Explicit Models of Motion to Understand Protein Side-Chain Dynamics

Nicolas Bolik-Coulon,[1,*,†] Olivier Languin-Cattoën,[2] Diego Carnevale,[1] Milan Zachrdla,[1,2] Damien Laage,[3] Fabio Sterpone,[2] Guillaume Stirnemann,[2,‡] and Fabien Ferrage[1,‖]

[1]*Laboratoire des Biomolécules, LBM, Département de chimie, École Normale Supérieure, PSL University, Sorbonne Université, CNRS, 24 rue Lhomond, 75005 Paris, France*

[2]*CNRS Laboratoire de Biochimie Théorique, Institut de Biologique Physico-Chimique, Université Paris Cité, PSL University, 13 rue Pierre et Marie Curie, 75005 Paris, France*

[3]*PASTEUR, Département de chimie, École Normale Supérieure, PSL University, Sorbonne Université, CNRS, 24 rue Lhomond, 75005 Paris, France*

(Received 13 April 2022; accepted 4 September 2022; published 10 November 2022)

Nuclear magnetic relaxation is widely used to probe protein dynamics. For decades, most analyses of relaxation in proteins have relied successfully on the model-free approach, forgoing mechanistic descriptions of motion. Model-free types of correlation functions cannot describe a large carbon-13 relaxation dataset in protein side chains. Here, we use molecular dynamics simulations to design explicit models of motion and solve Fokker-Planck diffusion equations. These models of motion provide better agreement with relaxation data, mechanistic insight, and a direct link to configuration entropy.

DOI: 10.1103/PhysRevLett.129.203001

Protein dynamics can be studied in the liquid state using a wide range of spectroscopic and scattering approaches [1–4]. For example, fluorescence anisotropy decay curves are direct measurements of the time-correlation function for the orientation of the fluorescence probe axis frame [5]. Nuclear magnetic resonance (NMR) relaxation rates probe the spectral density function, the Fourier transform of the time-correlation function for the orientation of the axis frames of spin interactions [6–9] in the fixed laboratory frame. These methods are sensitive to both the overall rotational diffusion of the molecule and internal motion of small, well-defined moieties.

The quantitative interpretation of relaxation rates in terms of motion requires parametrized models of correlation functions. A diversity of such models of motion were introduced in the second half of the 20th century: rotation on a cone [10], orientation jump [11], and diffusion in a cone [5]. Discriminating between these models is difficult, particularly when limited experimental datasets are available, which motivated the introduction of simplified correlation functions. For instance, in NMR spectroscopy, the model-free (MF) approach approximates the correlation function for internal motion $C_{\mathrm{int}}(t)$ to a single exponential decay term [12,13]:

$$C_{\mathrm{int}}(t) = \mathcal{S}^2 + (1-\mathcal{S}^2)e^{-t/\tau}, \qquad (1)$$

with $\mathcal{S}^2$ the generalized squared order parameter and τ an effective correlation time. The great simplification brought by the MF approach has led to hundreds of successful analyses of NMR relaxation data recorded on proteins [14] and a better understanding of internal protein motion.

A variety of NMR methods and new instruments make it possible to record extensive sets of relaxation rates and to probe the range of validity of the MF approach. We have recently reported site-specific isoleucine-δ1 carbon-13 relaxation measurements recorded on the protein ubiquitin over 2 orders of magnitude of magnetic fields [15,16]. Correlation functions derived from the MF approach were unable to describe both auto- and cross-correlated relaxation rates [16]. The MF approach was a necessary simplification four decades ago, but, today, the nature of motion is better known thanks to molecular dynamics (MD) simulations. Would explicit models of motion provide a more accurate and informative framework to analyze extensive relaxation datasets? Here, we use MD simulations to define key properties of motion (e.g., to identify which rotamer states are populated) and write Fokker-Planck diffusion equations that fulfill these properties. We employ the resulting correlation functions to analyze NMR relaxation data and to quantitatively describe the dynamics of protein side chains and the underlying molecular mechanisms.

Our approach is illustrated in the case of carbon-13 relaxation in isoleucine-δ1 specifically labeled $^{13}C^1H^2H_2$ methyl groups, where carbon-13 relaxation is driven by its dipole-dipole (DD) interactions with the proton and deuterons and its chemical shift anisotropy (CSA). We analyze a broad dataset consisting of carbon-13 longitudinal (R_1) and transverse (R_2) relaxation rates and carbon-proton DD cross-relaxation rates (σ^{NOE}) measured at four high magnetic fields, as well as 22 relaxometry carbon-13 R_1 rates measured over 2 orders of magnitude of magnetic field [15]. We also include CSA and DD cross-correlated

cross-relaxation (CCCR) rates (longitudinal η_z and transverse η_{xy}) measured at two high magnetic fields. We use for all models a Markov-chain Monte Carlo (MCMC) procedure [17] to analyze 38 relaxation rates per methyl group and extract the model parameters.

We previously used an MF-type of correlation function to describe the isoleucine $C_{\gamma 1}C_{\delta 1}$ bond motion, i.e., the extended model-free [18] correlation function. When all rates are analyzed together, experimental transverse relaxation rates R_2 cannot be reproduced for any of the six isoleucine residues with no sizable contribution of chemical exchange, especially at the highest magnetic fields (Figs. 2(a) and S1 [19]). These large discrepancies demonstrate that MF-type correlation functions cannot reproduce this set of relaxation rates.

Recent advances in MD simulations, with improved protein force fields, provide a wealth of information on the nature of motion [15,20,21]. We propose to use an MD trajectory to define appropriate explicit models of correlation functions to analyze experimental data, providing a mechanistic description of motion. Note that we are not (i) validating one method with the other, as can be done by calculating NMR parameters from MD simulations [20,22]; (ii) constraining an MD trajectory using NMR results [23,24]; or (iii) reweighing an MD trajectory with NMR constraints [20]. Rather, we use MD to identify the rotamers relevant for an explicit model that we use to analyze NMR relaxation. Our approach echoes a recent study of lipid bilayer dynamics [25], where MD results were refined using NMR data, but the detector approach used therein [25,26] yielded no mechanistic description of motion.

We performed a 1 μs MD simulation of ubiquitin with Gromacs [27–31], using the Amber ff99SB*-ILDN force field [32,33] modified with accurate energy barriers for methyl rotation [34] and the TIP4P-2005 water model [35]. The distributions of the dihedral angles χ_1 and χ_2 highlight the different possible motions of each isoleucine residue (Figs. 1 and S2 [19]). Accessible conformations of isoleucine side chains correspond to the nine possible (χ_1 and χ_2) rotamers, and their dynamics are well described by instantaneous jumps between rotamer states. Here, we neglect the local librations of $C_{\gamma 1}C_{\delta 1}$ bonds, as we assume a negligible contribution to relaxation [36], and use a model of infinitely fast jumps between discrete positions for $C_{\gamma 1}C_{\delta 1}$ bond dynamics [11].

The chemical environment for the Ile-$\delta 1$ methyl groups is rotamer dependent, which should lead to rotamer-specific chemical shift tensors [37]. Precise isoleucine side-chain conformations in water were obtained using density functional theory (DFT) calculations in Gaussian [38] (see Supplemental Material [19]). The CSA tensors calculated with the gauge-independent atomic orbital method [39,40] are rotamer dependent (Table S1 [19]). Within the Born-Oppenheimer approximation, any jump of the side chain

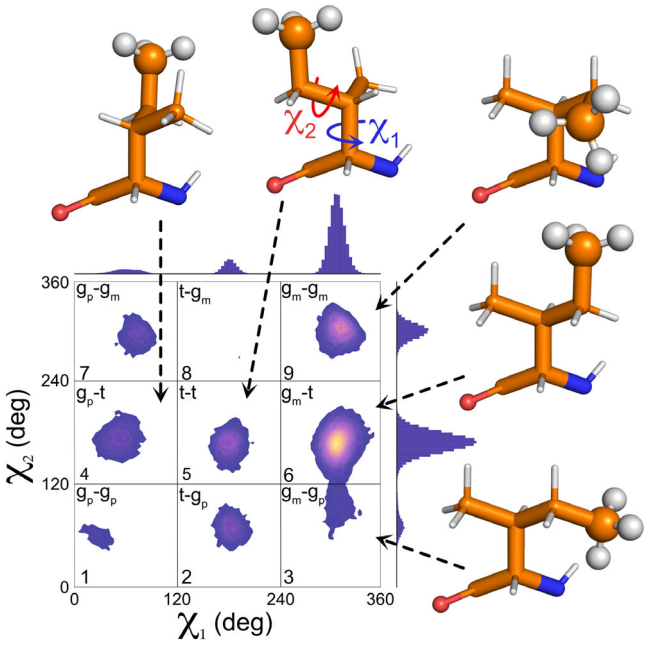

FIG. 1. Probability density distribution for χ_1 and χ_2 dihedral angles for Ile-36 obtained from a 1 μs MD trajectory. Nine rotamers can be defined, each numbered from 1 to 9 as shown. We also indicate the common states nomenclature with gauche+ (g_p), trans (t), and gauche− (g_m). Some conformers are shown as well with the $C_{\delta 1}H_3$ group presented with spheres.

from one rotamer to another would thus result in an instantaneous change of the CSA. We included this time-dependent interaction amplitude in our model to accurately account for the fluctuations of CSA tensors.

In the Bloch-Wangsness-Redfield relaxation theory, the correlation function between interactions i and j can include the amplitudes of the interactions, as denoted by the superscript (I):

$$C_{i,j}^{(I)}(t) = \langle \zeta_i(0)\zeta_j(t)\mathcal{D}_{q0}^{(2)*}(\Omega_{L,i},0)\mathcal{D}_{q0}^{(2)*}(\Omega_{L,j},t)\rangle, \quad (2)$$

where $\langle\cdots\rangle$ stands for an ensemble average, $\mathcal{D}_{q0}^{(2)*}(\Omega_{L,i},t)$ is a rank-2 Wigner matrix with the Euler angle set $\Omega_{L,i} = \{\alpha_{L,i},\beta_{L,i},\gamma_{L,i}\}$ for transformation from the laboratory to the interaction-i frame at time t, and $\zeta_i(t)$ is the strength of the interaction i at time t. Motion is modeled using Fokker-Plank diffusion equations, and the associated operators are diagonalized to write time-dependent bond-orientation conditional probabilities used to express the correlation function [36,41]. Intermediate frames are introduced to facilitate the description of individual motions, all assumed to be statistically independent unless otherwise stated (Fig. S3 [19]): the global tumbling (with the diffusion frame), the rotamer jump (with the jump and rotamer frames), and the methyl rotation (with the system frame). A diffusive motion for the methyl rotation is a relevant approximation, as discussed in Sec. I in Supplemental

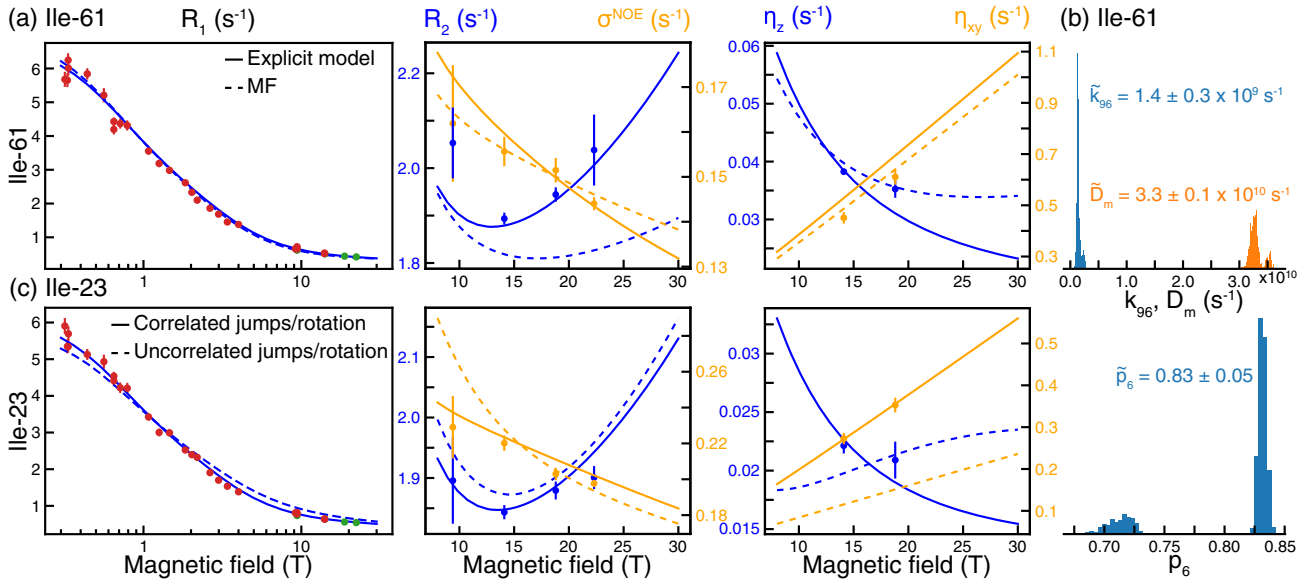

FIG. 2. Explicit models of motion for residues Ile-61 and Ile-23 in human ubiquitin were used to analyze longitudinal R_1 and transverse R_2 relaxation rates; dipolar cross-relaxation rates σ^{NOE}; and longitudinal η_z and transverse η_{xy} CCCR rates. (a) Measured and calculated relaxation rates following a MCMC procedure using the explicit model of motion (solid line) or MF-type correlation function (dashed line) for Ile-61. (b) Distributions of parameters of the explicit model of motion for Ile-61 from the MCMC analysis. MD simulations show that only rotamers 6 and 9 are populated, suggesting a model including jumps between rotamers 6 and 9. The free parameters of the model are the population of rotamer 6, p_6, the exchange rate from rotamer 6 to rotamer 9, k_{96}, and the diffusion constant for rotation on a cone of the methyl group, D_m. The median and standard deviation are indicated in each panel. (c) Measured and calculated relaxation rates following a MCMC procedure using the explicit model of motion with (solid line) or without (dashed line) correlation of jumps and methyl rotation for Ile-23. The high-field R_1 and the corrected relaxometry rates are shown in green and red, respectively.

Material [19]. Note that both internal and global motions are anisotropic *a priori*.

The complexity of the correlation functions rapidly increases with the number of rotamers N, as the sums contain N^3 terms, and $N(N-1)/2$ jump rates need to be determined to fully characterize the exchange matrix. Information from MD simulations is essential to define suitable explicit models with a minimal number of adjustable parameters. First, we included only rotamers with average fractional populations higher than 1% in the MD trajectory (Table S2 [19]). Second, jump rates for transitions which were not observed in the MD trajectory were fixed to 0 in the exchange matrix. The adjustable parameters of the model are populations of rotamers selected by MD, exchange rates, and diffusion constants for methyl rotation. For additional information on the data analysis, see Supplemental Material [19], which contains Refs. [42–49].

Explicit models of motion are compatible with experimental relaxation rates (Figs. 2(a) and S4 [19]) with better agreement than the MF analysis (Table S3 [19]), in particular, for R_2 and CCCR rates. Some deviations can be observed for the σ^{NOE}, which might be due to the neglected contribution of fast $C_{\gamma 1}C_{\delta 1}$ bond librations [36]. In addition, the explicit model gives a mechanistic picture of the motion: we estimate equilibrium populations (Figs. 2(b) and S5 and Table S4 [19]) which are accessible to only a few advanced experimental methods [37,50,51] and, uniquely, the kinetics of exchange.

Most distributions of parameters obtained from the MCMC analysis are well defined and allow a precise estimate of the parameters of the model (Eqs. (S7)–(S9) and Figs. 2(b), S5, and S6 [19]). However, as the number of states increases, the distributions become broader and the resulting jump matrices can be ill defined (see Supplemental Material [19]). Further improvement could consist in analyzing relaxation rates simultaneously with other data defining conformational ensembles, such as chemical shifts or scalar-coupling constants [37,52].

Explicit models of motion constitute an efficient framework to combine the imperfect information from MD simulations and experiments. Despite recent force-field improvements [34], MD data alone cannot reproduce relaxation rates, whether these are calculated directly from the MD correlation functions (Figs. S7 and S8 [19]) or from the analysis of MD with the explicit models of motion [53] (Fig. S9 and Table S3 [19]). Beyond force-field deficiencies, this is likely due to the limited sampling on the microsecond timescale. If the identification of the relevant rotameric states is probably reliable, the extraction of accurate transition rates would require many more transitions during the MD trajectory for good statistics.

Nevertheless, the MD input is critical to complement the information from experimental relaxation rates. We analyzed relaxation rates for isoleucines 30 and 61 for all combinations of two rotameric states. Several sets of two rotamers lead to comparable or slightly better χ^2 values (Fig. S10 [19]) than the combination of rotamers obtained from the MD simulation. Such an exhaustive analysis is challenging for a two-state model even assuming prior knowledge on the number of states. For larger numbers of states, such an approach would be unrealistic. Thus, relaxation rates alone are insufficient to determine the populations and kinetics of exchange among rotamer states. Our approach combines the most robust information from an MD simulation: the network of accessible rotamer states, with the information from NMR relaxation, which is sensitive to the populations of rotamer states and the kinetics of exchange. This combined analysis of MD simulation and NMR is necessary to obtain a quantitative mechanistic description of the dynamics.

Explicit models of motion can accommodate motion of increasing complexity including, for instance, different diffusion rates for methyl rotation in rotamer states. In the case of Ile-23, the explicit model of motion with uniform methyl rotation (Eq. (S7) [19]) is unable to reproduce the experimental data well (Fig. S4 [19]). We noticed that, for the major rotamer conformation of this residue, the $\delta 1$-methyl group is in close proximity of the H_α (rotamer 9 in Fig. 1). Such steric hindrance is not present in the other rotamer states. We built a model taking into account methyl rotation specific to rotamer states, with identical diffusion constants for methyl rotation in rotamers 3 and 6 and a specific diffusion constant for rotamer 9, which led to a clear improvement in the agreement of the model with the measured relaxation rates (Figs. 2(c) and S11 [19]). The resulting diffusion constants for methyl rotation $D_{m,3} = D_{m,6} = 5.5 \pm 1.9 \times 10^{10}$ s^{-1} and $D_{m,9} = 1.1 \pm 0.4 \times 10^{10}$ s^{-1} support the presence of correlated motion in the Ile-23 side chain, with methyl rotation 5 times slower in the major rotamer, as expected from the steric hindrance of the H_α.

Generalized order parameters quantify the width of the conformational space at equilibrium and thus can be linked to conformational entropy [54,55], which is defined by considering each accessible conformation as a microstate. The link between the generalized order parameter obtained in the MF approach and entropy is not direct and requires one either to reintroduce models of motion or rely on residue-specific estimates from MD simulations [54,56,57]. All these models have to use a single parameter to describe the amplitude of motion. One may ask whether a single parameter suffices to describe both order parameters and conformational entropy. In stark contrast to existing approaches, our explicit models of motion provide the distribution of rotamer states that can be directly used to estimate the configuration entropy S_c associated to this distribution of rotamer states:

$$S_c/k_B = -\sum_{\alpha=1}^{N} p_\alpha \ln p_\alpha, \tag{3}$$

where k_B is the Boltzmann constant and p_α is the equilibrium fractional population of rotamer α. The full conformational entropy includes both the configuration entropy S_c and the differential entropy [58] associated with substates within each rotamer and quantified from the amplitude of librations. The simulation analysis shows that order parameters predominantly reflect the equilibrium distribution of rotamers, suggesting a close link between order parameters and configuration entropy (Fig. S12(b) [19]). Importantly, we find that the differential entropy in a given rotameric state is rather constant, between rotamers and between residues (Fig. S12(a) [19]). Thus, variations in conformational entropy mostly arise from configuration entropy. The rotamer distributions enable one to evaluate the relationship between order parameters and configuration entropy. We generated 10 000 random distributions of rotamer populations, using optimized geometries for side chains obtained from DFT calculations (see above and Supplemental Material [19]) and computed both configuration entropy (Eq. (3)) and order parameters according to [36,52]

$$\mathcal{S}_J^2 = \sum_{\alpha,\beta=1}^{N} p_\alpha p_\beta \mathcal{P}_2(\cos\theta_{\alpha,\beta}), \tag{4}$$

where $\theta_{\alpha,\beta}$ is the angle between rotamers α and β. The random distributions were generated by first randomly choosing the number of populated rotamers and then the fractional population for each of them.

Interestingly, for a given value of order parameter, we find a range of possible configuration entropies as wide as ca. k_B (Fig. 3): Configuration entropy cannot be derived unambiguously from an order parameter. Similar results are obtained for leucine and valine side chains (Fig. S12(c) [19]). Our combined NMR and MD analysis gives distributions of rotamer populations that can be used to derive both configuration entropies and order parameters. Importantly, our experimental results confirm the calculations from random distributions (Fig. 3 and Table S5 [19]): Large variations of configuration entropy can be observed for small changes of order parameters (see Ile-13 and Ile-61), and similar configuration entropies can be obtained for residues with drastically different order parameters (compare Ile-13 and Ile-44). Considering additional limitations arising from limited conformational sampling on the timescale of overall rotational diffusion [59], our investigation confirms that the estimation of conformational entropy from NMR data alone is difficult.

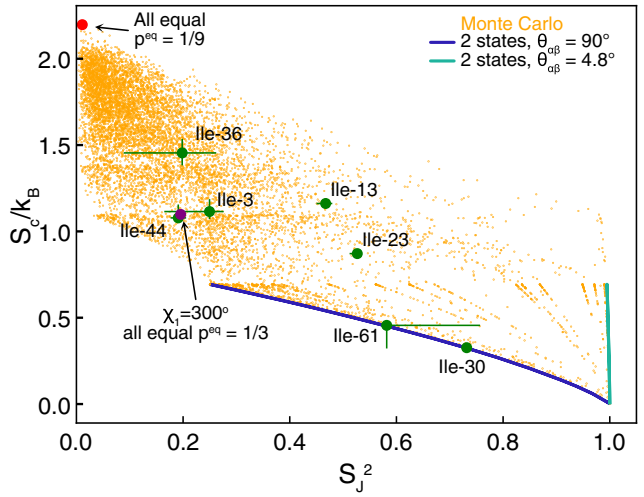

FIG. 3. Configuration entropy (S_c, Eq. (3)) and squared order parameter for rotamer jump (S_J^2, Eq. (4))calculated for 10 000 random rotamer distributions of an isoleucine side chain with geometry derived from DFT calculations. The generalized order parameters and configuration entropies obtained from the combined analysis of NMR relaxation and MD simulations for the seven isoleucines of ubiquitin are shown in green, with errors corresponding to the 16th and 84th percentile of the distributions following calculations of the order parameters and configuration entropy over the whole MCMC trajectories. The blue and green parametric curves describe a two-rotamer jump model with the lowest (blue) and highest (cyan) value for $\mathcal{P}_2(\cos\theta_{\alpha,\beta})$. The point where all rotamers have equal population (i.e., highest entropy) is shown in red, and the point where three rotamers have equal population and $\chi_1 = 300°$ (rotamers 3, 6, and 9) is shown in purple.

The combined analysis of NMR relaxation and MD simulations with a rotamer jump model uncovers the physical origin of conformational entropy and solves limitations that arise when a single-parameter model is used to link order parameters and conformational entropy. Further integration may open a path toward the identification of correlated motion in systems with a larger number of experimental side-chain probes [60,61].

Here, we have shown that molecular dynamics simulations could be used to build explicit models of motion in order to obtain a mechanistic description of motion for protein side chains from the analysis of NMR relaxation. We defined models of motion from MD simulations that include jumps between rotamer states and methyl rotation on a cone. We determined from NMR relaxation the populations of rotamers and their exchange rates. Such explicit models allow fine-tuning of motional properties, such as slower methyl rotation in one rotamer state, which may be understood by steric interactions. The description of motion as exchange between substates gives access to the molecular origin of conformational entropy, the variations of which are dominated by the configuration entropy associated to the distribution of rotamer states. Our approach, based on a combination of DFT calculations, MD simulations, and extensive NMR measurements, can be adapted for most probes of site-specific dynamics in macromolecules. We anticipate that further integration of NMR and MD simulations in combination with explicit models of motion will improve the mechanistic description of protein motion and lead to a better understanding of the physics and chemistry that sustain protein function.

The project was supported by the grant ALLODYN from the PSL Chemistry call of PSL University and the FET-Open program of the European Union, Grant Agreement No. 899683 (HIRES-MULTIDYN). This work received support from the "Initiative d'Excellence" program (Grant "DYNAMO", No. ANR-11-LABX-0011-01, G. S. and F. S.), and simulations benefited from access to the HPC resources of TGCC under the allocation A0070811005 made by GENCI (Grand Equipement National de Calcul Intensif).

*nicolas.bolikcoulon@utoronto.ca
†Present address: Department of Molecular Genetics, University of Toronto, Toronto, ON M5S 1A8, Canada.
‡stirnemann@ibpc.fr
‖fabien.ferrage@ens.psl.eu

[1] R. Roy, S. Hohng, and T. Ha, Nat. Methods **5**, 507 (2008).
[2] C. Charlier, S. F. Cousin, and F. Ferrage, Chem. Soc. Rev. **45**, 2410 (2016).
[3] M. Grimaldo, F. Roosen-Runge, F. Zhang, F. Schreiber, and T. Seybel, Q. Rev. Biophys. **52**, E7 (2019).
[4] G. W. Tumbic, M. Y. Hossan, and M. C. Thielges, Annu. Rev. Anal. Chem. **14**, 299 (2021).
[5] K. Kinosita, S. Kawato, and A. Ikegami, Biophys. J. **20**, 289 (1977).
[6] R. Wangsness and F. Bloch, Phys. Rev. **89**, 728 (1953).
[7] A. G. Redfield, IBM J. Res. Dev. **1**, 19 (1967).
[8] A. G. Palmer III, Chem. Rev. **104**, 3623 (2004).
[9] M. P. Nicholas, E. Erylmaz, F. Ferrage, D. Cowburn, and R. Ghose, Prog. Nucl. Magn. Reson. Spectrosc. **57**, 111 (2010).
[10] D. Wallach, J. Chem. Phys. **47**, 5258 (1967).
[11] R. J. Wittebort and A. Szabo, J. Chem. Phys. **69**, 1722 (1978).
[12] G. Lipari and A. Szabo, J. Am. Chem. Soc. **104**, 4546 (1982).
[13] B. Halle, J. Chem. Phys. **131**, 224507 (2009).
[14] V. A. Jarymowycz and M. J. Stone, Chem. Rev. **106**, 1624 (2006).
[15] S. F. Cousin, P. Kadeřávek, N. Bolik-Coulon, Y. Gu, C. Charlier, L. Carlier, L. Bruschweiler-Li, T. Marquardsen, J.-M. Tyburn, R. Brüschweiler, and F. Ferrage, J. Am. Chem. Soc. **140**, 13456 (2018).
[16] N. Bolik-Coulon, P. Kadeřávek, P. Pelupessy, J.-N. Dumez, F. Ferrage, and S. F. Cousin, J. Magn. Reson. **313**, 106718 (2020).
[17] D. Foreman-Mackey, D. W. Hogg, D. L. J., and Goodman, Publ. Astron. Soc. Pac. **125**, 306 (2013).
[18] G. M. Clore, A. Szabo, A. Bax, L. E. Kay, P. C. Driscoll, and A. M. Gronenborn, J. Am. Chem. Soc. **112**, 4989 (1990).

[19] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.129.203001 for details on the derivations of correlation functions, expressions of relaxation rates, calculations of CSA tensors, MD simulations, and supplementary figures.

[20] F. Kümmerer, S. Orioli, D. Harding-Larsen, F. Hoffmann, Y. Gavrilov, K. Teilum, and K. Lindorff-Larsen, J. Chem. Theory Comput. **17**, 5262 (2021).

[21] A. A. Smith, J. Magn. Reson. Open **10–11**, 100045 (2022).

[22] P. Maragakis, K. Lindorff-Larsen, M. P. Eastwood, R. O. Dror, J. L. Klepeis, I. T. Arkin, M. Ø. Jensen, H. Xu, N. Trbovic, R. A. Friesner, A. G. Palmer III, and D. E. Shaw, J. Phys. Chem. B **112**, 6155 (2008).

[23] K. Lindorff-Larsen, R. B. Best, M. A. DePristo, C. M. Dobson, and M. Vendruscolo, Nature (London) **433**, 128 (2005).

[24] C. Camilloni, P. Robustelli, A. De Simone, A. Cavalli, and M. Vendruscolo, J. Am. Chem. Soc. **134**, 3968 (2012).

[25] A. A. Smith, A. Vogel, O. Engberg, P. W. Hildebrand, and D. Huster, Nat. Commun. **13**, 108 (2022).

[26] A. A. Smith, M. Ernst, and B. H. Meier, J. Chem. Phys. **148**, 045104 (2018).

[27] H. J. C. Berendsen, D. van Der Spoel, and R. V. Drunen, Comput. Phys. Commun. **91**, 43 (1995).

[28] E. Lindahl, B. Hess, and D. van Der Spoel, J. Mol. Model. **7**, 306 (2001).

[29] D. van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. C. Berendsen, J. Comput. Chem. **26**, 1701 (2005).

[30] B. Hess, C. Kutzner, D. van Der Spoel, and E. Lindahl, J. Chem. Theory Comput. **4**, 435 (2008).

[31] S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van Der Spoel, B. Hess, and E. Lindahl, Bioinformatics **29**, 845 (2013).

[32] R. B. Best and G. Hummer, J. Phys. Chem. B **113**, 9004 (2009).

[33] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror, and D. E. Shaw, Proteins **78**, 1950 (2010).

[34] F. Hoffmann, F. A. A. Mulder, and L. V. Schäfer, J. Phys. Chem. B **122**, 5038 (2018).

[35] J. L. F. Abascal and C. Vega, J. Chem. Phys. **123**, 234505 (2005).

[36] N. Bolik-Coulon and F. Ferrage, J. Chem. Phys. **157**, 125102 (2022).

[37] L. Siemons, B. Uluca-Yazgi, R. B. Pritchard, S. McCarthy, H. Heise, and D. F. Hansen, Chem. Commun. **55**, 14107 (2019).

[38] M. J. Frisch *et al.*, *Gaussian 09 Revision A.01* (Gaussian Inc., Wallingford CT, 2009).

[39] D. Zeroka and H. F. Hameka, J. Chem. Phys. **45**, 300 (1966).

[40] R. Ditchfield, J. Chem. Phys. **56**, 5688 (1972).

[41] P. Luginbühl and K. Wüthrich, Prog. Nucl. Magn. Reson. Spectrosc. **40**, 199 (2002).

[42] D. Fushman, R. Xu, and D. Cowburn, Biochemistry **38**, 10225 (1999).

[43] C. Charlier, S. N. Khan, T. Marquardsen, P. Pelupessy, V. Reiss, D. Sakellariou, G. Bodenhausen, F. Engelke, and F. Ferrage, J. Am. Chem. Soc. **135**, 18665 (2013).

[44] R. Zwanzig, Proc. Natl. Acad. Sci. U.S.A. **85**, 2029 (1988).

[45] C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).

[46] A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).

[47] R. Ditchfield, W. J. Hehre, and J. A. Pople, J. Chem. Phys. **54**, 724 (1971).

[48] J. Tomasi, B. Mennucci, and R. Cammi, Chem. Rev. **105**, 2999 (2005).

[49] P. Virtanen *et al.*, Nat. Methods **17**, 261 (2020).

[50] D. F. Hansen, P. Neudecker, P. Vallurupalli, F. A. A. Mulder, and L. E. Kay, J. Am. Chem. Soc. **132**, 42 (2010).

[51] D. F. Hansen and L. E. Kay, J. Am. Chem. Soc. **133**, 8272 (2011).

[52] J. J. Chou, D. A. Case, and A. Bax, J. Am. Chem. Soc. **125**, 8959 (2003).

[53] K. A. Beauchamp, G. R. Bowman, T. J. Lane, L. Maibaum, I. S. Haque, and V. S. Pande, J. Chem. Theory Comput. **7**, 3412 (2011).

[54] M. Akke, R. Brüschweiler, and A. G. Palmer III, J. Am. Chem. Soc. **115**, 9832 (1993).

[55] K. K. Frederick, M. S. Marlow, K. G. Valentine, and A. J. Wand, Nature (London) **448**, 325 (2007).

[56] D. Yang and L. E. Kay, J. Mol. Biol. **263**, 369 (1996).

[57] D.-W. Li and R. Brüschweiler, J. Am. Chem. Soc. **131**, 7226 (2009).

[58] H. Singh, N. Misra, V. Hnizdo, A. Fedorowicz, and E. Demchuk, Am. J. Math. Manag. Sci. **23**, 301 (2003).

[59] F. Hoffmann, F. A. A. Mulder, and L. V. Schäfer, J. Phys. Chem. B **126**, 54 (2022).

[60] V. Tugarinov, V. Kanelis, and L. E. Kay, Nat. Protoc. **1**, 749 (2006).

[61] P. Gans, O. Hamelin, R. Sounier, I. Ayala, M. A. Durá, C. D. Amero, M. Noirclerc-Savoye, B. Franzetti, M. J. Plevin, and J. Boisbouvier, Angew. Chem., Int. Ed. **49**, 1958 (2010).