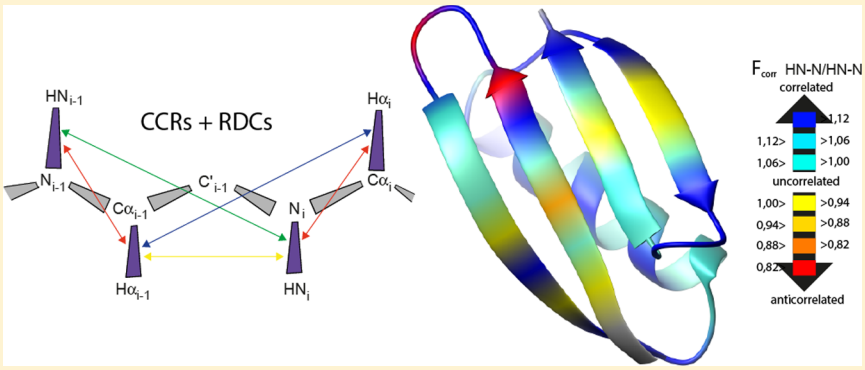
JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

Direct Investigation of Slow Correlated Dynamics in Proteins via Dipolar Interactions

R. Bryn Fenwick,*[,†,‡] Charles D. Schwieters,[§] and Beat Vögeli*[,∥]

[†]Institute for Research in Biomedicine (IRB Barcelona), Parc Científic de Barcelona, C/Baldiri Reixac 10, 08028 Barcelona, Spain
[‡]The Scripps Research Institute (TSRI), 10550 North Torrey Pines Road, La Jolla, California 92037, United States
[§]Division of Computational Bioscience, Center for Information Technology, National Institutes of Health, Building 12A, Bethesda, Maryland 20892-5624, United States
[∥]Laboratory of Physical Chemistry, Swiss Federal Institute of Technology, ETH-Hönggerberg, Vladimir-Prelog-Weg 2, CH-8093 Zürich, Switzerland

Supporting Information



ABSTRACT: The synchronization of native state motions as they transition between microstates influences catalysis kinetics, mediates allosteric interactions, and reduces the conformational entropy of proteins. However, it has proven difficult to describe native microstates because they are usually minimally frustrated and may interconvert on the micro- to millisecond time scale. Direct observation of concerted equilibrium fluctuations would therefore be an important tool for describing protein native states. Here we propose a strategy that relates NMR cross-correlated relaxation (CCR) rates between dipolar interactions to residual dipolar couplings (RDCs) of individual consecutive H^N–N and H^α–C^α bonds, which act as a proxy for the peptide planes and the side chains, respectively. Using Xplor-NIH ensemble structure calculations restrained with the RDC and CCR data, we observe collective motions on time scales slower than nanoseconds in the backbone for GB3. To directly access the correlations from CCR, we develop a structure-free data analysis. The resulting dynamic correlation map is consistent with the ensemble-restrained simulations and reveals a complex network. In general, we find that the bond motions are on average slightly correlated and that the local environment dominates many observations. Despite this, some patterns are typical over entire secondary structure elements. In the β-sheet, nearly all bonds are weakly correlated, and there is an approximately binary alternation in correlation intensity corresponding to the solvent exposure/shielding alternation of the side chains. For α-helices, there is also a weak correlation in the H^N–N bonds. The degree of correlation involving H^α–C^α bonds is directly affected by side-chain fluctuations, whereas loops show complex and nonuniform behavior.

INTRODUCTION

While protein folding is governed by a funneled energy landscape, native state motions constitute an ensemble of interconverting substates at the bottom of the funnel.[1−3] These inherent motions are part of the evolutionary selection criteria and therefore play a key role in protein function. The synchronization of these motions influences the kinetics of catalysis and is a mediator of allosteric interactions. It also impacts the thermodynamic stability of the protein through reduction of its conformational entropy.[4−7] While energy landscape theories have been successfully applied to understand protein folding, the native states are more difficult to model and are therefore less well understood.[2] Because they are usually minimally frustrated, it has proven difficult to adapt energy landscape theories to the description of native states and molecular dynamics simulations are usually limited to time scales shorter than many biologically relevant conformational changes.[8] Thus, experimental observation of equilibrium fluctuations in proteins are very much needed in order to further our understanding of protein dynamics and its role in stability and function.[9]

Received: February 8, 2016
Published: June 22, 2016

DOI: 10.1021/jacs.6b01447
J. Am. Chem. Soc. 2016, 138, 8412−8421

Nuclear magnetic resonance (NMR) is one of the few techniques that can provide such information because each atom may serve as a reporter on dynamics.[10−12] State-of-the-art NMR methods provide very powerful tools for assessing local bond motions, typically those of H^N−N bonds and methyl group symmetry axes.[13−15] However, these motions alone are not sufficient to understand the whole of protein dynamics and conformational entropy. In addition to measuring the magnitude and direction of individual atomic motions, the degree to which they are synchronized to each other must also be quantified. Novel NMR approaches have recently been proposed to achieve this goal. The first such method involves the analysis of the covariation of motional NMR parameters among a set of mutants.[16] Much attention has been dedicated to another approach, the calculation of ensembles of structures whose back-predicted averaged NMR parameters are in better agreement with experiments than those from single conformations.[17−23] Experimental input parameters are mostly nuclear Overhauser enhancement rates (NOEs), residual dipolar couplings (RDCs), and scalar couplings, and recently we have proposed the use of exact NOEs (eNOEs).[24] If different conformational states can be distinguished within the ensemble, then it may indicate modes of collective motions; however, the deconvolution of the motional time scales from these ensembles is difficult. The biological insights that such ensembles offer have recently been reviewed.[25] Both approaches (covariance and ensemble), however, reveal concerted motions in an indirect manner. Despite this, these methods have been used to great success to measure the degrees of correlated motions such as the crankshaft motions of the peptide plane, which occur on a time scale faster than the molecular tumbling time (sub-τ_c),[26] and correlated motions on time scales slower than τ_c (supra-τ_c) such as the β-lever motion.[21,27]

One direct observation of correlation between two events is offered by cross-correlated chemical shift modulation (CSM).[28,29] However, it is difficult to translate the modulation into a physical picture because it ultimately depends on the electron environment of the spins. A second direct observation can be made using cross-correlated dipolar relaxation rates (dipolar CCR rates) that depend on the relative orientation of two dipole−dipole interaction axes, although a careful analysis must be undertaken to bracket errors arising from approximations.[30,31] These are related to nucleus positions in a straightforward manner within the molecular frame. There is no fundamental restriction as to how far they may be apart as long as one spin of each interaction is involved in the same coherence. Importantly, CCR rates are sensitive to motion on all time scales.[32−34] Thus, multiple-quantum coherences are attractive for the study of the interplay between spatially separated bonds,[35] and numerous CCR experiments have been designed to determine the angles between H^N−N, H^α−C^α, H^N−C′, N−C′, and/or C′−C^α bonds.[32,35−42]

Internal motion modifies the CCR rate expected for a rigid molecule and can be quantified by a CCR order parameter S^2_{CCR}, which is obtained from the ratio of the experimental and the predicted rigid-molecule rate. The order parameters between consecutive H^N−N bonds in globular proteins have been shown to be ca. 0.75 on average and smaller than those obtained in autorelaxation rates.[32,43] Averaged heuristic order parameters estimated from consecutive H^α−C^α and H^N−N/H^α−C^α CCR rates cover a large range of 0.75−1.[36−40,42,44] Recently, dipolar CCR rates have also been used as a selection criterion for realistic NOE- and RDC-restrained molecular dynamics ensembles of ubiquitin.[27] Previous studies have used the CCR rates to identify slow correlated motions in a model-free manner because it has been theoretically shown that only correlation of slow bond motions (supra-τ_c) has a significant influence on the measured CCR rates for the pairs of bonds involved.[33,34,45] In one such case, a slow-motion (nano- to millisecond) CCR order parameter $S^2_{CCR,slow}$ was defined as $S^2_{CCR}/(S_{i,fast}S_{i+1,fast})$ and extracted by using relaxation H^N−N order parameters for S_{fast}.[43] The value of $S^2_{CCR,slow}$ can be used to detect the presence of slow motion; however, the degree of correlation is not encoded. An improved interpretation of the CCR order parameter requires both order parameters of the individual bond vectors that are sensitive to the supra-τ_c time-scale and in many cases a fully anisotropic description of molecluar tumbling. It is clear that insufficient experimental and structural accuracy would prevent reliable interpretation of the CCR order parameters.

Here, strategies to overcome all these caveats are employed to reliably extract from CCR rates a dense correlation network of H^N−N/H^N−N, H^α−C^α/H^α−C^α, and H^N−N/H^α−C^α bonds in the third immunoglobulin binding domain of protein G, GB3 (Figure 1).[46] To verify the accuracy of the measured CCR rate,

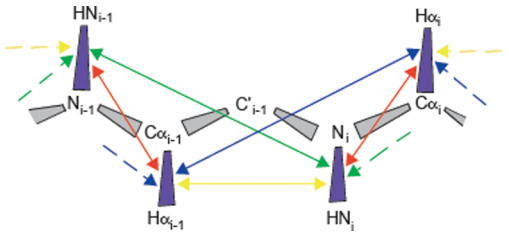


Figure 1. Cartoon representation of the CCR network and the RDC order parameters used in this study. Protein backbone bonds are shown in perspective rectangles. Purple rectangles indicate bond vectors whose RDC order parameters have previously been determined from multiple alignment conditions. Arrows connect bond vectors of the dipolar interactions between which CCRs are measured.

the following four points are crucial. (1) All CCR rates were measured with at least two methods that are as different as possible in order to obtain insight into systematic errors in addition to random errors inherent to each method. (2) Non-negligible corrections to the CCRs have been calculated from a full relaxation-matrix analysis.[47] (3) The orientations of the H^N−N and H^α−C^α bond vectors that have been used were previously determined to an accuracy of ca. 3 degrees from RDCs measured under multiple alignment conditions.[48,49] (4) CCR rates were back-calculated from structures using anisotropic overall tumbling.[34,50]

To a good approximation, there are two major angular contributions to the general-order parameter S^2. These have been defined by Bruschweiler as short-range collective motion, which is associated with adjacent dihedral angles, and long-range collective motions, which lead to angular fluctuations of largely rigid protein segments.[51] Bending and twisting of α-helicies and β-sheets, as well as more complex motions of loops and turns can produce both short- and long-range collective motions.

Here we report the combined analysis of RDCs and CCRs for GB3 from both ensemble structure calculations and a novel structure-free analysis of the same data. Because both CCRs

DOI: 10.1021/jacs.6b01447
J. Am. Chem. Soc. 2016, 138, 8412−8421

and RDCs are also sensitive to motion on the nano- to millisecond time scale, it is possible to characterize the degree and range of correlated motion exclusively on the supra-τ_c time scale. The degree and type of correlations on this time scale are not well charaterized in the literature; however, they are predicted to be distinct from those such as the peptide plane crankshaft motion, which is observed on the pico- to nanosecond time scale. By using the H^N–N bond vector as a proxy for the peptide plane and the H^α–C^α bond vector as a proxy for the side chain motion, we are able to probe the correlated motions of neighboring peptide planes and side chains as well as their interactions throughout the protein. Using the structure-free method with the two sets of previously determined RDC order parameters of the individual bonds as previously proposed,[52,53] our analysis of GB3 CCR data produces a highly complex correlation map for slow collective motions whose interpretation invokes not only secondary-structure-dependent geometry but also side chain dynamics.

THEORY AND SIMULATIONS

The cross-correlated relaxation rate between two dipolar interactions A and B ($\Gamma_{A/B}$) comprised of two spin pairs (A1, A2) and (B1, B2) is

$$\Gamma_{A/B} = \left(\frac{\mu_0}{4\pi}\right)^2 \frac{\gamma_{A1}\gamma_{A2}\gamma_{B1}\gamma_{B2}\hbar^2}{10\pi^2} \frac{1}{r_A^3 r_B^3} J_{A/B} \tag{1}$$

γ_X is the gyromagnetic ratio of nucleus X, r_X is the effective distance between nuclei X1 and X2,[54] μ_0 is the permeability of free space, and h is Planck's constant. The spectral density function $J_{A/B}$ depends on the orientation and dynamics of the vectors describing A and B. If the molecule undergoes anisotropic tumbling, then five effective overall tumbling times τ_k must be considered.[34,52]

It has been shown that $J_{A/B}$ is virtually independent of the time scale of internal motions as long as it is 1 order of magnitude faster or slower than τ_c (3.5–4 ns for GB3) and only the type of summation over the angular terms is altered when going from fast to slow motion.[34,45] When this condition is met, $J_{A/B}$ can be factorized into components G_k that depend on the orientation of only A, Ω_A, or only B, Ω_B:

$$J_{A/B} = \sum_{k=-2}^{2} \tau_k \langle G_k(\Omega_B)\, G_k(\Omega_A)\rangle \tag{2}$$

The averaging is carried out over an ensemble that samples all conformations present over a time longer than τ_c. If correlated slow motion between A and B is absent (but possibly all other types present: fast correlated, fast or slow uncorrelated), then the averaging can also be factorized:

$$J_{A/B} = \sum_{k=-2}^{2} \tau_k \langle G_k(\Omega_B)\rangle \langle G_k(\Omega_A)\rangle \tag{3}$$

This is possible because the fast-scale correlation is lost after a short time (much less than the overall molecular tumbling time) and the integral over the correlation function is nearly identical to one without any correlation at all.[45]

The correlation functions that do not decay over τ_c (slow motions), however, require averaging of their integrals over many molecules. Therefore, such an averaging causes a dependence of the observed CCR rate on the extent of correlation of slow motion. This opens a way to detect correlated slow motion if the CCR rates and the individual G_k factors (averaged over the slow time scale) can be assessed. This is indeed possible with the measurement of the individual RDC order parameters of A and B. Then eq 3 may be written as

$$J_{A/B} = S_A^{RDC} S_B^{RDC} J_{A/B}^{rigid} \tag{4}$$

if the A and B motions are sufficiently symmetric. $J_{A/B}^{rigid}$ is the spectral density function of a hypothetically rigid molecule. Deviation from eq 4 is then indicative of slow correlated motion between A and B. We introduce a correlation factor F_{corr} defined by the ratio of the experimental $\Gamma_{A/B}^{exp}$ and predicted CCR rates $\Gamma_{A/B}^{rigid}$ divided by the RDC order parameters. For all relevant cases, the following relationship is valid:

$$F_{corr,A/B} \equiv \frac{\Gamma_{A/B}^{exp}}{S_A^{RDC} S_B^{RDC} \Gamma_{A/B}^{rigid}} = \frac{J_{A/B}}{S_A^{RDC} S_B^{RDC} J_{A/B}^{rigid}} \begin{cases} = 1 \rightarrow \text{noncorrelated} \\ > 1 \rightarrow \text{syn-correlated} \\ < 1 \rightarrow \text{anti-correlated} \end{cases} \tag{5}$$

To restrain the CCRs during ensemble simulations, we have implemented an ensemble CCR restraint for isotropically tumbling molecules in Xplor-NIH.[55] In practice, the following sum of CCR rates is measured

$$\Gamma_{A/B}^{iso} = \Gamma_{A1,A2/B1,B2}^{iso} + \Gamma_{A1,B2/B1,A2}^{iso} \tag{6}$$

where the rate $\Gamma_{A1,A2/B1,B2}^{iso}$ accounts for the dipole interaction between A1–A2 and B1–B2 and depends on the ensemble-averaged angle between the two vectors A and B and their respective bond lengths as

$$\Gamma_{A1,A2/B1,B2}^{iso} = \left(\frac{\mu_0}{4\pi}\right)^2 \frac{\gamma_{A1}\gamma_{A2}\gamma_{B1}\gamma_{B2}\hbar^2}{10\pi^2} \frac{2\tau_c}{5} \left\langle \frac{3\cos^2\theta_{A1,A2/B1,B2} - 1}{r_{A1,A2}^3 r_{B1,B2}^3} \right\rangle \tag{7}$$

The second rate $\Gamma_{A1,B2/B1,A2}^{iso}$ in eq 7 accounts for the contribution from dipoles A1–B2 and B1–A2 as

$$\Gamma_{A1,A2/B1,B2}^{iso} = \left(\frac{\mu_0}{4\pi}\right)^2 \frac{\gamma_{A1}\gamma_{A2}\gamma_{B1}\gamma_{B1}\hbar^2}{10\pi^2} \frac{2\tau_c}{5} \left\langle \frac{3\cos^2\theta_{A1,B2/B1,A2} - 1}{r_{A1,B2}^3 r_{B1,A2}^3} \right\rangle \tag{8}$$

In most cases, this second contribution is so small that it can safely be neglected, and when the average structure is known, this additional contribution can be factored out to a large extent. Given that this contribution cannot be removed experimentally, we have opted to correct the restrained CCR rates for this contribution to the rate and also the effect of anisotropic diffusion. The restrained CCRs were calculated as

$$\Gamma_{A1,A2/B1,B2}^{restrained} = \Gamma_{A/B}^{expt} + \{\Gamma_{A/B}^{calc,iso} - \Gamma_{A/B}^{calc,aniso} - \Gamma_{A1,B2/B1,A2}^{calc,iso}\} \tag{9}$$

where the calculated isotropic and anisotropic CCRs were derived from the average structure of GB3.[48,49,56] Note that such a correction for anisotropy is only justified for small deviations. In the case of GB3, the maximum corrections are ca. 5%, but the majority is much smaller. These CCR rates were restrained using a harmonic restraint

$$E_{CCR} = w_{CCR}(\Gamma_{A1,A2/B1,B2}^{calc} - \Gamma_{A1,A2/B1,B2}^{restrained})^2 \tag{10}$$

DOI: 10.1021/jacs.6b01447
J. Am. Chem. Soc. 2016, 138, 8412–8421

Table 1. Validation of Ensembles and Single Structures with CCR and ^{3}J Coupling Data

	CCR rmsd (s^{-1})[a]				^{3}J coupling rmsd (s^{-1})[b]			
	H^N–N/H^NN	C^α–H^α/C^α–H^α	H^N–N/C^α–H^α	C^α–H^α/H^N–N	H^N–H^α	C′–C′	H^N–C′	H^N–C^β
expt. err	0.19	1.37	0.20	0.26	0.14	0.03	0.10	0.07
Single Structures								
2OED[c]	1.12	4.71	1.68	1.67	0.98	0.41	0.44	1.07
2OED[d]	1.03	4.10	1.91	1.42	0.99	0.39	0.42	1.04
CCR1[e,h]	0.65	4.11	0.98	1.23	0.87	0.62	0.63	1.12
Ensembles								
ENS8[e,f]	0.78	3.99	1.61	1.28	0.75	0.33	0.40	1.02
ENS16[e,g]	0.55	3.74	1.16	1.04	0.71	0.24	0.43	0.68
CCR16[e,h]	0.15	2.02	0.25	0.29	0.55	0.25	0.39	0.72

[a]Bond lengths of 1.041 and 1.117 Å were used to calculate CCRs from structures to account for libration motions that are not present in static structures. [b]Karplus parameters used were those for the fits to Ace–Ala–NMe[68] and experimental data.[69,70] [c]Coordinates from RDC-refined X-ray structure[56] whose H^N and H^α proton positions were subsequently optimized with RDCs.[48,49] [d]Coordinates rerefined with ^{3}J couplings.[71] [e]Average values calculated from 20 ensembles. [f]Previously determined dynamic ensemble.[22] [g]Ensemble generated using the same data as CCR16 except with exclusion of CCR data. [h]Ensemble generated using the CCR data and both 6 sets of H^N–N and H^α–C^α RDCs; J couplings are only used for independent validation.

where w_{CCR} is the weighting factor for the potential term. The weights w_{CCR} were set to 0.4, 0.8, 0.8, and 2.0 kcal/mol·s^2 for $\Gamma_{HNiNi/HNi+1Ni+1}$, $\Gamma_{HNiNi/HAiCAi}$, $\Gamma_{HNiNi/H\alpha i-1C\alpha i-1}$, and $\Gamma_{H\alpha iC\alpha i/H\alpha i-1C\alpha i-1}$, respectively. In the context of the current work, the value of $\Gamma^{calc}_{A1,A2/B1,B2}$ is given as a slightly modified version of eq 7 as

$$\Gamma^{calc}_{A1,A2/B1,B2} = \left\{\left(\frac{\mu_0}{4\pi}\right)^2 \frac{\gamma_{A1}\gamma_{A2}\gamma_{B1}\gamma_{B2}\hbar^2}{10\pi^2}\frac{2\tau_c}{5}\frac{1}{r^3_{A1,A2}r^3_{B1,B2}}\right\} \langle 3\cos^2\theta_{A1,A2/B1,B2} - 1\rangle \quad (11)$$

where the region in curly braces is a constant. This expression substantially simplifies the calcuations of derivatives during structure calculation.

RESULTS

We have recorded 30 data sets in total that encompass H^N–N/H^N–N, H^α–C^α/H^α–C^α, and both intraresidual and sequential H^N–N/H^α–C^α CCR rates. These four pairs of interactions describe the motions of the bond vectors in the backbone (see Figure 1).

We employed as diverse methods as possible but replicated each experiment at least twice. On the basis of extensive statistical analysis, we finally selected 21 sets to obtain the best averaged values and reliable errors for this study (see Tables S1–S4).[57] The errors are exceptionally small as they fall between 1.5 and 3% of the observed ranges of the rates. Typically, they are twice as large as the propagated random errors obtained from repetition of individual experiment types (H^α–C^α/H^α–C^α) or similar to them (all others CCR rates).

Agreement of Static and Dynamic Structures with CCRs. As an initial assessment of the structural and dynamic information contained in the CCRs, we compared the average solution structures of GB3 and the previously determined dynamic ensemble of GB3 to the CCRs.[22] In Table 1 is given the root mean square deviation (rmsd) agreement between the experimental and back-calculated CCRs using eq 1. Focusing first on the average solution structures for H^N–N/H^N–N CCRs, the rmsd of 1.12 s^{-1} for the RDC-refined average solution structure (Protein Databank ID: 2OED) is larger than the experimental error of 0.19 s^{-1}. A similar rmsd (1.03 s^{-1}) was obtained for a newly refined structure with ^{3}J couplings and RDCs. For the H^α–C^α/H^α–C^α CCRs, the average structure refined with RDCs has an rmsd of 4.71 s^{-1}, compared to the experimental error of 1.37 s^{-1}, whereas the structure refined with ^{3}J couplings shows improved agreement to the CCRs with an rmsd of 4.10 s^{-1}. Similarly, the agreement of the intraresidual and sequential H^N–N/H^α–C^α CCRs with experimental errors of 0.20 and 0.26 s^{-1}, respectively, ranges from 1.4 to 1.9 s^{-1}. A uniform scaling factor relating the experimental to the predicted CCR rates may be viewed as an "average" CCR order parameter S^2_{CCR}. For the first 2OED structure in Table 1, it is 0.81, 0.88, 0.86, and 0.85 when individually calculated for H^N–N/H^N–N, H^α–C^α/H^α–C^α, and intraresidual H^N–N/H^α–C^α and sequential H^α–C^α/H^N–N (see Table S9), which is in fair agreement with previous less quantitative studies.[32,37−40,42,44] Note that the values decrease to 0.72, 0.76, 0.75, and 0.74 when the bond lengths are not corrected for libration motion. Finally, we have calculated a single structure refined with both CCRs and RDCs called CCR1. These single structures show some improved agreement with the CCRs, compared to the 2OED structures; however, agreement with ^{3}J couplings was similar. Therefore, we choose to use the original 2OED structures as the representation of the average structure in solution.

Comparison of the CCR rmsd values of the refined average structures to those for dynamic ensemble of GB3 determined by Clore et al. shows better agreement with all four sets of CCRs when compared to the RDC-refined average structures.[22] In Table 1, we also present the independent validation of the average structures with ^{3}J couplings for the backbone (H^N–H^α, C′–C′, H^N–C′, and H^N–C^β). The agreement with the ^{3}J couplings for both structures refined with RDCs is very similar. The independent validation of the ensembles with ^{3}J couplings shows that the ensemble description of structure and dynamics leads to a superior agreement with the experimental data. For all of the structures and ensembles, the Pearson correlation coefficients are very similar (data not shown), indicating that the relative scaling of the dynamic contributions is the major cause of the variation in rmsd.

Generation of CCR-Restrained Ensemble CCR16. Motivated by the poor agreement of the experimental CCRs with the average structures and the improvement observed for the ensemble representation, we undertook to generate an ensemble restrained with the experimental CCRs. We used

DOI: 10.1021/jacs.6b01447
J. Am. Chem. Soc. 2016, 138, 8412−8421

ensemble structure calculations of GB3 restraining both six sets of RDCs for H^N–N and H^α–C^α bond vectors and four sets of CCR measurements converted to their approximate isotropic values using the average structure. The ensemble was determined using Xplor-NIH with the same general protocol used by Clore et al.[22] The angular CCR restraint for isotropic overall tumbling was added, and ensembles of varying sizes between 2 and 64 members were obtained with 20 replicate ensembles for each size. The optimum ensemble size of 16 members was determined through independent validation with ^3J couplings that similar to CCRs and RDCs are sensitive to fast and slow motions of the backbone. In Figure 2a, we show the

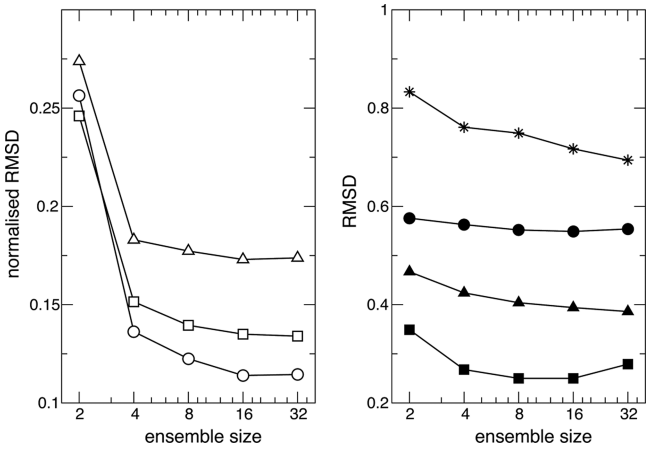


Figure 2. Ensemble-restrained data and independent validation of various ensemble sizes. Normalized rmsd for restrained data are shown in the left panel: RDCs (open circles), experimental CCR rates corrected for anisotropy (open squares), and indirectly restrained anisotropic CCRs (open triangles). The rmsd for unrestrained ^3J scalar couplings H^N–H^α (filled circles), C′–C′ (filled squares), H^N–C′ (filled triangles), and H^N–C^β (stars) are shown in the right panel.

impact of ensemble size on the ability to minimize the disagreement between the experimental and back-calculated RDCs and (anisotropic) CCRs. We also show the agreement with the CCRs after correction for anisotropy.

In Figure 2 we also show the independent validation of the optimum ensemble size using H^N–H^α, C′–C′, H^N–C′, and H^N–C^β ^3J couplings. It can be seen that no further improvement is obtained in increasing the size of the ensemble from 16 to 32 members. In Table 1 we present the validation of ensembles of 16 members with (CCR16) and without CCR (ENS16) restraints.

For the 16-member ensembles restrained with RDCs and CCRs, the agreement with the CCRs approach the experimental error of the CCRs. As a control, we repeated the 16-member ensemble without restraining the CCRs but kept the rest of the restraints and simulation parameters the same. This ensemble had good agreement with the experimental data and validated better than the single structures and the previous dynamic ensemble for GB3 (see Table 1). However, the independent validation for the H^N–H^α and H^N–C′ ^3J couplings is better for the CCR-restrained ensembles than for the CCR-unrestrained ensemble. The slightly poorer agreement for the H^N–C^β couplings is likely related to the inability of the Karplus parameters to accurately account for the influence of different side chain substituents. The correlation plots for the four sets of CCRs for the single structure and ensembles generated in this work are shown in Figure 3. We also conducted cross validation by systematically leaving out each set of CCRs from the calculation (see Table S10); these results show that even though the backbone is not overdetermined by the four sets of CCRs similar results are obtained in most cases. Also revealing in Table S10 is that the H^N–H^α ^3J coupling fit is sensitive to the presence of the intraresidual H^N–N/H^α–C^α CCR data, both of which are sensitive to the backbone ϕ angle.

Average Correlation from Predicted CCR Rates. The average F_{corr} calculated over the entire protein indicates the average degree of correlation throughout the protein. In Table 2 the average F_{corr} values for each pair of bond vectors is presented. The average F_{corr} values calculated from the back-calculated order parameters and CCR rates from the CCR-restrained ensemble are also presented.

The average F_{corr} for the H^N–N/H^N–N bond vectors quantifies the average correlation between neighboring peptide planes and that closest to one and indicates that the average degree of correlation between these probes is the weakest. The correlation between the side chains probed via the H^α–C^α/H^α–C^α CCR rates are of intermediate strength as is the correlation between side chain and its i + 1 peptide plane (sequential H^N–N/H^α–C^α). The intraresidue H^N–N/H^α–C^α correlation has the most correlated behavior of all probes.

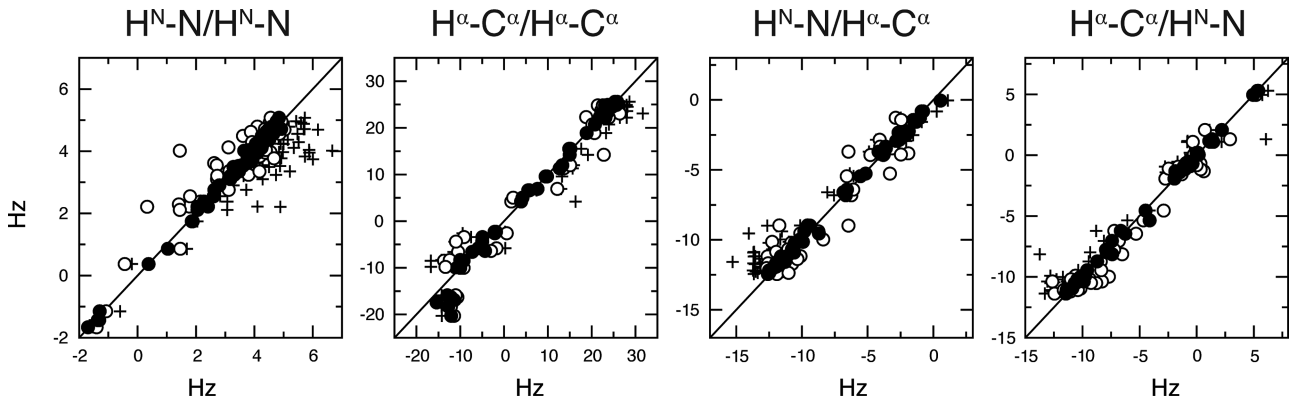


Figure 3. Comparison of experimental and back-calculated anisotropic CCR rates. The rates are calculated for the single average RDC-refined X-ray structure[56] whose H^N and H^α proton positions were subsequently optimized with RDCs[48,49] (+), the RDC restrained ensemble with 16 replicas (○), and the RDC- and CCR-restrained ensemble with 16 replicas (●).

DOI: 10.1021/jacs.6b01447
J. Am. Chem. Soc. 2016, 138, 8412–8421

Table 2. Average F_{corr} from Experimental S^2 and CCR Rates

	F_{corr} [a]			
	H^N–N/H^N–N	C^α–H^α/C^α–H^α	H^N–N/C^α–H^α	C^α–H^α/H^N–N
iterative DIDC[b]	1.01	1.05	1.10	1.02
ORIUM[b]	1.24	1.28	1.39	1.29
Ccr16[c]	1.10	1.14	1.14	1.04

[a]Bond lengths of 1.041 and 1.117 Å were used to calculate CCRs from structures to account for libration motions that are not present in static structures. [b]Coordinates from RDC-refined X-ray structure[56] whose H^N and H^α proton positions were subsequently optimized with RDCs[48,49] [c]Average values calculated from 20 ensembles.

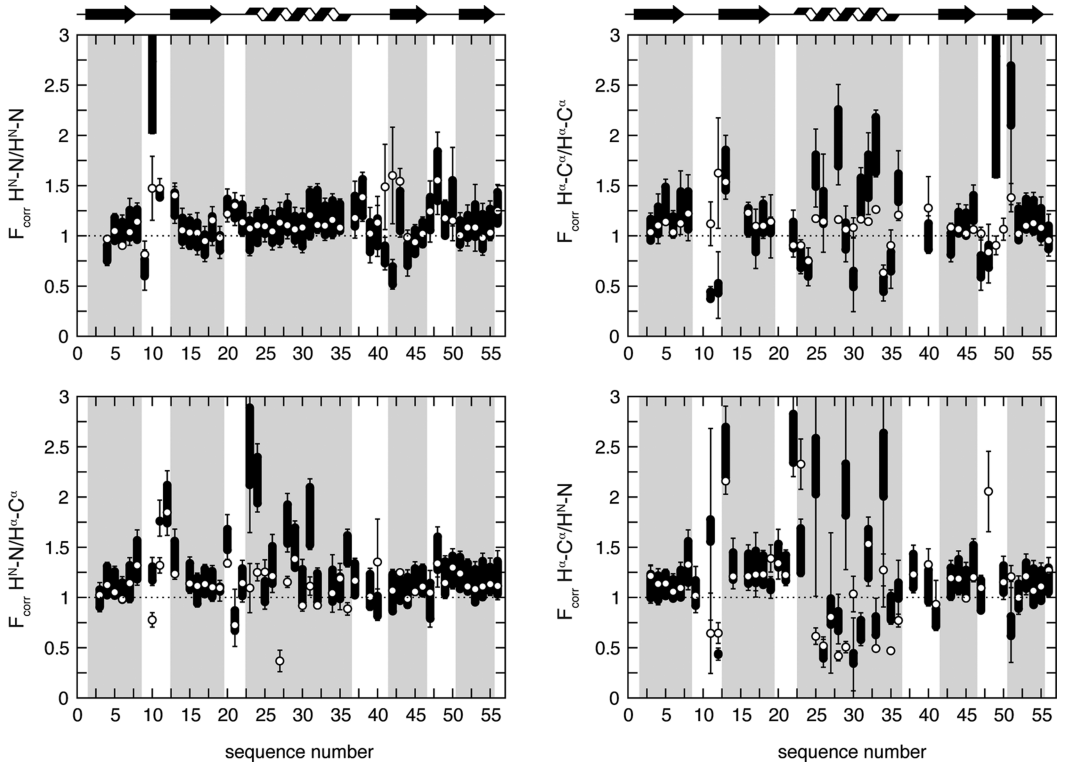


Figure 4. Residue specific F_{corr} values for GB3. Motional correlation between and among H^N–N and H^α–C^α bonds in GB3. F_{corr} versus residue numbers are shown for H^N_i–N_i/H^α_i–C^α_i and H^N_i–N_i/H^α_{i-1}–C^α_{i-1} on the left top and bottom, respectively, and for H^N_i–N_i/H^N_{i+1}–N_{i+1} and H^α_i–C^α_i/H^α_{i-1}–C^α_{i-1} on the right top and bottom. Black thick bars connect the lower F_{corr} estimate from the iDIDC order parameters with the higher estimate from the ORIUM order parameters. The error bars indicate the propagated error from the CCRs and order parameters. If a value or an error is not available from a specific CCR rate or S^{RDC} data set, then the bar is calculated as outlined in the Supporting Information. The white points indicate F_{corr} calculated from the ensemble that was restrained with the CCR and RDC data. Errors for this F_{corr} are the rmsd for the F_{corr} from 20 independent ensembles.

Despite the uniformity across different CCR types, there is considerable variation between individual bond pairs.

Accuracy of Model-Free F_{corr}. For a detailed quantification of motional correlation, the F_{corr} value is calculated for each CCR rate (listed in Tables S5–S8). In Figure 4, F_{corr} is plotted versus the residue number. It is crucial for the current study to establish the reliability of F_{corr}. A realistic error is obtained from error propagation of the well-defined errors of the experimental CCR rates and the RDC order parameters. A negligibly small error is assumed for the predicted CCR rate of the rigid GB3 model (see Supporting Information for details). The resulting error bars mostly lie above or below 1. This suggests that correlated motion (both syn and anti) can be deduced from individual CCR rates (see eq 5). In addition to the error propagated from the RDCs, there is also a model error that can impact on the determined order parameters. Here we use the order parameters derived from both the iterative DIDC (iDIDC) and the ORIUM methods,[48,53,58] which essentially correspond to upper and lower limits for the real order parameters. The ORIUM order parameters are shifted toward lower values (Figure S2; for a residue specific comparison see Figure 4 in ref 53); however, the two models result in a similar sequence-specific pattern (Figure 4). In this way the F_{corr} derived from the ORIUM order parameters is an upper estimate, whereas the F_{corr} derived from the iDIDC is a lower estimate. In Figure 4 we show the range of correlations as a result of changing the RDC analysis model. It is clear that changing the model shifts the values of F_{corr} but conserves the sequence-specific pattern.

Comparison of the Dynamic Ensemble and F_{corr} from the Model-Free Approach. To determine to what extent the dynamic ensemble agrees with the F_{corr} calculated directly from the experimental data (see below), we calculated the F_{corr} from the back-calculated CCRs and S^2 values from the restrained ensemble and its averaged coordinates. In Figure 4 we plot the values of the ensemble F_{corr} values compared with those from

DOI: 10.1021/jacs.6b01447
J. Am. Chem. Soc. 2016, 138, 8412–8421

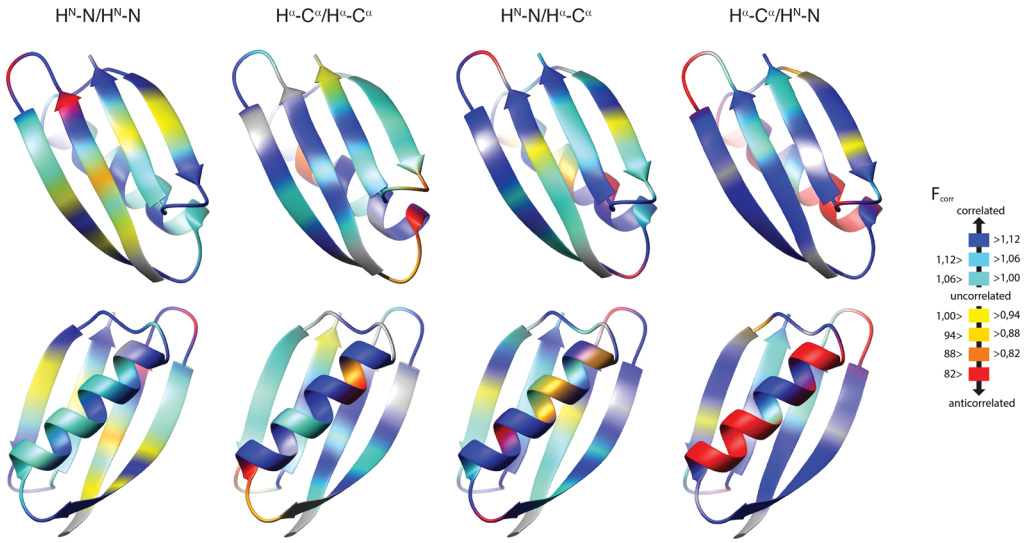


Figure 5. Residue specific F_{corr} values for GB3 from structure calculations. H^N-N/H^N-N, $H^\alpha-C^\alpha/H^\alpha-C^\alpha$, and intraresidual and sequential $H^N-N/H^\alpha-C^\alpha$ F_{corr} are mapped on 3D ribbon representations of GB3. The β-sheet is in the front in the top row, and the α helix is in the bottom row. The plots were prepared with the program MolMol.[72]

the experimental data. We observe good agreement between the ensemble determined F_{corr} and the experimental F_{corr} values. In Figure 5 we show the distribution of the F_{corr} value on the average structure of GB3 to indicate which parts of the structure undergo concerted motions. F_{corr} calculated from the CCR-ensemble-restrained structure calculations usually gives values that are intermediate between those for the ORIUM and iDIDC order parameter calculations. The largest exceptions are found for the loops and are mostly associated with large experimental errors. Large errors are typically obtained for small measured CCR rates. For example, intraresidue $H^N-N/H^\alpha-C^\alpha$ F_{corr} of residue 27 obtained from the DIDC and the ORIUM analysis is −0.04 and −0.06, respectively (see Table S7). Because the measured CCR rate is only 0.96 s^{-1}, the error propagated into F_{corr} is 0.36 and 0.48 s^{-1}. Therefore, these values are not in contradiction with F_{corr} obtained from the ensemble, which is 0.37 s^{-1}. All other exceptions are obtained for the helix, typically with small deviations from the structure-free F_{corr} values. As already inferred from the average F_{corr} values in Table 2, the bond motions are on average slightly syn-correlated. However, the per residue variation in the F_{corr} values is highly complex (see Figure 4), and while many correlations are dominated by the specific local environment, some patterns are typical for entire secondary structure elements.

For example, H^N-N/H^N-N $F_{corr} \approx 1$ for residue 4, which reports on the correlation of the amide bond vector of Tyr3 with that of the amide bond vector of residue Lys4 in the molecular frame. The side chain between these two residues (Tyr3) is buried in the core of the protein, whereas the neighboring side chains project out into the solvent. In contrast, $F_{corr} > 1$ for residue Leu5, which reports on the correlation between the amide bond vectors of Leu5 and Lys4, where the side chain between these two vectors (Lys4) is exposed to the solvent. This alternating pattern of F_{corr} repeats along the length of the strand with F_{corr} for residue Val6 having a value <1, and that of residue Ile7 having a value >1. The same zigzag type pattern is observed across the entire set of strands of GB3 as can be seen by the yellow strips that are present on the strands in Figure 5. This regular behavior of the strands contrasts with the irregular correlation pattern for the H^N-N/H^N-N F_{corr} in the loops. H^N-N/H^N-N $F_{corr} > 1$ throughout the helix. However, F_{corr} on the interior surfaces of the helix that are packed against the hydrophobic core of GB3 have slightly higher values. In the case of the $H^\alpha-C^\alpha/H^\alpha-C^\alpha$ vectors, $F_{corr} > 1$ in the strands with the only exception being the C-terminal residue. The behavior in the helix is more complex, and the majority of residues show $F_{corr} > 1$. For four residue pairs, 22/23, 23/24, 33/34, and 34/35, at the ends of the helices, however, the vectors are negatively correlated. For intraresidue $H^N-N/H^\alpha-C^\alpha$ F_{corr}, the majority of the bond vectors are positively correlated. These vectors span the shortest distance that we have measured here and are only separated by a single bond. In the strands, F_{corr} for residue Val6 is <1, as are the values for Trp43 and Thr44. These residues do not directly contact one another, but the side chains of Thr53 is positioned between their side chains on the surface of the protein. The residues with $F_{corr} < 1$ in the helix are 27, 30, 32, and 36. Residues 32 and 36 are hydrogen-bonded in the helix. Finally, the sequential $H^N-N/H^\alpha-C^\alpha$ $F_{corr} > 1.0$ for all residues in the strands with the exception of the $H^\alpha-C^\alpha$ vector of Thr44 with the H^N-N vector of Tyr45, which is very close to 1.0. In contrast the helix has F_{corr} values predominantly <1.0, indicative of negative correlations in this secondary element. The residues with $F_{corr} > 1$ in the helix are close to the center of the helix.

In general, when both bonds are located in residues with solvent-exposed side chains, strong syn correlation is present. However, anti correlation is more common if at least one side chain is anchored in the protein core. The loops show highly nonuniform behavior. A specific bond may be in strong anti correlation with one bond but in syn correlation with another.

Large-Scale Correlations. In addition to local correlations that are encoded in the input data, the high information density also offers insight into fluctuation of large segments of GB3. In the following, we investigate collective motions of the β-sheet and the α-helix as well as the motion of the helix relative to the sheet.

To quantify the bending of the α-helix, we choose to fit ideal helices to the segments on either end of the helix comprising residues 23−26 and 32−35, and calculate the angle between those. The ensembles have average angles of 14.0 ± 0.5° with a

DOI: 10.1021/jacs.6b01447
J. Am. Chem. Soc. 2016, 138, 8412−8421

rmsd of 6.2 ± 0.7°. The consistency of these numbers indicate that the extent of bending is reproducible with the ensemble. The bend per residue is 1.6°. The results are very similar when the CCR rates are omitted in the structure calculation with 13.9 ± 0.6 and 5.5 ± 0.7°, but the bending decreases to 8.0 ± 1.2° with a slightly smaller rmsd of 3.3 ± 0.6°.

For analysis of the β-sheet, we fitted a catenoid shape described by the inclined angle of β-strands and minimal radius[59] to the β-sheet.[60] The average inclined angle for the CCR+RDC ensemble is very similar to that of the RDC ensemble (42.7 ± 0.3 and 42.6 ± 0.3°), whereas the average angle was marginally larger in the control simulation (45.7 ± 0.6°). The rmsd that represents the structural heterogeneity in the ensemble is 1.6 ± 0.3° for both the CCR+RDC and RDC ensembles; however, the control ensemble is less heterogeneous with a rmsd of 1.1 ± 0.2°. The minimal radius of the catenoid is inversely related to the twisting of the β-sheet.[59] For the control simulation, the average radius is 12.83 ± 0.19 Å and decreases to 11.56 ± 0.19 Å for the RDC ensemble and 11.51 ± 0.15 Å for the CCR+RDC ensemble. This indicates that the β-sheet of the RDC and RDC+CCR ensembles are on average more twisted that of the control simulation. The rmsd of the minimal catenoid radius for the control, RDC, and CCR+RDC ensembles are 0.48 ± 0.08, 0.56 ± 0.12, and 0.65 ± 0.11 Å, respectively. Thus, we observe a broader distribution of minimal catenoid radii for the CCR+RDC ensemble. This increased heterogeneity indicates that the CCR+RDC ensemble samples a broader range of twists than the other ensembles.

DISCUSSION

The zigzag trend of the strong syn-correlated motions along the β-sheet sequence observed in all CCR types may be related to the fact that every other residue has a solvent-exposed side chain. Moreover, neighboring peptide planes in the helix have hydrogen bonds that are approximately opposite one another; thus, the observed anti-correlation of the sequential $H^N-N/H^\alpha-C^\alpha$ CCRs may be related to the slow bending mode of helices.[61] The apparent coupling between the side chains and the backbone results from the $H^\alpha-C^\alpha-C^\beta$ group acting as a hinge that propagates motions of the side chain to the neighboring peptide planes. Consequently, the degree of backbone correlation is affected by side-chain fluctuations, as revealed by stronger syn- or anti-correlated motions. Indeed, backbone motion slower than the overall tumbling has previously been observed for solvent exposed residues in ubiquitin.[62] The highly nonuniform behavior in the loops is caused by larger amplitude and possibly multiple-mode fluctuations.

What we learn from the restrained ensemble structure calculations is that much of the correlations that we observe are not encoded in the force field alone and that the inclusion of the additional experimental data leads to improved agreement of the resulting ensemble. In a previous analysis, Clore and Schwieters generated an ensemble of GB3 structures restrained by RDCs, J couplings, and relaxation order parameters, from which they indirectly obtained correlation functions among and between H^N-N and $H^\alpha-C^\alpha$ bonds (see Supporting Information for details on the calculation).[19] Although relaxation order parameters are only sensitive to subnanosecond motion and their model does not distinguish between syn- and anti-correlation, we compare their findings to our correlation measure F_{corr} that can distinguish between syn- and anti-correlations. Most strikingly, the patterns involving an $H^\alpha-C^\alpha$ bond are less uniform than those of H^N-N/H^N-N. Residue 11 has also the most pronounced F_{corr} for H^N-N/H^N-N, and those of residues 11 and 40 deviate strongly from 1 for $H^\alpha-C^\alpha/H^\alpha-C^\alpha$ and intraresidual and sequential $H^N-N/H^\alpha-C^\alpha$. More generally, the loops exhibit stronger deviation from 1, in particular loops $\beta1/\beta2$ (residues 9−12) and $\beta3/\alpha$ (37−41). One region where our analysis gives different results is the helix, where the F_{corr} values are not more pronounced for the N-terminal segment than in the C-terminal segment. The remarkable agreement for ensemble calculations is not unprecedented because other studies have also shown that experimental restraints often improve the quality of the underlying force field.[63] Comparison of our GB3 ensemble and the structure-free method indicates that the correlations are present in the backbone and can be detected using either the direct correlation via the calculation of F_{corr} or through the use of ensemble restrained simulations with RDCs and CCR meausrements.

In numerous studies of GB3, a crankshaft motion has been described.[19,22,24,64] Such anti-correlated fluctuations of the backbone φ_i and ψ_{i-1} angles is equivalent to tilting of the peptide plane around the $C^\alpha-C^\alpha$ axis. In the CCR-restrained ensemble, we observe two short-range correlated motions, the crankshaft and β-lever correlations (see Figure S3). However, on the longer time scales the long-range collective motions are predicted to dominate motions and hence the F_{corr}.[51] Consistent with this idea of angular fluctuations of largely rigid protein segments, we have observed a dominance of syn-correlated rather than anti-correlated motion.

Bouvignies et al. used the 3D Gaussian model with a large set of RDCs to determine the motional amplitudes (but not the degree of correlation).[64] Despite this, the correlated motion of the β-sheet on the nano- to millisecond time scale was observed to improve the agreement with experimentally measured hydrogen bond ^3J scalar couplings. In a related study, comparison of RDC and relaxation order parameters to those from accelerated molecular dynamics (AMD) simulations indicated the existence of micro- to millisecond motion in the loops and the β-sheet but not in the α-helix.[23] Also in the AMD study, an alternation of large and small motional amplitudes was observed in the strands $\beta1$, $\beta3$, and $\beta4$, pointing to motions coupled across the β-sheet. Interestingly, the pattern matches the alternation of strongly hydrophobic side chains buried in the protein core observed here and may be related to the weak zigzag pattern observed in this study.[63] Recently, we used exact NOEs to calculate an ensemble of structures.[24] Again, correlated motion across the β-sheet and within the loops was revealed. This is in line with the present observation, where CCR rates involving $H^\alpha-C^\alpha$ bonds indicate correlations to neighboring backbone bonds.

In the biological context of GB3, the N-terminal segment of strand $\beta2$ is a site of interaction with the Fc immunoglobulin region.[56] This region is often observed to exhibit increased H^N-N bond mobility in experiments that probe dynamic fluctuations.[19,21,48,64,65] The current study suggests that this additional mobility is organized in a correlated fashion. From the ensemble we see indications that the sheet is twisting and that the helix is bending. Despite the consistencies with other data, the CCR data does not strictly probe the same aspect of motion. It is rather a complement to other NMR measures, both in terms of time scale and spatial dependence. In conclusion, we introduced a method that quantifies the

DOI: 10.1021/jacs.6b01447
J. Am. Chem. Soc. 2016, 138, 8412−8421

correlation of protein native state motion. In principle, the method can be extended to include other protein side chain bonds and to study nucleic acids. Ultimately, we hope that the methods presented here that allow the determination of local disorder and correlated motions will open an avenue to refine local entropy calculations in biomolecules.

MATERIALS AND METHODS

Sample Expression and Purification. GB3 was expressed and purified as described previously.[66] The $^{13}C,^{15}N$- and $^{2}H,^{13}C,^{15}N$-labeled NMR samples contained 350 and 500 μL of 4 and 2 mM protein solution, respectively, in 97%/3% and 95%/5% H_2O/D_2O, 50 mM potassium phosphate buffer, pH 6.5 and 7.0, and 0.5 mg/mL sodium azide.

NMR Spectroscopy. All experiments were performed on a BRUKER DRX600 MHz spectrometer, equipped with a z-axis gradient cryogenic probe, respectively, at 298 K. All spectra were processed and analyzed using the software package NMRPipe.[67]

$\Gamma_{HNiNi/HNi+1Ni+1} + \Gamma_{HNiNi+1/HNi+1Ni}$ were obtained from the triply labeled sample and "reference" and "trans" spectra of 3D ct-$^{13}C'$–HN(CA)CON [ct-$^{13}C^{\alpha}$–HNCA(CO)N] experiments (derived from a 2D version)[45] recorded with 36(N, t_1) × 40(C′, t_2)[20(C^{α}, t_2)] × 512(H^N, t_3) complex points, t_{1max} = 18.0 ms, t_{2max} = 26.4[6.6] ms, t_{3max} = 51.2 ms, an interscan delay of 1 s, τ_{MQ} = 43 ms, and typically 16[32] scans per increment resulting in a measurement time of 1[2] day(s). Because the trans spectra are considerably less sensitive they were typically recorded twice and added up thereby effectively doubling the number of scans. The time domain data were multiplied with a square cosine function in the direct dimension and cosine functions in the indirect dimensions and zero-filled to 256 × 128 × 2048 complex points. Spectra with $^{13}C'$ and with $^{13}C^{\alpha}$ evolution were recorded twice and three times, respectively, such that averaged CCR rates with associated random errors could be obtained.

$\Gamma_{H\alpha iC\alpha i/H\alpha i-1C\alpha i-1} + \Gamma_{H\alpha iC\alpha i-1/H\alpha i-1C\alpha i}$ were obtained from "reference" and "trans" spectra of the 2D HNCA(CA) experiment previously put forward,[37] and a 3D $^{13}C^{\alpha}$–HNCA(CO)CA experiment recorded with 64(N, t_1) × 512(H^N, t_2) and 36(N, t_1) × 24(C^{α}, t_2) × 512(H^N, t_3) complex points, t_{1max} = 32.0 ms, t_{2max} = 51.2 ms and t_{1max} = 18.0 ms, t_{2max} = 7.92 ms, t_{3max} = 51.2 ms, an interscan delay of 1 s, τ_{MQ} = 28 ms, and typically 256 and 32 per increment resulting in measurement times of 0.5 and 2 days, respectively. Because the trans spectra are considerably less sensitive than the reference spectra, they were typically recorded twice and added up, thereby effectively doubling the number of scans. The time domain data were multiplied with a square cosine function in the direct dimension and cosine functions in the indirect dimensions and zero-filled to 1024 × 4096 and 256 × 128 × 2048 complex points. All spectra were recorded twice so that average CCR rates and associated random errors could be obtained.

$\Gamma_{HNiNi/H\alpha iC\alpha i} + \Gamma_{H\alpha iNi/HNiC\alpha i}$ were obtained from three experiments. The two experiments used previously were repeated.[52] The first approach relies on the ACE (all component evolution) method[52] realized in a 3D ct-HNCA experiment, where all components of the quadruplets of the multiquantum coherences are resolved. In the second approach, the DIAI (double in-phase/anti-phase interconversion) method is realized with a pair of 3D HNCA pulse sequences ("reference" and "trans") in 2 days.[38] In addition, a 3D ct-HNCA MMQ (mixed multiquantum, where zero- and double-quantum coherence evolution is averaged) experiment was used. The ZQ and DQ coherences are superimposed, resulting in four components to be evaluated.[44] The spectrum was recorded twice with τ_{MQ} = 31.0 ms or τ_{MQ} = 33.5 ms, 50(MQ[N,C^{α}], t_1) or 55(MQ[N,C^{α}], t_1) × 36(N, t_2) × 512(H^N, t_3) complex points, t_{1max} = 25.0 or 27.5 ms, t_{2max} = 18.0 ms, t_{3max} = 63.28 ms, interscan delays of 1 or 0.92 s, and 16 or 24 scans per increment resulting in measurement times of 1.5 or 2 days. The time domain data were multiplied with a square cosine function in the direct dimension and cosine functions in the indirect dimensions and zero-filled to 256 × 128 × 2048 complex points.

$\Gamma_{HNiNi/H\alpha i-1C\alpha i-1} + \Gamma_{H\alpha i-1Ni/HNi-1C\alpha i-1}$ were obtained by repetition of three previously used experiments.[52] The first approach relies on ACE realized in a 3D ct-HN(CO)CA experiment. In the second and third approach, the DIAI method is realized with two types of 2D HN(CO)CA experiments.[36,38]

Corrections to the apparent CCR rates were calculated with a full matrix analysis as outlined in reference 47.

Prediction of CCR Rates. To obtain reliable predicted CCR rates and thus F_{corr}, the average positions of the H^N and H^{α} protons are of particular importance. Highly accurate RDCs have been used to orient H^N–N and H^{α}–C^{α} bond vectors in the NMR-optimized X-ray structure with Protein Databank entry 2OED.[48,49,66] The impact of angular fluctuation is removed by scaling the H^N–N and H^{α}–C^{α} bond lengths to 1.041 and 1.117 Å, respectively.[54] For the prediction of the CCR rates expected for uncorrelated motion of the two vectors, the previously published RDC order parameters were used (Table S11).[48,49,53] Note that structural inconsistencies between the different mutants were identified for residues 19 and 41 for the H^N–N and residues 11, 25, 30, and 40 for the H^{α}–C^{α} bonds, four of which are sites of mutations or immediate neighbors.[48] The errors may be larger than reported and are accordingly propagated into F_{corr}. The CCR rates predicted for the rigid and dynamic models are shown in Tables S5–S8.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.6b01447.

Brief description of NMR experiments and extraction of experimental CCR rates; details on F_{corr} calculations; details of ensemble calculations; correlation plots of experimental and predicted CCR rates assuming a rigid or dynamic molecule; correlation plots of order parameters obtained from the ensembles or from the DIDC and ORIUM analysis; correlation plots for the dihedral ϕ and ψ angles; tables showing experimental CCR rates, errors and corrections for anisotropic tumbling; tables showing F_{corr} values obtained from model-free analysis and ensembles; table comparing experimental CCR rates to predictions using a rigid molecule model; and table listing RDC order parameters obtained from iDIDC and ORIUM analysis (PDF)

AUTHOR INFORMATION

Corresponding Authors

*E-mail: fenwick@scripps.edu. Tel.: (+41)-44-633-4405.
*E-mail: beat.voegeli@phys.chem.ethz.ch. Tel.: (+41)-44-633-4405.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Prof. Roland Riek, Dr. Jason Greenwald, Dr. Michael Sabo, and Dr. Donghan Lee for valuable discussion. We note that a similar correlation factor F_{corr} has been presented by Dr. Donghan Lee at the Swiss Federal Institute of Technology and Dr. Michael Sabo at the 54th ENC conference 2013 in Asilomar. This work was supported by the Swiss National Science Foundation with Grant 140214 to B.V. and by the Intramural Research Program of C.I.T, National Institutes of Health, to C.D.S.

REFERENCES

(1) Zhuravlev, P.; Papoian, G. *Q. Rev. Biophys.* **2010**, *43*, 295.

DOI: 10.1021/jacs.6b01447
J. Am. Chem. Soc. 2016, 138, 8412–8421

(2) Bahar, I.; Lezon, T.; Bakan, A.; Shrivastava, I. *Chem. Rev.* **2010**, *110*, 1463.
(3) Dill, K.; Chan, H. *Nat. Struct. Biol.* **1997**, *4*, 10.
(4) Fersht, A. *Structure and mechanism in protein science: a guide to enzyme catalysis and protein folding*; W.H. Freeman: New York, 1999.
(5) Kern, D.; Zuiderweg, E. *Curr. Opin. Struct. Biol.* **2003**, *13*, 748.
(6) Tsai, C.; Del Sol, A.; Nussinov, R. *Mol. BioSyst.* **2009**, *5*, 207.
(7) Hilser, V. *Science* **2010**, *327*, 653.
(8) Shaw, D.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R.; Eastwood, M.; Bank, J.; Jumper, J.; Salmon, J.; Shan, Y.; Wriggers, W. *Science* **2010**, *330*, 341.
(9) Lindorff-Larsen, K.; Maragakis, P.; Piana, S.; Eastwood, M.; Dror, R.; Shaw, D. *PLoS One* **2012**, *7*, e32131.
(10) Cavanagh, J. *Protein NMR spectroscopy: principles and practice*, 2nd ed.; Academic Press: Amsterdam, The Netherlands, 2007.
(11) Kleckner, I.; Foster, M. *Biochim. Biophys. Acta, Proteins Proteomics* **2011**, *1814*, 942.
(12) Torchia, D. *J. Magn. Reson.* **2011**, *212*, 1.
(13) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546.
(14) Tolman, J.; Al-Hashimi, H.; Kay, L.; Prestegard, J. *J. Am. Chem. Soc.* **2001**, *123*, 1416.
(15) Kay, L. E. *J. Magn. Reson.* **2005**, *173*, 193.
(16) Mayer, K.; Earley, M.; Gupta, S.; Pichumani, K.; Regan, L.; Stone, M. *Nat. Struct. Biol.* **2003**, *10*, 962.
(17) Brüschweiler, R.; Blackledge, M.; Ernst, R. *J. Biomol. NMR* **1991**, *1*, 3.
(18) Bonvin, A.; Brunger, A. *J. Mol. Biol.* **1995**, *250*, 80.
(19) Clore, G.; Schwieters, C. *Biochemistry* **2004**, *43*, 10678.
(20) Lindorff-Larsen, K.; Best, R.; Depristo, M.; Dobson, C.; Vendruscolo, M. *Nature* **2005**, *433*, 128.
(21) Bouvignies, G.; Bernado, P.; Meier, S.; Cho, K.; Grzesiek, S.; Brüschweiler, R.; Blackledge, M. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 13885.
(22) Clore, G.; Schwieters, C. *J. Mol. Biol.* **2006**, *355*, 879.
(23) Markwick, P.; Bouvignies, G.; Blackledge, M. *J. Am. Chem. Soc.* **2007**, *129*, 4724.
(24) Vogeli, B.; Kazemi, S.; Guntert, P.; Riek, R. *Nat. Struct. Mol. Biol.* **2012**, *19*, 1053.
(25) Fenwick, R.; Esteban-Martin, S.; Salvatella, X. *Eur. Biophys. J.* **2011**, *40*, 1339.
(26) Fadel, A.; Jin, D.; Montelione, G.; Levy, R. *J. Biomol. NMR* **1995**, *6*, 221.
(27) Fenwick, R.; Esteban-Martin, S.; Richter, B.; Lee, D.; Walter, K.; Milovanovic, D.; Becker, S.; Lakomek, N.; Griesinger, C.; Salvatella, X. *J. Am. Chem. Soc.* **2011**, *133*, 10336.
(28) Früh, D.; Tolman, J.; Bodenhausen, G.; Zwahlen, C. *J. Am. Chem. Soc.* **2001**, *123*, 4810.
(29) Lundstrom, P.; Mulder, F.; Akke, M. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 16984.
(30) Hubbard, P. *Phys. Rev.* **1958**, *109*, 1153.
(31) Werbelow, L. G.; Marshall, A. G. *J. Magn. Reson.* **1973**, *11*, 299.
(32) Pelupessy, P.; Ravindranathan, S.; Bodenhausen, G. *J. Biomol. NMR* **2003**, *25*, 265.
(33) Vugmeyster, L.; Pelupessy, P.; Vugmeister, B. E.; Abergel, D.; Bodenhausen, G. *C. R. Phys.* **2004**, *5*, 377.
(34) Vogeli, B. *J. Chem. Phys.* **2010**, *133*, 14501.
(35) Reif, B.; Hennig, M.; Griesinger, C. *Science* **1997**, *276*, 1230.
(36) Chiarparin, E.; Pelupessy, P.; Ghose, R.; Bodenhausen, G. *J. Am. Chem. Soc.* **1999**, *121*, 6876.
(37) Chiarparin, E.; Pelupessy, P.; Ghose, R.; Bodenhausen, G. *J. Am. Chem. Soc.* **2000**, *122*, 1758.
(38) Pelupessy, P.; Chiarparin, E.; Ghose, R.; Bodenhausen, G. *J. Biomol. NMR* **1999**, *13*, 375.
(39) Pelupessy, P.; Chiarparin, E.; Ghose, R.; Bodenhausen, G. *J. Biomol. NMR* **1999**, *14*, 277.
(40) Carlomagno, T.; Maurer, M.; Hennig, M.; Griesinger, C. *J. Am. Chem. Soc.* **2000**, *122*, 5105.
(41) Yang, D.; Mittermaier, A.; Mok, Y.; Kay, L. *J. Mol. Biol.* **1998**, *276*, 939.
(42) Takahashi, H.; Shimada, I. *J. Biomol. NMR* **2007**, *37*, 179.
(43) Vugmeyster, L.; McKnight, C. *Biophys. J.* **2008**, *95*, 5941.
(44) Yang, D.; Kay, L. E. *J. Am. Chem. Soc.* **1998**, *120*, 9880.
(45) Pelupessy, P.; Espallargas, G.; Bodenhausen, G. *J. Magn. Reson.* **2003**, *161*, 258.
(46) Gronenborn, A.; Filpula, D.; Essig, N.; Achari, A.; Whitlow, M.; Wingfield, P.; Clore, G. *Science* **1991**, *253*, 657.
(47) Vögeli, B. *J. Magn. Reson.* **2013**, *226*, 52.
(48) Yao, L.; Vogeli, B.; Torchia, D.; Bax, A. *J. Phys. Chem. B* **2008**, *112*, 6045.
(49) Yao, L.; Vogeli, B.; Ying, J.; Bax, A. *J. Am. Chem. Soc.* **2008**, *130*, 16518.
(50) Deschamps, M.; Bodenhausen, G. *ChemPhysChem* **2001**, *2*, 539.
(51) Brüschweiler, R. *J. Chem. Phys.* **1995**, *102*, 3396.
(52) Vogeli, B.; Yao, L. *J. Am. Chem. Soc.* **2009**, *131*, 3668.
(53) Sabo, T. M.; Smith, C. A.; Ban, D.; Mazur, A.; Lee, D.; Griesinger, C. *J. Biomol. NMR* **2014**, *58*, 287.
(54) Case, D. *J. Biomol. NMR* **1999**, *15*, 95.
(55) Schwieters, C.; Kuszewski, J.; Tjandra, N.; Clore, M. G. *J. Magn. Reson.* **2003**, *160*, 65.
(56) Derrick, J.; Wigley, D. *J. Mol. Biol.* **1994**, *243*, 906.
(57) Olsson, S.; Vogeli, B. R.; Cavalli, A.; Boomsma, W.; Ferkinghoff-Borg, J.; Lindorff-Larsen, K.; Hamelryck, T. *J. Chem. Theory Comput.* **2014**, *10*, 3484.
(58) Tolman, J. *J. Am. Chem. Soc.* **2002**, *124*, 12020.
(59) Koh, E.; Kim, T. *Proteins: Struct., Funct., Genet.* **2005**, *61*, 559.
(60) Novotny, J.; Bruccoleri, R. E.; Newell, J. *J. Mol. Biol.* **1984**, *177*, 567.
(61) Emberly, E.; Mukhopadhyay, R.; Wingreen, N.; Tang, C. *J. Mol. Biol.* **2003**, *327*, 229.
(62) Lakomek, N.; Fares, C.; Becker, S.; Carlomagno, T.; Meiler, J.; Griesinger, C. *Angew. Chem., Int. Ed.* **2005**, *44*, 7776.
(63) De Simone, A.; Richter, B.; Salvatella, X.; Vendruscolo, M. *J. Am. Chem. Soc.* **2009**, *131*, 3810.
(64) Bouvignies, G.; Markwick, P.; Brüschweiler, R.; Blackledge, M. *J. Am. Chem. Soc.* **2006**, *128*, 15100.
(65) Bernadó, P.; Blackledge, M. *J. Am. Chem. Soc.* **2004**, *126* (25), 7760.
(66) Ulmer, T.; Ramirez, B.; Delaglio, F.; Bax, A. *J. Am. Chem. Soc.* **2003**, *125*, 9179.
(67) Delaglio, F.; Grzesiek, S.; Vuister, G.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277.
(68) Case, D. A.; Scheurer, C.; Brüschweiler, R. *J. Am. Chem. Soc.* **2000**, *122*, 10390.
(69) Lee, J.; Li, F.; Grishaev, A.; Bax, A. *J. Am. Chem. Soc.* **2015**, *137*, 1432.
(70) Vogeli, B.; Ying, J.; Grishaev, A.; Bax, A. *J. Am. Chem. Soc.* **2007**, *129*, 9377.
(71) Li, F.; Lee, J. H.; Grishaev, A.; Ying, J.; Bax, A. *ChemPhysChem* **2015**, *16*, 572.
(72) Koradi, R.; Billeter, M.; Wüthrich, K. *J. Mol. Graphics* **1996**, *14*, 51.

DOI: 10.1021/jacs.6b01447
J. Am. Chem. Soc. 2016, 138, 8412−8421