

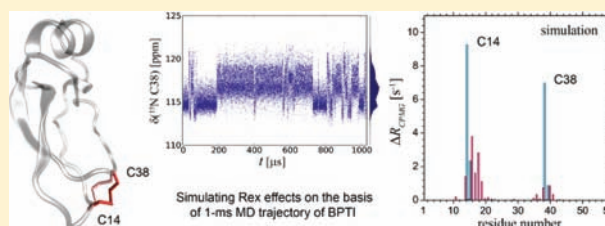
Microsecond Time-Scale Conformational Exchange in Proteins: Using Long Molecular Dynamics Trajectory To Simulate NMR Relaxation Dispersion Data

Yi Xue, Joshua M. Ward,[†] Tairan Yuwen, Ivan S. Podkorytov,[‡] and Nikolai R. Skrynnikov*

Department of Chemistry, Purdue University, 560 Oval Drive, West Lafayette, Indiana 47907-2084, United States

S Supporting Information

ABSTRACT: With the advent of ultra-long MD simulations it becomes possible to model microsecond time-scale protein dynamics and, in particular, the exchange broadening effects (R^{ex}) as probed by NMR relaxation dispersion measurements. This new approach allows one to identify the exchanging species, including the elusive “excited states”. It further helps to map out the exchange network, which is potentially far more complex than the commonly assumed 2- or 3-site schemes. Under fast exchange conditions, this method can be useful for separating the populations of exchanging species from their respective chemical shift differences, thus paving the way for structural analyses. In this study, recent millisecond-long MD trajectory of protein BPTI (Shaw et al. *Science* **2010**, 330, 341) is employed to simulate the time variation of amide ^{15}N chemical shifts. The results are used to predict the exchange broadening of ^{15}N lines and, more generally, the outcome of the relaxation dispersion measurements using Carr–Purcell–Meiboom–Gill sequence. The simulated R^{ex} effect stems from the fast (~ 10 – $100 \mu\text{s}$) isomerization of the C14–C38 disulfide bond, in agreement with the prior experimental findings (Grey et al. *J. Am. Chem. Soc.* **2003**, 125, 14324).



INTRODUCTION

Protein conformational variability plays a key role in enzymatic catalysis, ligand binding, folding, and other processes of functional importance. Notably, transient conformational species with populations on the order of 1% and lifetimes in the microsecond to millisecond range are often indispensable for protein function.^{1–4} Such species, termed “invisible” or “excited” states, are difficult to characterize using conventional experimental methods. So far the most detailed information about the invisible states has been obtained from the NMR relaxation dispersion measurements.^{5–8} These experiments focus on chemical shift (CS) evolution of spin magnetization and, particularly, on the dephasing effect caused by stochastic transitions between different conformational states (i.e., conformational exchange). The CS evolution and the concomitant dephasing effect can be partially canceled by application of 180° pulses or continuous-wave rf field. This constitutes the basis for Carr–Purcell–Meiboom–Gill (CPMG) and spin-lock experiments.^{9–12}

The relaxation dispersion data contain a trove of structural and dynamic information. Under favorable circumstances, it is possible to extract the exchange correlation time $\tau_{ex} = (k_1 + k_{-1})^{-1}$, the populations of the exchanging species, p_A and p_B , and the respective CS differences for the individual spin sites, $\Delta\delta_{ab} = \delta_a - \delta_b$. The latter quantity can be used to recover the chemical shifts of minor species, δ_b , which further serve as a basis for structural characterization.^{13–16}

While this approach has been remarkably successful, it nevertheless faces a number of limitations. First, the detailed

data interpretation is only possible when the exchange occurs on the time scale $\tau_{ex} \sim 1$ ms. It is also required that the relaxation dispersion data are collected at more than one magnetic field strength.¹⁷ In contrast, when exchange is fast, ca. 10 – $300 \mu\text{s}$, it is impossible to separate the populations of the exchanging species from the chemical shifts—only the product $p_A p_B (\Delta\delta_{ab})^2$ can be determined. In the absence of the chemical shift information, structural analysis also becomes impossible.

Second, the two-state exchange scheme provides only a crude approximation for protein conformational dynamics, which is governed by an extremely complex energy landscape. Indeed, the observed site-to-site variations in the extracted values of p_A , p_B , and τ_{ex} immediately suggest that the actual dynamics is more complex than implied by the two-state model. In a number of systems the existence of the third state has been demonstrated experimentally.^{18–20} In general, it can be expected that conformational diversity goes beyond two or three states. Especially in the situations involving disorder, many conformational species coexist in dynamic equilibrium, interconverting on nanosecond and microsecond–millisecond time scales.

Third, interpretation of chemical shift data (δ_b) usually requires some structural guidance. In principle, it is possible to generate a de novo structure based mainly on δ_b data.^{21,22} However, in practice the number of available δ_b shifts is usually small and their accuracy is limited, which makes it difficult to

Received: July 19, 2011

Published: December 29, 2011

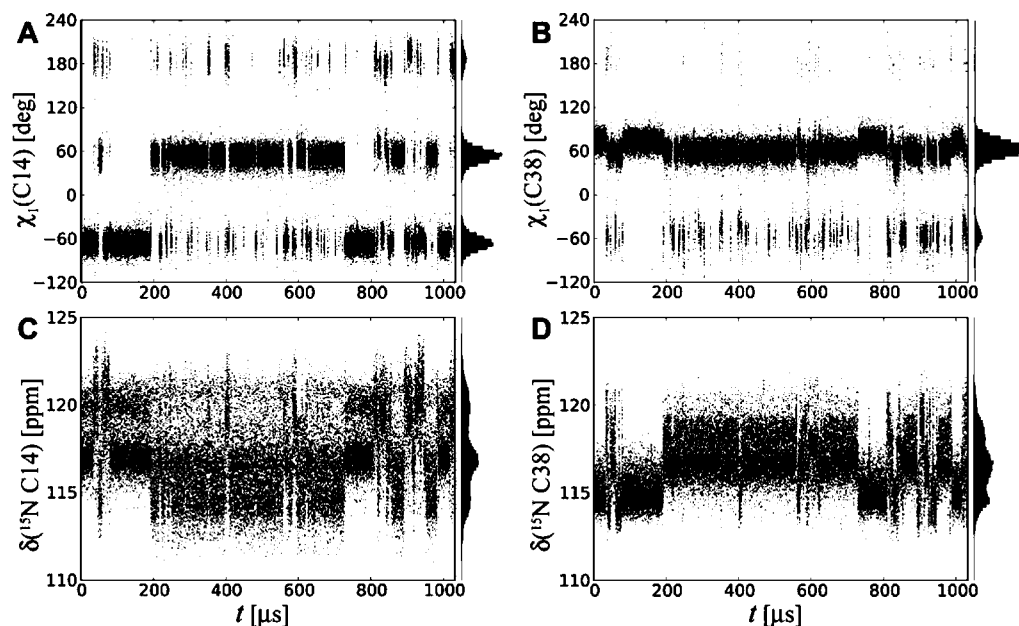


Figure 1. (A,B) Time variation of χ_1 torsional angles in disulfide-bonded residues C14, C38 as extracted from 1-ms MD trajectory of BPTI.³² (C,D) Time variation of backbone $\delta^{15\text{N}}$ chemical shifts in residues C14, C38 as calculated by application of the program SHIFTX+³⁵ to the individual frames from the BPTI trajectory. SHIFTX+ is a part of the recently released package SHIFTX2 (the other module, SHIFTY+, relies on protein primary sequence and thus does not help to predict the dynamic CS variations).³³ Similar strategies for computation of chemical shifts on the basis of MD trajectories have been implemented earlier.^{34–36} Note that $\delta^{15\text{N}}$ undergoes large fluctuations on a fast time scale (<25 ns), as especially well seen in the central portion of the plot C. As it turns out, the peptide plane P13–C14 is “destabilized” in the conformational state m_{C14} and experiences large axial motions (involving a concerted change in $\psi(\text{P13})$ and $\varphi(\text{C14})$).

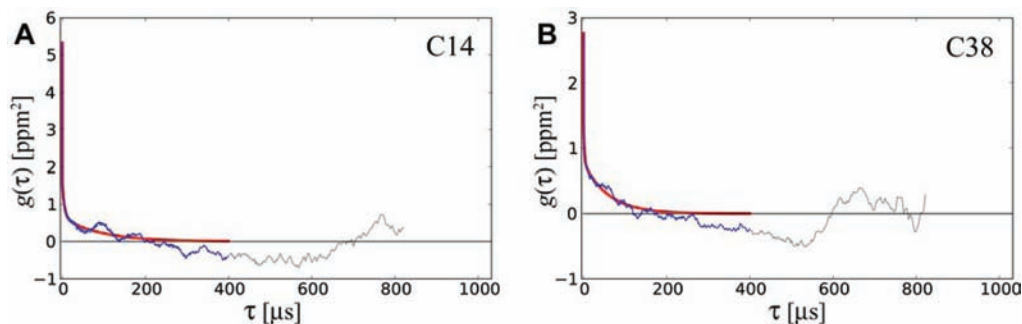


Figure 2. Chemical shift correlation functions for backbone $^{15\text{N}}$ spins from residues C14, C38 in BPTI. Blue/gray curves: $g(\tau)$ calculated via eq 1 from $\delta^{15\text{N}}$ traces shown in Figure 1C,D. Red curves: triexponential fits of $g(\tau)$. Since the tail of the correlation function is statistically unreliable, the fitting is limited to the first 400 μs . The fitting procedure is set up such as to capture the initial fast decay in $g(\tau)$ (time constant < 25 ns, that is, less than the time step of the resampled MD trajectory) plus the two slower components. Use of 4-, 5-, or 6-exponential fitting functions does not change the results. The complete set of correlation functions for all nonproline residues in BPTI is shown in Supporting Information Figure S5.

calculate a reasonable structure. Instead, pre-existing structural models have been widely used to assist in the interpretation of relaxation dispersion data.^{23–27} Alternatively, many investigators sought to boost the population of minor species, for example, by adding a ligand, changing sample conditions, or introducing a point mutation.^{21,28–30} None of these solutions is fully satisfactory—a suitable model may not be available or readily identifiable, whereas changing the experimental conditions may bias the results.

In summary, there is a good deal of uncertainty in the transition from relaxation dispersion data to chemical shifts and further to the structural model of the excited state. Under these circumstances, the insight from MD simulations may prove to be extremely valuable.^{27,31} Indeed, MD trajectory presents a picture of internal motion in all its complexity. Provided that the trajectory is sufficiently long, the results can be used to

simulate the relaxation dispersion profiles which can be then directly compared with the experimental data. Thus a direct connection can be established between the experimental data and the highly realistic structural/dynamic model, without suffering from all of the parametrization issues described above.

RESULTS

Until recently, it has been unthinkable that unbiased explicit-atom MD simulation can reach the length that is necessary to sample microsecond protein dynamics. However, a year ago Shaw and co-workers reported 1-ms-long trajectory of BPTI in explicit water.³² The simulation was conducted at 300 K under the control of (slightly altered) Amber ff99SB force field^{37,38} using a special-purpose supercomputer Anton. The snapshots from this simulation, resampled at 25 ns steps, have been released for use by the research community.

In this report we demonstrate how the ultra-long MD trajectory of BPTI can be used to reproduce the results from NMR relaxation dispersion measurements. As a first step, each frame from the MD trajectory has been processed using the chemical shift prediction program SHIFTX+.³³ The resulting chemical shift traces $\delta(t)$ are illustrated in Figure 1C,D for backbone ^{15}N spins of residues C14 and C38.

Next, we evaluate the autocorrelation functions for the individual spin sites in BPTI:

$$g(\tau) = \langle (\delta(t) - \bar{\delta})(\delta(t + \tau) - \bar{\delta}) \rangle \quad (1)$$

Here $\bar{\delta} = \langle \delta(t) \rangle$, and the angular brackets denote averaging with respect to t . The examples of the correlation functions calculated according to eq 1 are shown in Figure 2. In the fast exchange regime, when Redfield relaxation theory is valid, $g(\tau)$ can be related to the exchange-induced spin relaxation rate:^{39,40}

$$R_2^{\text{ex}} = (2\pi\nu_0)^2 \int_0^\infty g(\tau) d\tau \quad (2)$$

Here ν_0 is the Larmor frequency for a given sort of spins. This result can be extended to describe the outcome of the spin lock experiment:⁴⁰

$$R_{1\rho,\text{SL}}^{\text{ex}}(\nu_1) = (2\pi\nu_0)^2 \int_0^\infty g(\tau) \cos(2\pi\nu_1\tau) d\tau \quad (3)$$

where ν_1 is the frequency of spin lock (or, in other words, the strength of *rf* field).

Finally, the equivalent expression can be derived for the CPMG pulse train, $(\tau_{\text{CP}} - 180^\circ - \tau_{\text{CP}})_n$:

$$R_{1\rho,\text{CPMG}}^{\text{ex}}(\nu_{\text{CP}}) = (2\pi\nu_0)^2 \int_0^\infty g(\tau) \text{tri}(2\pi\nu_{\text{CP}}\tau) d\tau \quad (4)$$

Here $\nu_{\text{CP}} = 1/4\tau_{\text{CP}}$ and $\text{tri}(x)$ is a triangle-wave function which consists of linear segments connecting the points $(x,y) = (0,1)$, $(\pi,-1)$, $(2\pi,1)$, $(3\pi,-1)$, and so forth (i.e., the extrema of the cosine curve). The derivation of eq 4, which is similar to our recent treatment of the phase-alternated spin lock,⁴¹ is described in Appendix A (see Supporting Information). While eq 4 reflects the discrete nature of the CPMG experiment, the result is strikingly similar to the spin-lock case, eq 3. If the correlation function $g(\tau)$ is comprised of several exponentials, then the expression in eq 4 can be readily evaluated in analytical form,

$$g(\tau) = \sum_{j=1}^N c_j \exp(-\tau/\tau_{\text{ex}}^j) \quad (5.1)$$

$$\begin{aligned} R_{1\rho,\text{CPMG}}^{\text{ex}}(\nu_{\text{CP}}) \\ = (2\pi\nu_0)^2 \sum_{j=1}^N c_j \tau_{\text{ex}}^j \left\{ 1 - \frac{\tau_{\text{ex}}^j}{\tau_{\text{CP}}} \tanh \frac{\tau_{\text{CP}}}{\tau_{\text{ex}}^j} \right\} \end{aligned} \quad (5.2)$$

in agreement with the early findings of Luz and Meiboom.⁴² In summary, eqs 1–5 allow for calculation of relaxation dispersion profiles for faster (μs) forms of conformational dynamics on the basis of MD data.

Conformational exchange in BPTI has been investigated in depth in Wüthrich's group^{43,44} and later in Palmer's group.^{17,25,45,46} In particular, Palmer and co-workers reported the results of the ^{15}N CPMG measurements in a form of

(partial) dispersion amplitude, $\Delta R = R_{1\rho,\text{CPMG}}^{\text{ex}}(\nu_{\text{CP}}^{\text{low}}) - R_{1\rho,\text{CPMG}}^{\text{ex}}(\nu_{\text{CP}}^{\text{high}})$. Their experimental data are summarized in Figure 3A.^{25,45}

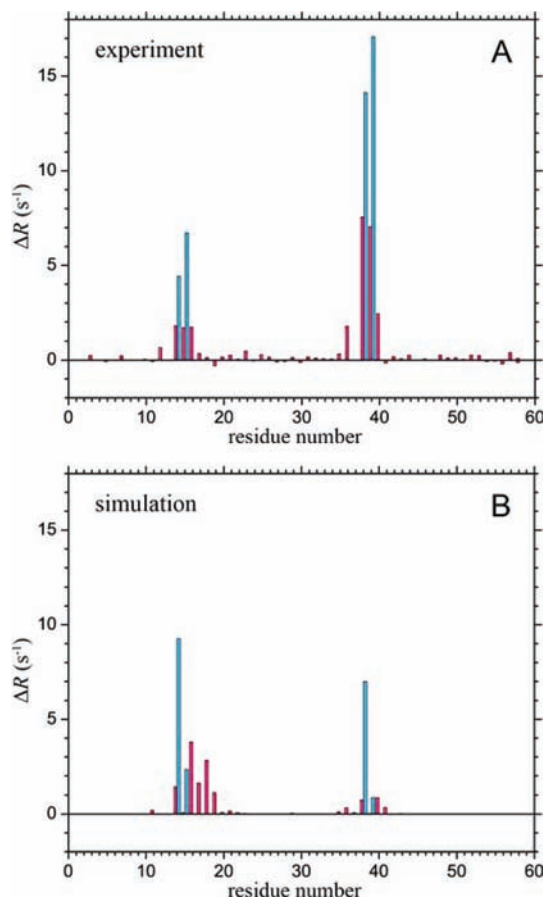


Figure 3. Experimental (A) and simulated (B) values of ^{15}N ΔR for BPTI at 300 K, spectrometer frequency 600 MHz. Pink bars represent $\Delta R = R_{1\rho,\text{CPMG}}^{\text{ex}}(50 \text{ Hz}) - R_{1\rho,\text{CPMG}}^{\text{ex}}(500 \text{ Hz})$; the experimental data are from Loria et al.⁴⁵ Cyan bars represent $\Delta R = R_{1\rho,\text{CPMG}}^{\text{ex}}(8 \text{ Hz}) - R_{1\rho,\text{CPMG}}^{\text{ex}}(\infty) \approx R_2^{\text{ex}}$; the experimental data are from Grey et al.,²⁵ reported only for C14, K15, C38, and R39.

To simulate ΔR , we first evaluated the correlation functions $g(\tau)$, eq 1, and then fitted the results with 3-exponential curves (red profiles in Figure 2). The fit parameters were subsequently used to predict $R_{1\rho,\text{CPMG}}^{\text{ex}}(\nu_{\text{CP}})$ via eq 5. Finally, the values of ΔR were calculated, Figure 3B (defined such as to imitate the experimental data).

Clearly, the pattern obtained for the simulated ΔR is similar to the one that has been observed experimentally, cf. parts B and A in Figure 3. In both cases the exchange broadening is attributable to isomerization of the C14–C38 disulfide bond. This is the only one of the three disulfide bonds in BPTI that shows the evidence of microsecond–millisecond dynamics (for comparison, we have generated the equivalent of Figures 1 and 2 for C5–C55, see Supporting Information).

DISCUSSION

Figure 3 directly connects the experimental results with a highly realistic motional model. In choosing this approach, we sidestep the problem of model selection, that is, we do not limit ourselves to 2- or 3-site exchange scheme involving certain representative conformations with distinct chemical shifts.

Table 1. Conformational Species of BPTI Classified According to the Isomeric State of the C14–C38 Disulfide Bridge

| | $\chi_1(\text{C14})$ | $\chi_2(\text{C14})$ | χ_3 | $\chi_2(\text{C38})$ | $\chi_1(\text{C38})$ | $d_{\text{C}\beta\text{C}\alpha}$ (Å) | population (%) | $E_{\chi}^{\text{AMBER}^b}$ (kcal/mol) | E_{χ}^c (kcal/mol) | occurrences in PDB ^d |
|---|----------------------|----------------------|----------|----------------------|----------------------|--|-------------------|---|----------------------------|------------------------------------|
| SPTI ⁴⁷ | −72° | 106° | 95° | −114° | 61° | 5.7 | | 4.9 | 4.3 | 214 |
| M (Grey et al. ²⁵) | −60° | 101° | 90° | −130° | 74° | | ~95 | 4.5 | 4.2 | 212 |
| M1 (MD) ^a | −66° | 100° | 95° | −126° | 73° | 5.8 | 25 | 4.8 | 4.3 | 215 |
| M2 (MD) ^a | −62° | 173° | 90° | −168° | 67° | 6.5 | 3 | 1.0 | 1.5 | 44 |
| M3 (MD) ^a | −63° | −78° | 89° | −172° | 54° | 5.0 | 6 | 1.2 | 1.6 | 277 |
| m _{C14} (Grey et al. ²⁵) | 50° | −110° | −135° | −70° | 66° | | ~1 ^e | 6.9 | 6.0 | 7 |
| m _{C14} (MD) ^a | 54° | −96° | −88° | −90° | 60° | 6.3 | 50 | 3.0 | 2.2 | 105 |
| m _{C38} (Grey et al. ²⁵) | −68° | 150° | −90° | 130° | −48° | | ~4 | 4.0 | 4.2 | 13 |
| m _{C38} (MD) ^a | −67° | 164° | −88° | 154° | −57° | 6.0 | 6 | 1.8 | 2.4 | 205 |

^aThe results represent averages over the respective clusters, see Supporting Information Figure S1. ^bTorsional angle energy of the disulfide bridge according to the Amber potential function.^{37,48} ^cTorsional angle energy based on MP2 calculations by Haworth et al.⁵⁷ Although E_{χ} represents only one term in the complex energy balance that governs the conformation of disulfide bridge, this parameter has proven to be informative—in particular, the energies $E_{\chi} > 6$ kcal/mol are often indicative of structural artifacts.⁵⁸ ^dThe parameters of disulfide bonds in crystallographic structures have been harvested from the Protein Data Bank using the program by Wong and Hogg.⁵⁹ A disulfide conformation from the PDB was considered to be equivalent to the given C14–C38 conformation if all five of their respective χ_i angles agreed to within 30°. ^eThe temperature-dependent CPMG data have been used to separate $p_A p_B$ from $\Delta\delta_{ab}$; this is normally not possible under the fast exchange conditions.²⁵

Instead, the entire bulk of information contained in the MD trajectory is translated into relaxation dispersion data, which can be then directly compared with the experimental results. In this sense the above treatment is more general than the commonly used exchange models.

With this consideration in mind, it is instructive to compare the current set of MD data with the existing model of microsecond–millisecond dynamics in BPTI.^{17,25,43–46} Grey et al. identified three relevant conformations of the disulfide bridge C14–C38, which can be conveniently classified on the basis of torsional angles $\chi_1(\text{C14})$ and $\chi_1(\text{C38})$, namely, M (−60°; 60°), m_{C14} (60°; 60°), and m_{C38} (−60°; −60°).²⁵ At 300 K, the line broadening in BPTI is caused by two distinct motional processes: (i) relatively fast exchange between the major state M and the minor state m_{C14} and (ii) much slower exchange involving M and m_{C38}:



The characteristic time of the first process was determined to be approximately 150 μs , whereas for the second process it is much longer, approximately 2 ms. Grey et al. also built structural models for M, m_{C14}, and m_{C38} and used these models to predict the chemical shift differences, $\Delta\delta_{ab}$. In what follows we discuss the relationship between these earlier NMR-based results and the current MD simulation.

Conformational States and Their Populations. The analysis of conformational space spanned by the torsional angles of the C14–C38 disulfide bridge, $\chi_1(\text{C14})$, $\chi_2(\text{C14})$, χ_3 , $\chi_2(\text{C38})$, and $\chi_1(\text{C38})$, reveals the presence of several clusters that can be directly related to the earlier models by Grey et al.,²⁵ see Supporting Information Figure S1. The structural statistics for these states is summarized in Table 1. The inspection of these data demonstrates that MD state M1 is nearly identical to the crystallographic conformation (structure SPTI⁴⁷) and to the energy-minimized model M described by Grey et al.²⁵ On the other hand, the minor state m_{C14} as observed in the MD trajectory is significantly different from the reconstruction by Grey et al.²⁵ Specifically, the MD simulation predicts a χ_3 angle of −88° which is close to the canonical value −90°; in contrast, the earlier model displays the angle of −135° which entails a significant energy penalty, approximately 4 kcal/mol.^{25,48,49} This discrepancy can be readily understood considering that the original model has been built from the

crystal structure SPTI by forcibly changing the angle $\chi_1(\text{C14})$ and energy-minimizing the resulting coordinates.²⁵ In reality, it appears that the conformational transition M → m_{C14} involves subtle but significant rearrangement of the two proximal loops, L1 (residues 9–18) and L2 (residues 36–40).⁵⁰ As a result of this long-range structural adjustment, the disulfide bridge is accommodated without undue strain on χ_3 . This leads us to suggest that the current MD representation is superior to the earlier energy-minimized model, as also confirmed by the PDB statistics (see Table 1). Similar observations can be made with respect to another minor state, m_{C38}.

While conformational species M1, m_{C14}, and m_{C38} are correctly identified in the current trajectory, the predicted populations lack quantitative agreement (see Table 1). In particular, the MD data predict that the m_{C14} state is populated at the level of 50%, whereas the experimentally determined population is approximately 1%. As pointed out by Shaw and co-workers,³² such differences between the simulations and the experiment are not unexpected. Indeed, the imbalance in populations corresponds to 2.3 kcal/mol difference in free energy, which is within the uncertainty range of the current force fields. One should also keep in mind that MD-derived populations are strongly affected by statistical uncertainty (the consequence of the fact that the trajectory is relatively short on the scale of forward and reverse exchange rates). Of note, the temperature-dependent relaxation dispersion study by Grey et al. suggests that m_{C14} has favorable enthalpy (consistent with the data in Table 1) and the M state is preferred on entropic grounds.²⁵ The fact that relative populations of m_{C14} and M are governed by entropy puts further emphasis on the problem of statistical sampling in the MD simulation.

Finally, the MD data display a greater variety of conformational species than previously envisaged. M2 and M3 states, differing mainly by the $\chi_2(\text{C14})$ angle, summarily account for 9% of the trajectory. The lifetime of these states reaches several microseconds (see below), which makes them potentially relevant for R^{ex} analyses. The conformations other than M1–M3, m_{C14}, and m_{C38} account for further 10% of the trajectory.

In the context of our study it appears that the conformational states of BPTI can be reasonably well classified according to the conformational state of the C14–C38 disulfide bridge

(see Supporting Information Figure S2). In general, however, one should reckon with the possibility that this limited parametrization may be insufficient. Recent analysis of the same BPTI trajectory identified many discrete substates sampled by the L1 and L2 loops (counting backbone conformations only).⁵⁰ The alternative loop conformations have been experimentally observed in some of the BPTI mutants.⁵¹ The conformational rearrangements involving L1 and L2 appear to be important in their own right since microsecond–millisecond dynamics in this region persists even when the disulfide bond is reduced.⁵² Finally, slow flips of the Y35 aromatic ring^{53,54} as well as the slow exchange of the resident water molecule that bridges C14 with C38^{55,56} point toward other potentially relevant degrees of freedom.

Exchange Rates. To derive the information about the characteristic time scale of exchange we turn to the correlation functions $g(\tau)$, as shown in Figure 2. The initial steep drop in $g(\tau)$ reflects fast fluctuations of chemical shifts and does not directly contribute to the R^{ex} relaxation – it is the slowly decaying component of $g(\tau)$ that is responsible for the R^{ex} effect. In the case of C14, the relevant decay time is 91 μs (cf. red curve in Figure 2A). For C38, the corresponding time constant is 60 μs . It should be emphasized that there is a large margin of uncertainty associated with the time constants τ_{ex}^i . Indeed, the analyzed 1-ms MD trajectory contains a relatively small number of transitions between the main conformational states (see below). The resulting correlation functions are clearly misshapen due to the poor statistics, Figure 2, which causes the uncertainty in the extracted τ_{ex}^i values. Considering all residues affected by exchange broadening, it is only safe to conclude that the relevant time constants fall in the range 10–100 μs , with the average value of ca. 50 μs . Note that τ_{ex}^i values of this magnitude correspond to the fast exchange regime, which ensures the validity of eqs 2–5.

Is it possible to link this time constant to the exchange between specific conformational species? The inspection of the MD trajectory reveals multiple transitions $M1 \rightleftharpoons M2 \rightleftharpoons M3$, as well as $m_{\text{C14}} \rightleftharpoons m_{\text{C38}}$. The majority of these transitions occur in a form of fast fluctuations, i.e. “in-and-out” excursions. To factor out this effect, we counted only those transitions that lead to a formation of relatively long-lived conformational states, i.e. those with a lifetime of at least 1 μs . We have also ignored for a moment the presence of the species other than $M1$ – $M3$, m_{C14} , and m_{C38} . Using these updated counting rules, we found 30 transitions between m_{C14} and m_{C38} , 20 transitions within the ($M1$, $M2$, $M3$) group, and 9 transitions between (m_{C14} , m_{C38}) and ($M1$, $M2$, $M3$). In principle, all of these transitions may be relevant for R^{ex} line broadening.

As described above, the MD trajectory is dominated by two conformational species, m_{C14} and $M1(M)$, and display the evidence of fast exchange process with characteristic correlation time of ca. 50 μs . These observations are broadly consistent with the experimental study by Grey et al., where the motional process with time constant of approximately 150 μs has been attributed to $m_{\text{C14}} \rightleftharpoons M$ exchange.²⁵ On the other hand, the situation with minor species m_{C38} raises additional questions. The experimental data suggest that m_{C38} is created in the time frame of ca. 40 ms and has a lifetime of approximately 2 ms.^{17,25} The MD trajectory, however, indicates that this state is created much more frequently and readily interconverts with m_{C14} . Several comments can be made in this regard.

- (i) The MD simulation suggests that the barrier separating m_{C38} from m_{C14} is relatively low. This is not unreasonable, given that the $m_{\text{C38}} \rightleftharpoons m_{\text{C14}}$ transition does not involve the crossing of the χ_3 barrier. The inspection of the original trajectory (sampled at a finer time interval of 250 ps) shows many “direct” transitions between m_{C38} and m_{C14} that do not involve any detectable intermediates.
- (ii) Nevertheless, it is possible that MD simulation underestimates by several kcal/mol the free-energy barrier that separates m_{C38} from m_{C14} and other conformational states. This would account for the observed differences in the simulated and experimental transition rates.
- (iii) Alternatively, it cannot be ruled out that the simulated m_{C38} state is, in fact, not identical to the experimentally observed conformer. The simulated state may represent a metastable intermediate, which converts to other conformational states rather than proceeding to a fully stable form of m_{C38} . One can further speculate that the long-lived form of m_{C38} is stabilized by conformational changes elsewhere in the protein; this form remains unsampled in the current 1-ms MD simulation.
- (iv) The MD results suggest that it may be advisable to generalize the previously proposed linear exchange scheme, eq 6, to a triangular scheme, allowing for a possibility of $m_{\text{C38}} \rightleftharpoons m_{\text{C14}}$ exchange. The assumption that $m_{\text{C38}} \rightarrow M$ and $m_{\text{C38}} \rightarrow m_{\text{C14}}$ transitions are (roughly) equiprobable does not seem to contradict any of the experimental observations.

Finally, it is worth noting that there is a certain amount of compensatory effect in the R^{ex} calculations, Figure 3B. The MD simulation overestimates the population factors $\sim p_{\text{A}}p_{\text{B}}$, while at the same time underestimating the exchange correlation times τ_{ex} . On balance, this translates into reasonable values of R^{ex} and ΔR . Note that the absence of the slow exchange mode in the MD simulation leads to lower-than-expected magnitude of ΔR for residues C38 and R39 (cf. Figure 3 and discussion by Grey et al.²⁵).

Chemical Shift Differences. As illustrated in Figure 1, individual MD frames can be used as an input for chemical shift calculation software (in our case, SHIFTX+). The results can be grouped according to the above classification ($M1$, $M2$, $M3$, m_{C14} , and m_{C38} ; see Supporting Information Figure S1) and the average chemical shift values can be computed for each state. Obviously, this approach assumes that conformational fluctuations within each of the five states are fast. To mimic as closely as possible the previous NMR-based study, we chose to compute the chemical shifts and chemical shift differences for the states $M1$, m_{C14} , and m_{C38} . The summary of these results is shown in Table 2. Also shown in Table 2 are $\Delta\delta_{\text{ab}}$ values computed by Grey and co-workers²⁵ using the earlier program

Table 2. ¹⁵N Chemical Shift Differences between the Conformational States $M1(M)$, m_{C14} , and m_{C38}

| res. | $\Delta\delta_{\text{M1},m_{\text{C14}}}$ (ppm) | | | $\Delta\delta_{\text{M1},m_{\text{C38}}}$ (ppm) | | |
|------|---|---------------------------|----------|---|---------------------------|----------|
| | exptl ²⁵ | prev. calcd ²⁵ | calcd MD | exptl ⁴³ | prev. calcd ²⁵ | calcd MD |
| C14 | 3.6 | 3.7 | 1.8 | −0.4 | 0.1 | 0.6 |
| K15 | 4.7 | 1.4 | −1.4 | −0.5 | −0.4 | −1.5 |
| C38 | 0.8 | 0.6 | −2.2 | −1.7 | −4.9 | −1.9 |
| R39 | 1.2 | 0.4 | −0.5 | −3.7 | −2.3 | −2.6 |

SHIFTX⁶⁰ and the energy-minimized structures of M (M1), m_{C14}, and m_{C38}.

There is a tentative correlation between the experimental and MD-derived values, as characterized by the rms deviation of 2.7 ppm (calculated for the shifts with known sign). This is appreciably worse than the previous calculations by Grey et al.,²⁵ which show the rms deviation of 2.0 ppm. Below we argue that one should not attach too much significance to this kind of difference.

- (i) The experimental $\Delta\delta_{M,mC14}$ values have been extracted from the relaxation dispersion data assuming linear 3-site exchange scheme, eq 6. This assumption may be restrictive. Indeed, as pointed out by Grey et al., the exchange rates derived from C14 data and K15 data systematically differ by a factor of approximately 1.4. "This difference, which is roughly the same size as the experimental uncertainties, may indicate that additional kinetic processes remain to be discovered in BPTI."²⁵ Of note, our calculations produce much better agreement with the experimental $\Delta\delta_{M,mC38}$ values, which have been extracted directly from the HSQC-type spectral map⁴³ (rms deviation 0.9 ppm).
- (ii) A major limiting factor is the relatively low accuracy of the chemical shift predictions. The error margin of δ^{15N} calculations using SHIFTX+ is reported to be 2.4 ppm.³³ A priori it is not clear whether this error is amplified or attenuated upon calculating the chemical shift differences $\Delta\delta_{ab} = \delta_a - \delta_b$. In any event, the error of this size can easily account for the observed discrepancy between the experimental and computational results, 2.7 ppm. It is worth noting that BPTI is a part of the training set of SHIFTX+ (crystallographic coordinates 5PTI,⁴⁷ chemical shifts from the TALOS database⁶¹). As a consequence, the program closely reproduces the experimental δ^{15N} shifts of BPTI when used with the structure 5PTI. Understandably, the agreement deteriorates when 5PTI is replaced with the MD ensemble M1. From our perspective, the presence of BPTI in the training set is not necessarily good since the program may be biased toward the dominant form of the protein at the expense of the "invisible" minor states. To investigate this matter further, we asked the authors of SHIFTX+ to prepare a customized version of the program where BPTI would be excluded from the training set. Using this customized program, we recalculated the data in Table 2. The results proved to be only slightly different, at most by 0.3 ppm (see Supporting Information Table S1).
- (iii) It is instructive to compare the results from SHIFTX+ with predictions from other similar programs. Toward this goal we have repeated the computations using the program SPARTA+,⁶² which achieves the same level of accuracy as SHIFTX+. Generally, the outcome proves to be very similar and the results in Table 2 are well reproduced. However, in some cases the predicted $\Delta\delta$ values differ by as much as 1.9 ppm (see Supporting Information Table S2).

In principle, the comparison between experimental and calculated values of $\Delta\delta$ opens the way for validation of structural models representing protein excited states. However, given all of the limitations described above and the small number of the available shifts, we conclude that the effective resolution of this method remains relatively low.

CONCLUSION

While the mechanism of μ s-ms exchange in proteins can often be intuited, in other cases the origin of R^{ex} remains unclear. This can be true even for well-characterized proteins, such as ubiquitin.⁶³ In this report we show that long MD simulations can shed light on the details of microsecond dynamics. The matching pattern of the simulated and experimental R^{ex} rates, such as shown in Figure 3, signals that MD data encode the information on relevant motional modes. We suggest that MD-based simulations should be used as a qualitative tool to guide the interpretation of experimental data. A similar view has been advocated in MD/NMR studies of fast protein dynamics.^{64,65}

It is anticipated that MD-based analysis will be used to identify relevant conformational species and map out the exchange network. For conformers of interest this approach can yield the estimates of chemical shift differences, $\Delta\delta_{ab}$. These estimates can be further included in the interpretation of the experimental relaxation dispersion data, thus allowing one to isolate and quantify the populations p_A, p_B —the feat that is otherwise impossible for microsecond time-scale exchange.

Given a high-quality MD simulation of sufficient length, one may be able to establish a three-way correspondence between the experimental data, MD simulation, and the judiciously chosen n -site exchange model. In particular, the n -site model can be used to fit the simulated relaxation dispersion data, thus allowing one to determine the rates of exchange between the individual sites. This method should successfully capture slow conformational dynamics, while filtering out fast fluctuations where protein fails to reach a stable conformational state (in this sense it is preferable to simply counting the transitions in the MD trajectory). As can be appreciated from the above discussion, even small, well-studied proteins such as BPTI can display an unexpectedly complex pattern of μ s motions. We expect that further progress in ultra-long MD simulations will shift the focus of dynamics studies from 2-site schemes to more general n -site models.

The Redfield theory treatment leading to eqs 1–5 is perfectly suited for studies of μ s exchange. It can also be easily expanded to treat multiquantum R^{ex} effect^{66–68} (see Appendix B, Supporting Information) and R^{ex} effect due to modulation of residual dipolar couplings.^{69,70} A more general approach, going beyond the Redfield formalism, can also be readily implemented. If one records a collection of long MD simulations, such data can be converted into chemical shift trajectories, $\delta_m^{(q)}(t)$. A sum over these trajectories,

$$FID(t) = \sum_{m,q} \exp(-i2\pi\nu_0\delta_m^{(q)}(t)t - (R_2^{m,dip} + R_2^{m,CSA})t) \quad (7)$$

would produce a FID-like signal corresponding to direct or indirect domain of an NMR experiment (here index m enumerates the atoms of interest, q enumerates the trajectories, and R_2^{dip} and R_2^{CSA} stand for conventional relaxation rates; if needed, the contributions from individual trajectories can be supplemented with statistical weights). The function $FID(t)$ encodes exchange-induced dephasing and thus faithfully recreates the broadened spectrum (given that the number of trajectories is sufficiently large). Along the same lines one can simulate the outcome of the spin-lock and CPMG measurements without being restricted to the Redfield fast-exchange limit. More exotic variants of the relaxation dispersion experiments^{41,71,72} can be treated as well.

To successfully simulate R^{ex} effects, the MD trajectory needs to be sufficiently long on the time scale of both forward, k_1 , and reverse, k_{-1} , exchange rates. In practice this means that only fast forms of exchange, 10–100 μs , are amenable to such studies and that the populations of the exchanging species should not be too skewed. While at present only a handful of trajectories generated at D. E. Shaw Research may satisfy these requirements, in the future the scope of opportunities will grow. The development of faster chips, better networks, and more efficient MD algorithms will make 1 ms trajectories more routine and bring about longer trajectories.^{73–75} Such ultra-long MD simulations used in conjunction with experimental NMR data will significantly enhance the potential for studies of μs -ms dynamics, offering unique insight in many functionally important forms of motion.

■ ASSOCIATED CONTENT

■ Supporting Information

Appendix with derivation of eqs 4 and 5; scatter plot illustrating the conformational states of the disulfide bond C14–C38; equivalent of Figure 1 colored according to the isomeric state of C14–C38; equivalents of Table 2 for the calculations using Sparta+ and the customized version of SHIFTX+; equivalents of Figures 1 and 2 for the disulfide bond C5–C55; appendix with discussion of multiquantum relaxation rates $\Delta R_{\text{MQ}}^{\text{ex}}$; simulated correlation functions $g(\tau)$ for all nonproline residues in BPTI. Complete ref 14. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

nikolai@purdue.edu

Present Addresses

[†]Chemistry Department, University of Oulu, 90014 Oulu, Finland.

[‡]Monitoring Ltd., St. Petersburg 190013, Russia.

■ ACKNOWLEDGMENTS

This work was supported by the National Science Foundation (MCB 0445643). We acknowledge the profound contribution of the scientists at D. E. Shaw Research who shared with us their MD simulation data. We are indebted to L. E. Kay and K. Lindorff-Larsen for critically reading this manuscript and to B. Han for his help with SHIFTX+.

■ REFERENCES

- (1) Baldwin, A. J.; Kay, L. E. *Nat. Chem. Biol.* **2009**, *5*, 808.
- (2) Korzhnev, D. M.; Kay, L. E. *Acc. Chem. Res.* **2008**, *41*, 442.
- (3) Boehr, D. D.; Dyson, H. J.; Wright, P. E. *Chem. Rev.* **2006**, *106*, 3055.
- (4) Henzler-Wildman, K.; Kern, D. *Nature* **2007**, *450*, 964.
- (5) Mulder, F. A. A.; Mittermaier, A.; Hon, B.; Dahlquist, F. W.; Kay, L. E. *Nat. Struct. Biol.* **2001**, *8*, 932.
- (6) Palmer, A. G.; Kroenke, C. D.; Loria, J. P. *Methods Enzymol.* **2001**, *339*, 204.
- (7) Akke, M. *Curr. Opin. Struct. Biol.* **2002**, *12*, 642.
- (8) Vallurupalli, P.; Hansen, D. F.; Kay, L. E. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 11766.
- (9) Orekhov, V. Y.; Pervushin, K. V.; Arseniev, A. S. *Eur. J. Biochem.* **1994**, *219*, 887.
- (10) Akke, M.; Palmer, A. G. *J. Am. Chem. Soc.* **1996**, *118*, 911.
- (11) Ishima, R.; Wingfield, P. T.; Stahl, S. J.; Kaufman, J. D.; Torchia, D. A. *J. Am. Chem. Soc.* **1998**, *120*, 10534.
- (12) Mulder, F. A. A.; van Tilborg, P. J. A.; Kaptein, R.; Boelens, R. *J. Biomol. NMR* **1999**, *13*, 275.
- (13) Cavalli, A.; Salvatella, X.; Dobson, C. M.; Vendruscolo, M. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 9615.
- (14) Shen, Y.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 4685.
- (15) Shen, Y.; Vernon, R.; Baker, D.; Bax, A. *J. Biomol. NMR* **2009**, *43*, 63.
- (16) Wishart, D. S.; Arndt, D.; Berjanskii, M.; Tang, P.; Zhou, J.; Lin, G. *Nucleic Acids Res.* **2008**, *36*, W496.
- (17) Millet, O.; Loria, J. P.; Kroenke, C. D.; Pons, M.; Palmer, A. G. *J. Am. Chem. Soc.* **2000**, *122*, 2867.
- (18) Korzhnev, D. M.; Salvatella, X.; Vendruscolo, M.; Di Nardo, A. A.; Davidson, A. R.; Dobson, C. M.; Kay, L. E. *Nature* **2004**, *430*, 586.
- (19) Grey, M. J.; Tang, Y. F.; Alexov, E.; McKnight, C. J.; Raleigh, D. P.; Palmer, A. G. *J. Mol. Biol.* **2006**, *355*, 1078.
- (20) Sugase, K.; Dyson, H. J.; Wright, P. E. *Nature* **2007**, *447*, 1021.
- (21) Korzhnev, D. M.; Religa, T. L.; Banachewicz, W.; Fersht, A. R.; Kay, L. E. *Science* **2010**, *329*, 1312.
- (22) Korzhnev, D. M.; Vernon, R. M.; Religa, T. L.; Hansen, A. L.; Baker, D.; Fersht, A. R.; Kay, L. E. *J. Am. Chem. Soc.* **2011**, *133*, 10974.
- (23) Ishima, R.; Freedberg, D. I.; Wang, Y. X.; Louis, J. M.; Torchia, D. A. *Structure* **1999**, *7*, 1047.
- (24) Volkman, B. F.; Lipson, D.; Wemmer, D. E.; Kern, D. *Science* **2001**, *291*, 2429.
- (25) Grey, M. J.; Wang, C. Y.; Palmer, A. G. *J. Am. Chem. Soc.* **2003**, *125*, 14324.
- (26) Wolf-Watz, M.; Thai, V.; Henzler-Wildman, K.; Hadjipavlou, G.; Eisenmesser, E. Z.; Kern, D. *Nat. Struct. Mol. Biol.* **2004**, *11*, 945.
- (27) Massi, F.; Wang, C. Y.; Palmer, A. G. *Biochemistry* **2006**, *45*, 10787.
- (28) Evenäs, J.; Forsén, S.; Malmendal, A.; Akke, M. *J. Mol. Biol.* **1999**, *289*, 603.
- (29) Mittag, T.; Schaffhausen, B.; Günther, U. L. *Biochemistry* **2003**, *42*, 11128.
- (30) Hass, M. A. S.; Thuesen, M. H.; Christensen, H. E. M.; Led, J. J. *J. Am. Chem. Soc.* **2004**, *126*, 753.
- (31) Morcos, F.; Chatterjee, S.; McClendon, C. L.; Brenner, P. R.; Lopez-Rendon, R.; Zintsmaster, J.; Ercsey-Ravasz, M.; Sweet, C. R.; Jacobson, M. P.; Peng, J. W.; Izaguirre, J. A. *PLoS Comput. Biol.* **2010**, *6*, e1001015.
- (32) Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y. B.; Wriggers, W. *Science* **2010**, *330*, 341.
- (33) Han, B.; Liu, Y. F.; Ginzinger, S. W.; Wishart, D. S. *J. Biomol. NMR* **2011**, *50*, 43.
- (34) de Dios, A. C.; Pearson, J. G.; Oldfield, E. *Science* **1993**, *260*, 1491.
- (35) Li, D. W.; Bruschweiler, R. *J. Phys. Chem. Lett.* **2010**, *1*, 246.
- (36) Markwick, P. R. L.; Cervantes, C. F.; Abel, B. L.; Komives, E. A.; Blackledge, M.; McCammon, J. A. *J. Am. Chem. Soc.* **2010**, *132*, 1220.
- (37) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins* **2006**, *65*, 712.
- (38) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. *Proteins* **2010**, *78*, 1950.
- (39) Abragam, A. *The Principles of Nuclear Magnetism*; Clarendon Press: Oxford, 1961.
- (40) Deverell, C.; Morgan, R. E.; Strange, J. H. *Mol. Phys.* **1970**, *18*, 553.
- (41) Podkorytov, I. S.; Skrynnikov, N. R. *J. Magn. Reson.* **2004**, *169*, 164.
- (42) Luz, Z.; Meiboom, S. *J. Chem. Phys.* **1963**, *39*, 366.
- (43) Otting, G.; Liepinsh, E.; Wuthrich, K. *Biochemistry* **1993**, *32*, 3571.
- (44) Szyperski, T.; Lugnbuhl, P.; Otting, G.; Guntert, P.; Wuthrich, K. *J. Biomol. NMR* **1993**, *3*, 151.
- (45) Loria, J. P.; Rance, M.; Palmer, A. G. *J. Biomol. NMR* **1999**, *15*, 151.
- (46) Massi, F.; Johnson, E.; Wang, C. Y.; Rance, M.; Palmer, A. G. *J. Am. Chem. Soc.* **2004**, *126*, 2247.
- (47) Wlodawer, A.; Walter, J.; Huber, R.; Sjolin, L. *J. Mol. Biol.* **1984**, *180*, 301.

- (48) Katz, B. A.; Kossiakoff, A. *J. Biol. Chem.* **1986**, *261*, 15480.
- (49) Jiao, D.; Barfield, M.; Combariza, J. E.; Hruby, V. J. *J. Am. Chem. Soc.* **1992**, *114*, 3639.
- (50) Long, D.; Bruschiweiler, R. *J. Am. Chem. Soc.* **2011**, *133*, 18999.
- (51) Housset, D.; Kim, K. S.; Fuchs, J.; Woodward, C.; Wlodawer, A. *J. Mol. Biol.* **1991**, *220*, 757.
- (52) Beeser, S. A.; Oas, T. G.; Goldenberg, D. P. *J. Mol. Biol.* **1998**, *284*, 1581.
- (53) Wagner, G.; DeMarco, A.; Wuthrich, K. *Biophys. Struct. Mech.* **1976**, *2*, 139.
- (54) Battiste, J. L.; Li, R. H.; Woodward, C. *Biochemistry* **2002**, *41*, 2237.
- (55) Otting, G.; Liepinsh, E.; Wuthrich, K. *J. Am. Chem. Soc.* **1991**, *113*, 4363.
- (56) Denisov, V. P.; Peters, J.; Horlein, H. D.; Halle, B. *Biochemistry* **2004**, *43*, 12020.
- (57) Haworth, N. L.; Gready, J. E.; George, R. A.; Wouters, M. A. *Mol. Simul.* **2007**, *33*, 475.
- (58) Haworth, N. L.; Liu, J. Y.; Fan, S. W.; Gready, J. E.; Wouters, M. *Aust. J. Chem.* **2010**, *63*, 379.
- (59) Wong, J. W. H.; Hogg, P. J. *J. Thromb. Haemost.* **2010**, *8*, 2345.
- (60) Neal, S.; Nip, A. M.; Zhang, H. Y.; Wishart, D. S. *J. Biomol. NMR* **2003**, *26*, 215.
- (61) Cornilescu, G.; Delaglio, F.; Bax, A. *J. Biomol. NMR* **1999**, *13*, 289.
- (62) Shen, Y.; Bax, A. *J. Biomol. NMR* **2010**, *48*, 13.
- (63) Sidhu, A.; Surolia, A.; Robertson, A. D.; Sundd, M. *J. Mol. Biol.* **2011**, *411*, 1037.
- (64) Bremi, T.; Brüschweiler, R.; Ernst, R. R. *J. Am. Chem. Soc.* **1997**, *119*, 4272.
- (65) Lipari, G.; Szabo, A.; Levy, R. M. *Nature* **1982**, *300*, 197.
- (66) Kloiber, K.; Konrat, R. *J. Biomol. NMR* **2000**, *18*, 33.
- (67) Wang, C. Y.; Palmer, A. G. *J. Biomol. NMR* **2002**, *24*, 263.
- (68) Orekhov, V. Y.; Korzhnev, D. M.; Kay, L. E. *J. Am. Chem. Soc.* **2004**, *126*, 1886.
- (69) Igumenova, T. I.; Brath, U.; Akke, M.; Palmer, A. G. *J. Am. Chem. Soc.* **2007**, *129*, 13396.
- (70) Vallurupalli, P.; Hansen, D. F.; Stollar, E.; Meirovitch, E.; Kay, L. E. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 18473.
- (71) Konrat, R.; Tollinger, M. *J. Biomol. NMR* **1999**, *13*, 213.
- (72) Ulzega, S.; Salvi, N.; Segawa, T. F.; Ferrage, F.; Bodenhausen, G. *ChemPhysChem* **2011**, *12*, 333.
- (73) Klepeis, J. L.; Lindorff-Larsen, K.; Dror, R. O.; Shaw, D. E. *Curr. Opin. Struct. Biol.* **2009**, *19*, 120.
- (74) Vendruscolo, M.; Dobson, C. M. *Curr. Biol.* **2011**, *21*, R68.
- (75) Voelz, V. A.; Singh, V. R.; Wedemeyer, W. J.; Lapidus, L. J.; Pande, V. S. *J. Am. Chem. Soc.* **2010**, *132*, 4702.