort, M.; Griffin, R. G. Proton-Driven Spin Diffusion in Rotating Solids via Reversible and Irreversible Quantum

- Dynamics. *J. Chem. Phys.* **2011**, *135* (13), 134509.
- (51) Giraud, N.; Blackledge, M.; Böckmann, A.; Emsley, L. The Influence of Nitrogen-15 Proton-Driven Spin Diffusion on the Measurement of Nitrogen-15 Longitudinal Relaxation Times. *J. Magn. Reson.* **2007**, *184* (1), 51–61.
- (52) Torchia, D. A.; Szabo, A. Spin-Lattice Relaxation in Solids. *J. Magn. Reson.* **1982**, *49* (1), 107–121.
- (53) Mantsch, H. H.; Saitō, H.; Smith, I. C. P. Deuterium Magnetic Resonance, Applications in Chemistry, Physics and Biology. *Prog. Nucl. Magn. Reson. Spectrosc.* **1977**, *11* (4), 211–272.
- (54) Rinné, M.; Depireux, J. Nuclear Quadrupole Coupling Constant of Deuterium Bound to Carbon in Organic Molecules. In *Advances in Nuclear Quadrupole Resonance, Vol. 1*; Smith, J. A. S., Ed.; Heyden: London, 1972; pp 357–374.
- (55) Vold, R. R.; Vold, R. L. Deuterium Relaxation in Molecular Solids. *Adv. Magn. Opt. Reson.* **1991**, *16*, 85–171.
- (56) Vold, R. L.; Hoatson, G. L. Effects of Jump Dynamics on Solid State Nuclear Magnetic Resonance Line Shapes and Spin Relaxation Times. *J. Magn. Reson.* **2009**, *198* (1), 57–72.
- (57) Vugmeyster, L.; Ostrovsky, D.; Ford, J. J.; Burton, S. D.; Lipton, A. S.; Hoatson, G. L.; Vold, R. L. Probing the Dynamics of a Protein Hydrophobic Core by Deuteron Solid-State Nuclear Magnetic Resonance Spectroscopy. *J. Am. Chem. Soc.* **2009**, *131* (38), 13651–13658.
- (58) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. GROMACS: A Message-Passing Parallel Molecular Dynamics Implementation. *Comput. Phys. Commun.* **1995**, *91* (1–3), 43–56.
- (59) Lindahl, E.; Hess, B.; van der Spoel, D. GROMACS 3.0: A Package for Molecular Simulation and Trajectory Analysis. *J. Mol. Model.* **2001**, *7*, 306–317.
- (60) Neria, E.; Fischer, S.; Karplus, M. Simulation of Activation Free Energies in Molecular Systems. *J. Chem. Phys.* **1996**, *105* (5), 1902–1921.
- (61) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713.
- (62) Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. a; Case, D. a. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25* (9), 1157–1174.
- (63) 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* (1995), 8577–8593.
- (64) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, *18* (12), 1463–1472.
- (65) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; et al. Gaussian 09 Revision A.02. Gaussian, Inc.: Wallingford, CT 2009.
- (66) Chatfield, D. C.; Wong, S. E. Methyl Motional Parameters in Crystalline L-Alanine: Molecular Dynamics Simulation and NMR. *J. Phys. Chem. B* **2000**, *104* (47), 11342–11348.
- (67) Krushelnitsky, A.; Zinkevich, T.; Mukhametshina, N.; Tarasova, N.; Gogolev, Y.; Gnezdilov, O.; Fedotov, V.; Belton, P.; Reichert, D. 13C and 15N NMR Study of the Hydration Response of T4 Lysozyme and αB-Crystallin Internal Dynamics. *J Phys Chem B* **2009**, *113*, 10022–10034.
- (68) Ferreira, H. E.; Drobný, G. P. Solid State Deuterium NMR Study of LKα14 Peptide Aggregation in Biosilica. *Biointerphases* **2017**, *12* (2), 02D418.
- (69) Vugmeyster, L.; Ostrovsky, D.; Penland, K.; Hoatson, G. L.; Vold, R. L. Glassy Dynamics of Protein Methyl Groups Revealed by Deuteron NMR. *J. Phys. Chem. B* **2013**, *117* (4), 1051–1061.
- (70) Xue, Y.; Pavlova, M. S.; Ryabov, Y. E.; Reif, B.; Skrynnikov, N. R. Methyl Rotation Barriers in Proteins from 2H Relaxation Data. Implications for Protein Structure. *J. Am. Chem. Soc.* **2007**, *129* (21), 6827–6838.
- (71) Bryant, R. G. The Dynamics of Water-Protein Interactions. *Annu. Rev. Biophys. Biomol. Struct.* **1996**, *25* (1), 29–53.
- (72) Vugmeyster, L.; Ostrovsky, D.; Villafranca, T.; Sharp, J.; Xu, W.; Lipton, A. S.; Hoatson, G. L.; Vold, R. L. Dynamics of Hydrophobic Core Phenylalanine Residues Probed by Solid-State Deuteron NMR. *J. Phys. Chem. B* **2015**, *119* (47), 14892–14904.

TOC Graphic:
