JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

Dependence of Internal Friction on Folding Mechanism

Wenwei Zheng,[†] David De Sancho,[‡,§,⊥] Travis Hoppe,[†] and Robert B. Best*[,†]

[†]Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892, United States

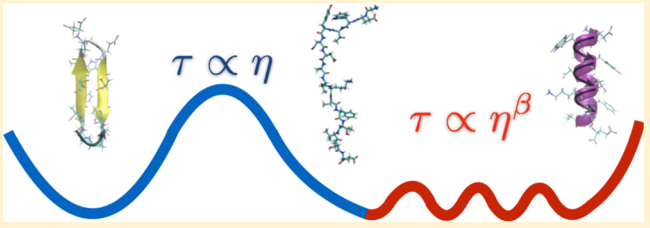
[‡]Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom

[§]CIC nanoGUNE, 20018 Donostia−San Sebastián, Spain

[⊥]IKERBASQUE, Basque Foundation for Science, Maria Diaz de Haro 3, 48013 Bilbao, Spain

Supporting Information

ABSTRACT: An outstanding challenge in protein folding is understanding the origin of "internal friction" in folding dynamics, experimentally identified from the dependence of folding rates on solvent viscosity. A possible origin suggested by simulation is the crossing of local torsion barriers. However, it was unclear why internal friction varied from protein to protein or for different folding barriers of the same protein. Using all-atom simulations with variable solvent viscosity, in conjunction with transition-path sampling to obtain reaction rates and analysis via Markov state models, we are able to determine the internal friction in the folding of several peptides and miniproteins. In agreement with experiment, we find that the folding events with greatest internal friction are those that mainly involve helix formation, while hairpin formation exhibits little or no evidence of friction. Via a careful analysis of folding transition paths, we show that internal friction arises when torsion angle changes are an important part of the folding mechanism near the folding free energy barrier. These results suggest an explanation for the variation of internal friction effects from protein to protein and across the energy landscape of the same protein.



A large body of theoretical and experimental work has been devoted to characterizing the free energy landscape for protein folding, and has broadly converged toward a description based on a biased energy landscape ("funnel"), selected by evolution for efficient folding.[1−3] Folding models with an energetic bias toward the native protein structure, resulting in a funneled landscape,[4] can be used to explain trends in protein folding rates with protein topology,[5] the effects of mutations on folding rates (via ϕ-values),[6,7] coupled folding-binding,[8] the response of proteins to a pulling force,[9] and many other examples. An area of growing interest is the manner in which local features of the energy landscape may influence the folding dynamics. In the energy landscape picture, folding can be viewed as diffusion on a low-dimensional free energy surface, in which local features enter in the diffusion coefficient for moving along the folding coordinate(s), particularly near the top of the folding barrier.[10] Projection onto a low-dimensional energy surface puts protein folding on the same footing as other chemical reactions in solution, which can be described by Kramers theory.[11] Clearly, determining local properties of the energy landscape by experiment is very challenging, but recent developments in experimental techniques as well as the discovery of several interesting prototypical examples have yielded new insights into this process. These studies are primarily concerned with the phenomenon of "internal friction", which is empirically identified from the change of rates as solvent viscosity is varied (by adding extrinsic viscogenic agents).[12,13] The rationale for this approach is that if the solvent were the dominant source of friction, folding times τ (or relaxation times in the unfolded state) would be proportional to solvent viscosity η, following Stokes' law. However, if there were a significant contribution from the protein ("internal friction") to folding dynamics, there may be a deviation from this anticipated first power dependence of relaxation times on viscosity.

Recent experimental work has identified internal friction in protein unfolded states,[14] folding rates,[15−17] folding transition-path times[18] and in the molecular phase in temperature-jump experiments, also related to barrier transit times.[19] The magnitude of the effect does, however, vary with the protein studied as well as control variables such as temperature and denaturant concentration.[14,15,18] However, in earlier studies on other proteins there was no deviation from a first power dependence of folding times on viscosity, outside of experimental errors, suggesting zero internal friction.[20−22] Interpreting these results at a molecular level, and explaining the differences between proteins, is clearly challenging. Molecular simulations have the potential to provide detailed insights into biomolecular dynamics,[23] including the origin of internal friction. Simulations can also circumvent one of the key challenges in experiment, namely the frequent stabilization of proteins by viscogens. The change in stability must usually be corrected for by adding

Received: November 17, 2014
Published: February 27, 2015

DOI: 10.1021/ja511609u
J. Am. Chem. Soc. 2015, 137, 3283−3290

compensating amounts of chemical denaturant in experiment.[15,17,18] Only in a few cases[16,21,24] have conditions been found where a viscogen had no effect on stability.

In simulations with implicit solvent, friction can be varied directly without altering the free energy landscape by using Langevin dynamics,[25,26] but this uses a simplified memoryless model for the friction. A more detailed approach when comparing with experiment is to vary the solvent mass in explicit solvent simulations, altering the viscosity by effectively rescaling time,[27,28] allowing for solvent memory effects to be included. The free energy landscape is again unchanged owing to the statistical independence of momenta and positions in classical statistical mechanics.[29] Simulation studies using mass-scaling have found evidence for internal friction in peptide dynamics, protein folding, and the reconfiguration of unfolded proteins,[28,30,31] similar to that seen in experiments. A common source of apparent internal friction has been identified as the crossing of sharp local barriers in the energy landscape, particularly dihedral barriers,[30−32] as originally proposed by Kuhn in the context of polymer dynamics.[33] This appears to give rise to a deviation from first power relaxation time−viscosity dependence because of the effect of the solvent relaxation spectrum on the crossing of these local barriers.[30,32]

Despite the identification of internal friction in these all-atom simulation studies, variation of internal friction from one protein to another has not been fully accounted for, in particular, the apparent lack of internal friction for some proteins. It has been predicted that differences in folding mechanism will make some proteins more sensitive to solvent friction than others:[32] it is expected that proteins whose mechanism involves large scale movements of the chain through the solvent near the transition state might exhibit little internal friction, in contrast to cases where torsion angle changes are more important, such as helix formation. Indeed, many of the examples in which internal friction has been identified, in both experiment and simulation, are small helical proteins,[15−18] while many of the cases lacking internal friction are either all-β[21,34] or α/β[22] proteins. This would be expected if the mechanism were determined by the native topology, as has been found in many cases,[7] although there are certainly exceptions, e.g., mutants of lambda repressor may fold downhill with a single barrier or via a folding intermediate,[35,36] or experimentally observed intermediates may result from non-native interactions.[37,38] There is also a second mechanism by which helical proteins may tend to exhibit more internal friction, since many of those where internal friction has been detected are small, fast-folding proteins. A lower (or vanishing) folding free energy barrier will result in the folding rate being sensitive to dynamics on a broader region of the energy landscape, and therefore more likely to exhibit internal friction. Indeed, when the barrier is reduced sufficiently, there is even evidence for internal friction in the all-β protein Fip35.[19]

In this paper we investigate the relationship between the degree of internal friction for a given protein and its native structure. We start with the folding of the GB1-hairpin, the smallest system that does not show any internal friction in experiment.[24] We study its folding in explicit solvent, but we bias the energy landscape such that different native states are favored: either a hairpin (as in the original GB1 protein from which it comes[39]) or a helix, which it can also form in the context of tertiary structure.[40] This bias reduces the importance of non-native interactions in the simulations, allowing us to assess the influence of folding mechanism and native topology on internal friction effects. On the other hand it makes it more difficult for us to comment directly on the possible influence of non-native interactions. We use a transition-path sampling scheme to determine accurately the folding rate for the hairpin, and we find no evidence for a deviation from a first power dependence of folding time on viscosity, in agreement with experiment. By contrast, when the native state is a helix, there is internal friction, similar to what has been observed in both experiment and simulation on other model helix systems.[24,28,30] Lastly, we consider how internal friction may vary across the energy landscape (recently demonstrated for spectrin domains[41]). We find that the folding of a model of Trp cage exhibits two free energy barriers, a smaller barrier for α-helix formation and a major barrier for formation of the tertiary hairpin structure. Using transition-path sampling, we find that the two barriers have different apparent internal friction, with clear internal friction for the first barrier and weaker evidence of internal friction for the second barrier. For both GB1 and Trp cage, we find that the sensitivity of the folding rates to internal friction depends on the role of torsion angle isomerization in the folding mechanism, consistent with the previous work that torsion transition is the molecular origin of internal friction;[30] the expectation that lower folding barriers will also be sensitive to internal friction is also consistent with our results. Overall, our results support a dominant role for folding mechanism in determining the internal friction in the folding process, by determining the role of torsion angle transitions in the folding mechanism.

METHODS

Molecular Simulation Methods. Molecular dynamics was performed in GROMACS 4.5.3 and 4.6.5[42] with a leapfrog integrator. A velocity rescaling thermostat[43] is used for all cases involving water-mass scaling, except for replica exchange molecular dynamics (REMD)[44] in which Langevin dynamics is used. Pressure is controlled with the Parrinello−Rahman method.[45] The force field in all cases is Amber ff03ws, a derivative of Amber ff03[46] with a backbone modification to match the population of the helical states[47] and a scaling of protein and water interaction to capture the dimensions of unfolded structures.[48] The sequence of 16-residue GB1 is cut from residues 41 to 56 of the full length GB1 protein (PDB: 1PGB[39]); the hairpin native structure is taken from the same PDB structure. The native structure of Trp cage is taken from PDB: 1L2Y.[49] All simulations except REMD are performed at 350 K. The biased all-atom model is constructed by modifying the Lennard−Jones parameters of the original force field. Protein−protein atom pairs separated by less than 6 Å in the native configuration were stabilized by multiplying the ϵ of their LJ potential by 1.15, while the remaining protein−protein Lennard−Jones interactions were modified by using a small $\epsilon = 10^{-5}$ kJ mol^{-1} and defining σ such that the new and old potentials coincide at an energy of 2.5 kJ mol^{-1}. In doing so, the attractive part of the non-native nonbonded interactions was removed, while the effective radius of the atom was preserved. We keep all the other terms of the original force field.

Folding Relaxation Time from Transition-Path Sampling (TPS). Rate coefficients in Trp cage and hairpin-biased GB1 are calculated by sampling transition paths using the method proposed by Hummer.[50] Briefly, an ensemble of structures on a dividing surface Q_0 near the top of the apparent barrier in the REMD-derived free energy surface $F(Q)$, for the fraction of native contacts Q, are generated by umbrella sampling. Structures within $\Delta Q = |Q - Q_0| \leq 0.001$ of the dividing surface are selected from this ensemble. "Forward" and "reverse" trajectory pairs were generated by reversing the sign of the initial Maxwell−Boltzmann velocities, and each trajectory was terminated once it reached a region of Q defined as a stable state. Successful transition paths were those where the forward and reverse parts terminated in different end states. For simulation convenience, we discarded the trajectories longer than a chosen maximum time. However, this maximum trajectory length was selected to be long

DOI: 10.1021/ja511609u
J. Am. Chem. Soc. 2015, 137, 3283−3290

enough so that over 99% of the trajectories are terminated in the stable states. The boundaries of Q used to define the stable states, the maximum length of each trajectory and the probability of finished trajectories are all shown in Table S1 (Supporting Information).

By considering the time spent by the system in each stable state, and on transition paths, the rate coefficients for two-state kinetics can be expressed as

$$k' \equiv \frac{2}{k_1^{-1} + k_2^{-1}} = p_{eq}(Q_0)\langle \theta_{TP} w \rangle|_{Q_0, p_{eq}} \quad (1)$$

where k_1 and k_2 are the rate coefficients for the forward and backward transitions between the two states, $p_{eq}(Q_0)$ is the equilibrium probability density at Q_0, θ_{TP} takes a value of one for trajectory pairs forming a transition path and zero otherwise, and the w is the inverse of the time spent at the dividing surface. The weight w is used to correct the bias toward transition paths that cross the dividing surface many times, arising from the TPS method, and is also used as the relative weight of each trajectory when calculating the average transition path time in different solvent viscosities. We refer interested readers to the Supporting Information and the original ref 50 for more details. A particular advantage of this method is that the error in $p_{eq}(Q_0)$, arguably the most challenging quantity to determine accurately from equilibrium simulation, does not affect relative folding rates (at different viscosities), thus improving the statistics of the rate calculation.

For comparison to the Markov state model (MSM, see next section), we report the relaxation time τ, defined as the slowest relaxation of the system when it is perturbed from equilibrium. For a two-state approximation, τ is inverse of the sum of the folding and unfolding rate coefficients. Therefore, the relation between τ and k' can be written as

$$\tau = \frac{2P_1(1 - P_1)}{k'} \quad (2)$$

where P_1 is the equilibrium probability of the first state.

500 pairs of forward and backward shooting were performed for each barrier of Trp cage and hairpin-biased GB1 by using the fraction of native contacts Q as a reaction coordinate. Further details are given in the Supporting Information.

Markov State Model. To complement the analysis of the simulations we use a Markov state model (MSM).[51] This analysis is particularly important in the case of the helix-biased model, since the dynamics of this system does not conform to a two-state picture. First, the simulation data at each value of the viscosity were discretized into microscopic states using a hydrogen bond or a torsion angle description. Assignment of the transitions was performed using transition-based assignment.[52] In the case of the torsion angle discretization, typically hundreds to thousands of states are populated during the simulation. To simplify the analysis and increase the statistics of observed transitions, at each value of the viscosity we use the most populated set of states, accounting for 99% of the simulation data, and discard any states that are not part of the largest strongly connected set. The transition count matrix $\mathbf{N}(\Delta t)$, was then constructed by counting the number of transitions between every pair of states, after a lag time Δt. Then the column transition probability matrix $\mathbf{T}(\Delta t)$, was computed using the maximum likelihood estimator[53]

$$T_{ji}(\Delta t) = N_{ji}(\Delta t) / \sum_k N_{ki}(\Delta t) \quad (3)$$

The relaxation times were calculated from the eigenvalues λ_i of $\mathbf{T}$ as $\tau_i = -\Delta t/\ln(\lambda_i)$. The errors of the relaxation times were derived from a bootstrap analysis. The lag times for the MSM were chosen so that the relaxation times were well converged within error (see Figure S1).

RESULTS

There are many sources of roughness in protein energy landscapes that determine the diffusion coefficient on the folding coordinate,[54] including the making and breaking of both native and non-native contacts, and local barriers to chain rearrangements. However, as we are specifically interested in the effect of native topology rather than other factors, we have designed our study to focus on this effect. Thus, we investigate the peptide with the sequence from the GB1 hairpin, using all-atom simulations with explicit water, using a recent atomistic force field.[48] We use a weak energetic bias to favor one of two possible structures: either (i) the GB1 hairpin structure ("hairpin-biased"), defined by the crystal structure of the complete protein, or (ii) an α-helix ("helix-biased"). In each case, this is achieved by reducing the Lennard–Jones (LJ) pair interaction for non-native contacts (relative to a given structure) to a short-range repulsion. Non-native interactions are not eliminated, as non-native hydrogen bonds and salt bridges are still possible, and the hydrophobic effect is present. In the absence of attractive non-native LJ interactions, however, non-native interactions are weaker and the energy landscape is naturally "funneled" toward a given native structure. In this way we can more directly probe the effect of native topology, rather than other factors, in determining internal friction. However, it also means that, for most of the systems studied, we can only comment indirectly on the possible influence of non-native interactions.

Absence of Internal Friction in GB1 Hairpin. First we focus on the hairpin-biased GB1, which is the smallest folded peptide lacking internal friction in experiment.[24] Its folding relaxation time is ~3 μs in T-jump experiments[55] and ~13 μs in all-atom simulations[56] at 300 K. Even at 350 K, simulations still yielded a microsecond relaxation time.[56] Therefore, using brute force simulation to predict the rate coefficient with sufficient accuracy using different solvent viscosities would require orders of magnitude longer simulations than are currently feasible. We used replica exchange molecular dynamics (REMD)[44] to get an estimate of the free energy surface of the hairpin-biased GB1. Using the fraction of native contacts (Q) to project the free energy, we obtain a two-state landscape with a barrier of about ~4 kcal mol^{-1}, shown in Figure 1. A similar barrier was estimated by fitting an Ising-like model to experimental data for this system.[57] With the hairpin-biased LJ terms, the misfolded hairpin found in the previous work[56,58] is not populated. In order to calculate rate coefficients, we use a variant of transition-path sampling (TPS),[59] in which transition paths are harvested by shooting trajectories from a well-chosen dividing surface.[50] The main idea behind TPS is to focus sampling on the rare transitions

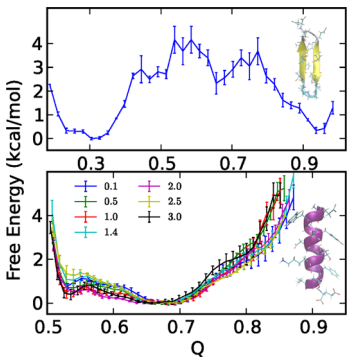


Figure 1. Free energy surfaces of GB1 projected onto the fraction of native contacts Q at 350 K. Top: hairpin-biased model from REMD. Bottom: helix-biased model at different viscosities. Different colors show the free energy surfaces determined independently at different solvent viscosities (η/η_0), confirming the similarity of sampling when changing viscosities. The native structures of the two models are shown on the right-hand side.

DOI: 10.1021/ja511609u
J. Am. Chem. Soc. 2015, 137, 3283–3290

from unfolded to folded, and to avoid the long periods spent wandering in each of these free energy basins between transitions. We can thus sample orders of magnitude more folding events than in brute force simulation, greatly reducing the statistical error in the rate estimate. As in experiment, we determine internal friction based on the dependence of folding relaxation times on solvent viscosity, which we vary in the simulations by scaling the water mass.[27]

The viscosity-dependent two-state folding relaxation times τ of the hairpin-biased GB1 are shown in Figure 2 (the folding

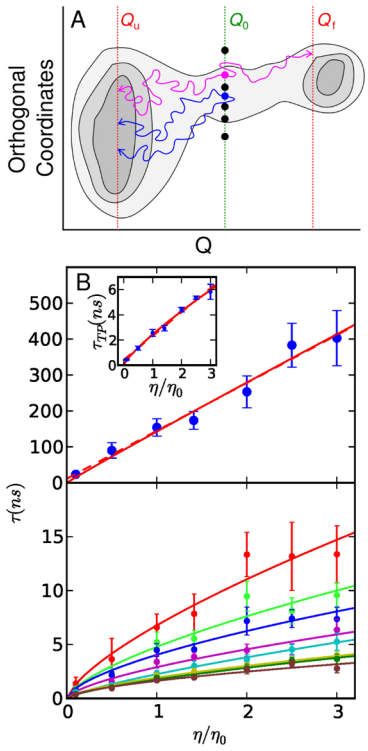


Figure 2. Viscosity dependence of the relaxation times of GB1. (A) Schematic illustration of transition-path sampling scheme used to determine rates for the hairpin-biased GB1: configurations were randomly selected from an equilibrium distribution at the chosen dividing surface (Q_0 = 0.65); pairs of "forward" trajectories initiated from these configurations with momenta $\mathbf{p}_{fwd}$ chosen from a Maxwell–Boltzmann distribution and "reverse" trajectories with initial momenta $\mathbf{p}_{rev} = -\mathbf{p}_{fwd}$ form a transition path if they end in opposite free energy minima, defined by Q_u = 0.35 and Q_f = 0.85. The magenta and blue trajectory pairs are examples of successful and unsuccessful transition paths, respectively. "Orthogonal coordinates" are illustrative of collective coordinates orthogonal to Q. (B) Folding relaxation times τ for the hairpin-biased model from transition path sampling. The inset shows the average transition path time, τ_{TP} for this model. The full distributions of the transition path times are in Figure S2. Relaxation times of different modes of the Markov state model for the helix-biased GB1. Error bars are derived from a bootstrap analysis. Solid lines correspond to power-law fits to the data. For the hairpin, the broken line is a linear fit.

relaxation time is defined here as the slowest relaxation time when the system is perturbed from equilibrium, i.e., the inverse of the sum of the folding and unfolding rates for a two-state folder). Because changing the solvent mass does not alter the free energy surface, folding and unfolding times will have the same dependence on viscosity as the relaxation time. Fits to the two most commonly used empirical relations, a linear equation $\tau = \tau_0 + b(\eta/\eta_0)$ or power-law $\tau = A(\eta/\eta_0)^{\beta}$, are almost superimposable (η_0 is the viscosity of the water model with normal masses), with the power-law exponent β being close to unity (0.97 ± 0.11) and the normalized intercept of linear fit $\tau_0/\tau(\eta_0)$ being close to zero (0.08 ± 0.06; Table 1). Thus, the folding and unfolding times of

Table 1. Coefficient β of Power-Law Fits and Normalized Intercept $\tau_0/\tau(\eta_0)$ of Linear Fits[a]

	β	$\tau_0/\tau(\eta_0)$
GB1		
hairpin-biased τ	0.97 (0.11)	0.08 (0.06)
hairpin-biased τ_{TP}	0.82 (0.05)	0.18 (0.03)
helix-biased λ_1	0.71 (0.11)	0.27 (0.08)
helix-biased λ_1–λ_9[b]	0.67 (0.02)	0.31 (0.02)
Jas et al. 2001[24] 20 °C	1.07 (0.25)	
Trp cage		
Barrier A τ	0.66 (0.04)	0.30 (0.04)
Barrier A τ_{TP}	0.73 (0.04)	0.25 (0.03)
Barrier B τ	1.20 (0.14)	−0.05 (0.04)
Barrier B τ_{TP}	1.04 (0.15)	0.05 (0.06)
Barrier A+B τ	0.92 (0.09)	0.12 (0.04)
Qiu et al. 2005[15] 35 °C	0.80 (0.03)	0.34 (0.01)

[a]Numbers in round brackets are the errors. [b]Average fitting coefficients to the relaxation times from the first nine eigenvalues.

the hairpin-biased GB1 are proportional to the external friction from solvent, and so the hairpin topology of GB1 exhibits a "zero-internal friction" behavior, consistent with the experimental finding of $\beta = 1.07 \pm 0.25$.[24]

The transition-path sampling results provide another probe of the viscosity-dependent dynamics, via the transition-path duration, $\langle\tau_{TP}\rangle$. This quantity corresponds to the time taken to cross the barrier, much shorter than the folding or unfolding time. For a one-dimensional energy landscape, Szabo has shown that $\langle\tau_{TP}\rangle$ is insensitive to the barrier height, and inversely proportional to the diffusion coefficient near the top of the barrier.[60] Since the free energy barrier is unchanged by varying viscosity with water mass-scaling, $\langle\tau_{TP}\rangle$ should scale in the same way as the relaxation time when changing the solvent viscosity, assuming one-dimensional diffusion is a reasonable approximation. Although transition-path durations are in general not correlated with folding times, there is some correlation in our calculation, because the same relative transition-path weights needed to compute $\langle\tau_{TP}\rangle$ are also used in calculating the rate coefficient in eq 1; nonetheless, they are not trivially related and so provide a different probe of the folding dynamics. The viscosity dependence of $\langle\tau_{TP}\rangle$ (inset of Figure 2) fits a power-law with exponent 0.82 ± 0.05, in reasonable agreement with that for the relaxation time τ. Thus, the folding time is sensitive to the variation of solvent viscosity, suggesting that the energy landscape near the folding barrier of the hairpin-biased GB1 lacks the local roughness causing internal friction.

Internal Friction in a GB1 Helix. Next, we check for internal friction in the helix-biased GB1. The model is generated based on the helix structure shown in Figure 1, constructed with perfect helical dihedral angles [−60°,−45°] for the backbone. Because of the absence of a folding barrier, TPS would confer no advantage, and several long μs time scale simulations with different water-mass scaling covered the relaxation time of helix formation in the peptide reasonably well. The free energy landscapes at different viscosities are shown in Figure 1. The system involves no barriers along Q but many (meta)stable states corresponding to the

DOI: 10.1021/ja511609u
J. Am. Chem. Soc. 2015, 137, 3283−3290

formation of helices with different lengths in different regions of the peptide. The helix-biased GB1 shows an average of 23% helical population (Table S2). Checking the time series of the secondary structure maps (using DSSP,[61] Figure S3) at different viscosities showed no evidence of residual hairpin formation.

We employ a Markov state model (MSM)[51,62] to analyze the time scales of the slowest motions of helix-biased GB1, defining the microstates based on backbone torsion angles.[63] Here we show the relaxation times from the first nine nontrivial eigenvalues of the transition matrix in Figure 2. All of them behave similarly and are relatively insensitive to the variation of the solvent viscosity, with an average power-law exponent of $\sim 0.67 \pm 0.02$, suggesting the presence of internal friction for all the slowest motions of the helix topology. This is in excellent agreement with the experimental result for α-helices, 0.64 ± 0.07,[24] and with both experiments[15−18] and theoretical work on the existence of internal friction in small helical proteins.[28,30]

Since the view that internal friction is caused by crossing of torsion barriers may appear biased by our choice of torsion angles to define the MSM states (since torsion angles have previously been shown to be a cause of internal friction), we have constructed a second MSM using native hydrogen bond formation. The hydrogen-bond MSM gives almost the same relaxation times and viscosity dependence as that of the torsion-angle MSM in the helix-biased GB1 (Figure S4), suggesting that both discretizations of configuration space are equally good at capturing the global folding dynamics; this may be expected due to the correlation between the formation of helical hydrogen bonds and fixing the backbone in the helical region of the Ramachandran map. Both MSMs indicate a similar deviation from zero-internal friction for the helix forming peptide.

In order to place the folding kinetics of the hairpin-biased GB1 in the same framework as that of the helix-biased version, and to ensure that the results are independent of the rate calculation method, we also constructed MSMs for the GB1-hairpin using similar discretizations based on torsion angles or native hydrogen bonds. In this case, rather than using equilibrium sampling, we harvested short trajectories starting from the same configurations used to initiate the TPS runs, but for a fixed length of time in order to also sample the stable end states, rather than stopping when they reached the stable states as in TPS. We note that we obtain equilibrium folded populations from these MSMs based on short trajectories which are similar to those from REMD (Table S3). The larger folded population from REMD is likely due to the limited length of REMD trajectory, which started from the folded structure.

In the hydrogen-bond MSM, both the slowest relaxation time and the corresponding power-law exponents (Figure S4, Table S4) for the hairpin are consistent with those from TPS (Figure 2). Interestingly, the relaxation times for the faster modes of the GB1 hydrogen-bond MSM do show smaller power-law exponents and therefore more internal friction, an effect that had been predicted theoretically for high frequency modes.[32] The slowest relaxation time of the torsion-angle MSM is however much shorter than that of either the hydrogen-bond MSM or TPS (Figure S4). This suggests that in this case the torsion-angle discretization does not capture the global folding dynamics, probably because most of the torsion angles are in an extended state also when the peptide is unfolded, and so the slowest motion of hairpin-biased GB1 is weakly correlated to the torsion transitions.

Variation of Internal Friction Across the Trp Cage Folding Landscape. Trp cage is a mini-protein in which both hairpin and helix topologies are involved in folding.[49] Internal friction has been observed in Trp cage in both experiment[15] and a recent computational study.[30] In this paper, a native-structure-biased LJ term is applied to Trp cage to further investigate the dependence of internal friction on native topology, in the same way as for GB1. We resolved two barriers, A and B, in the projection of the free energy onto Q from an REMD simulation (Figure 3). The contacts corresponding to the transitions of the

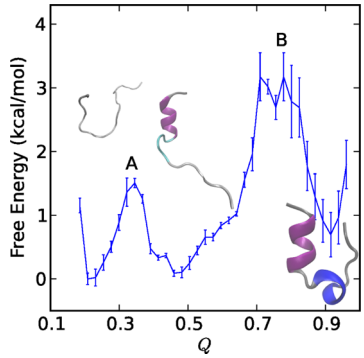


Figure 3. Free energy $F(Q)$ of native-biased Trp cage at 350 K and the typical configurations in the three minima. Transition-path sampling simulations were conducted separately for the two barriers labeled A and B.

two barriers are shown in Figure S5. The small barrier A corresponds to the folding of the first half of the peptide, comprising two turns of α-helix. The major barrier B corresponds to the docking of the C-terminal half of the peptide (polyproline helix) onto the N-terminal helix, similar to the formation of a hairpin structure. This hairpin barrier is roughly ~ 1.5 kcal mol^{-1} higher than the helix barrier. There is no experimental evidence for an intermediate on the unfolded side of the main barrier for Trp cage, although this may be difficult to determine because folding would be dominated by the larger barrier. Intermediates were identified in a previous transition-path sampling study of this system, however.[64] Regardless, the native-biased Trp cage is a useful model system to probe the variation of the internal friction along a folding pathway.

We applied TPS to each barrier in turn to illustrate the viscosity dependence of the relaxation times (Figure 4). A linear fit for the relaxation time of barrier A showed a nonzero intercept, and an exponent for the corresponding power-law fit of $\sim 0.66 \pm 0.04$, close to the helix-biased GB1. We also observed near linear dependence of both the relaxation time and the transition path time on the solvent viscosity for barrier B. Though the power-law/linear fits of the relaxation time of the hairpin barrier are less well determined, the difference between the power-law exponents for the hairpin and helix barriers is outside of the errors, suggesting a statistically meaningful difference of internal friction. These results coincide with our observations in the helix-biased and hairpin-biased GB1 that the internal friction is correlated to the nature of the structural transition. They are also consistent with the expectation that lower barriers tend to exhibit more evidence for internal friction.

We merged the dynamics for crossing the two barriers together to obtain a global relaxation between the folded and unfolded state to compare with experiment. The power-law exponent for the viscosity dependence of the global relaxation is $\sim 0.92 \pm 0.09$, compared to $\sim 0.80 \pm 0.03$ from experiment.[15] The fact that we obtained a smaller internal friction than experiment might arise

DOI: 10.1021/ja511609u
J. Am. Chem. Soc. 2015, 137, 3283−3290

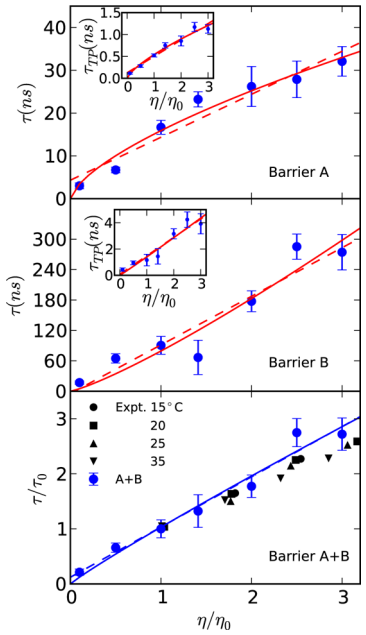


Figure 4. Viscosity dependence of the relaxation time of Trp cage. Top: barrier A. Middle: barrier B. Bottom: global relaxation over both barriers A and B. The insets are the average transition path times corresponding to each barrier, with full distributions given in Figure S2. Dashed lines show linear fits whereas solid lines show power-law fits to the data. Blue symbols are relaxation times from simulation, while black symbols in bottom plot are experimental relaxation times taken from Qiu and Hagen.[15]

from the relative barrier heights for forming the helical secondary structure and the tertiary hairpin being inaccurately captured by our model, owing to the native-centric bias: in a recent work,[30] we have observed stronger internal friction for Trp cage in a different force field (Amber ff03 and TIP3P water) without the native bias used in the present work. The native bias results in qualitative differences in the 2D free energy surface for the fraction of native contacts and radius of gyration, resulting in a more expanded unfolded state and higher folding barrier (Figure S6).

DISCUSSION

There are many experimental examples of proteins showing a wide-range of internal friction,[15−18,20−22] with α-helical proteins having the highest and all-β or α/β proteins the lowest internal friction. These trends are very suggestive of a correlation between native state topology and internal friction. By carefully constructing a model of the GB1 hairpin sequence that can fold into either an α-helix or β-hairpin, we are able to relate unambiguously the observed internal friction to the native topology; a similar difference of internal friction is found for the two barriers of Trp cage. What are the general principles governing this topology dependent internal friction?

In our previous work,[30] we have shown that torsion barriers contribute significantly to the internal friction, and were able to explain the dependence of helix formation rates on solvent viscosity. We showed that crossing of sharp barriers on the energy landscape could be a source of internal friction.[32] Other work has suggested a role for correlated torsion barrier crossings in causing internal friction in unfolded state dynamics.[31] Therefore, the different extent of internal friction for hairpin and helix topologies may be related to the importance of crossing local dihedral barriers near the folding barrier, the region of the energy landscape where the diffusion coefficient contributes most to the folding rate.

To investigate this, we plot in Figure 5 the fraction of native dihedral angles, Q_{dih}, as a function of the fraction of native

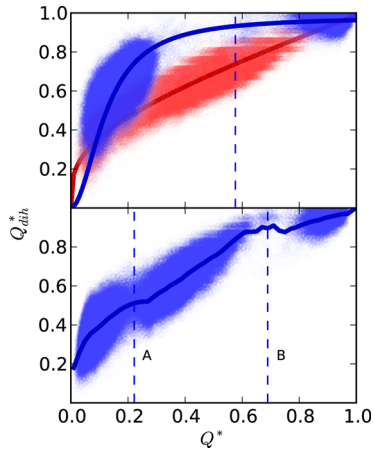


Figure 5. Scatter plots of fraction of native dihedrals Q_{dih} as a function of fraction of native contacts Q. Here * indicates both Q and Q_{dih} are normalized to the scale [0,1] for comparison. Top: Hairpin-biased GB1 is in blue and helix-biased GB1 is in red. The dashed lines are the position of the dividing surface near the barrier top, where we initialized transition-path sampling. A function $Q_{dih} = 1/(aQ^b + c)$ is used to fit the data to guide the eye. Bottom: The same plot for Trp cage from REMD. The solid lines are the average of Q_{dih} to guide the eye. The broken lines are the positions of the dividing surfaces for barriers A and B. Note that because the vertical scales have been normalized, the absolute change in the number of dihedrals between unfolded and folded may differ (e.g., it is much larger for the GB1 helix than for the GB1 hairpin).

contacts, Q, for the two GB1 models and for Trp cage. In helix-biased GB1 there is an almost linear increase in fraction of native torsion angles with fraction of native contacts, as would be expected; moreover, since there is no barrier to folding, the relaxation time should be sensitive to the diffusion coefficient at all points on the landscape. The behavior of the GB1 hairpin is more complex, with Q_{dih} being initially steeply dependent on Q, but becoming almost independent of Q near the barrier top (barrier position indicated with broken lines). In the hairpin, only a few torsion angles need to be rearranged in the turn, since extended or polyproline II conformations dominate the disordered states,[65] and that appears to happen before entering the transition state region. A factor that may further localize the internal friction effect to the transition state region for the hairpin is its significant folding barrier of $\sim 4\ k_B T$: in the Kramers theory of one-dimensional barrier crossing, only the diffusion coefficient for positions close (energetically) to the barrier top contributes significantly to the rate coefficient.[11]

Trp cage presents an interesting comparison, because both the helix and hairpin formation involve barriers. The plot of Q_{dih} vs Q shows some differences between the transition states of the two barriers (Figure 5). The dependence of Q_{dih} on Q is slightly greater for the first barrier (A) than for the second (B). By itself this is not a clear-cut explanation of the difference in internal friction between the two barriers. However, it may be that because the second folding barrier is higher than the first, it is less sensitive to torsional changes.

Although torsion barriers appear to explain the internal friction, an additional possibility is that the mechanism of helix

DOI: 10.1021/ja511609u
J. Am. Chem. Soc. 2015, 137, 3283−3290

formation is less sensitive to solvent friction than hairpin formation, because helix formation displaces less solvent, as has been suggested previously.[30] We have tested this intuitively appealing argument by developing a metric describing the atom displacement of proteins, related to the solvent that needs to be displaced during transition paths. There is no substantial difference of the atom displacement during the transition path of a helix or hairpin (Figure S7). This suggests that for these short peptides, torsion angle friction is most likely the main difference between the helix and hairpin; for longer chains, however, other effects could play a more important role.

Another possibility that has previously been considered as a source of internal friction is contact formation (either native, or non-native). In a previous simulation study of unfolded chain dynamics, this was not found to be the main cause of internal friction.[31] Analogous to our earlier work in which we used a dipeptide to test for internal friction effects on an isolated torsion angle,[30] we have used bimolecular association of two blocked amino acids as a model for contact formation. We examined contact formation between either two blocked tryptophan residues, or between blocked lysine and blocked aspartic acid (Figure S8). Both association and dissociation times for the tryptophans showed a first power dependence on solvent viscosity. For lysine and aspartic acid, there was some evidence of weak internal friction, but smaller than that seen for helix formation in the folding simulations. This suggests that contact formation by itself does not explain the observed internal friction in our simulations. However, it should be noted that while the bimolecular model has the advantage of a simple interpretation, additional complexity may arise when considering the association of more extended protein interfaces.

Non-native interactions have also been suggested as a cause of internal friction.[18] However, in all of the systems discussed here, the applied native bias reduces the importance of non-native interactions, such that we cannot easily comment on their role. In a previous study, we considered Trp cage folding without a native centric bias and a different force field. While this suggests a comparison between the results of that work and the present one, this is complicated by the substantial differences in their free energy landscapes (Figure S6), with the simulation without bias having essentially no barrier to folding. Interestingly, however, the dependence of Q_{dih} on Q for that simulation is quite similar to that in the present work (data not shown). Thus, a possible explanation for the difference is the absence of a barrier in the earlier study, making the folding relaxation sensitive to the diffusion coefficient over a larger region of the energy landscape.

We have also assessed the potential role of non-native interactions on helix formation in GB1: the unfolded state of GB1 with the original force field (absent helix or hairpin bias), has a substantial propensity to form helix (secondary structure map shown in Figure S3). The high helix population is most likely a force field artifact due to the excessively high helical propensity of the two aspartate residues in the middle of the GB1 sequence in the force field, relative to experiment.[66,67] We find that, with the inclusion of non-native interactions, the relaxation times for forming this helix have very similar viscosity dependence to those for the helix-biased potential (Figure S9). Thus, at least in this case, non-native interactions do not appear to significantly increase the internal friction, relative to the native-centric model. Of course, non-native interactions may still add roughness to the energy landscape, without increasing the internal friction probed by the viscosity dependence.

Studies on unfolded states of a variety of unfolded proteins and intrinsically disordered proteins have revealed internal friction to be a universal feature of these states, in the absence of denaturant.[14,41] Therefore, it is interesting to check for internal friction in the unfolded chain dynamics of a system where internal friction is not evident in the dynamics of barrier crossing. We have built a MSM on only the unfolded part of the trajectories of hairpin-biased GB1. Although the statistical errors are substantial, the slowest relaxation time reflects an insensitivity to solvent viscosity similar to that for helix formation, suggesting an internal friction behavior different from that for the barrier crossing (Figure S9, Table S4). This is consistent with internal friction being identified in the reconfiguration dynamics of many proteins in the unfolded state, even when there is no internal friction in their folding dynamics.[41]

CONCLUSION

We find that the deviation of protein folding relaxation times from a first-power dependence on solvent viscosity, as is commonly used to define internal friction, can be clearly correlated with native topology, at least for the small systems considered here. The explanation we propose, suggested by a careful analysis of the folding dynamics, is the following: when folding is limited by a significant free energy barrier, folding mechanisms in which torsional isomerization plays an important role in the folding mechanism near the barrier top (as in helix formation) will tend to exhibit more evidence of internal friction than the folding mechanisms for forming other structures. An additional factor that cannot be directly assessed here is that lower (or vanishing) folding barriers are more likely to exhibit internal friction, because folding dynamics should be sensitive to the diffusion coefficient over a wider region of the energy landscape. This proposal may help to resolve the variations in internal friction observed for different proteins in experiment, although further investigation would be needed to generalize the results to other systems, and to identify the role of other possible effects such as non-native interactions.

ASSOCIATED CONTENT

Supporting Information

Additional text, figures and tables. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

robertbe@helix.nih.gov

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank William Eaton for helpful comments on the manuscript. This work was supported by the Intramural Research Programme of the National Institute of Diabetes and Digestive and Kidney Diseases of the National Institutes of Health (R.B.B., T.H. and W.Z) and a grant from the Engineering and Physical Sciences Research Council (D.d.S.). This study utilized the high-performance computational capabilities of the Biowulf Linux cluster at the National Institutes of Health, Bethesda, MD (http://biowulf.nih.gov).

DOI: 10.1021/ja511609u
J. Am. Chem. Soc. 2015, 137, 3283−3290

REFERENCES

(1) Wolynes, P. G.; Onuchic, J. N.; Thirumalai, D. *Science* **1995**, *267*, 1619−1620.
(2) Dill, K. A.; Chan, H. S. *Nat. Struct. Biol.* **1997**, *4*, 10−19.
(3) Oliveberg, M.; Wolynes, P. G. *Q. Rev. Biophys.* **2005**, *38*, 245−288.
(4) Miller, M. A.; Wales, D. J. *J. Chem. Phys.* **1999**, *111*, 6610−6616.
(5) Chavez, L. L.; Onuchic, J. N.; Clementi, C. *J. Am. Chem. Soc.* **2004**, *126*, 8426−8432.
(6) Shea, J.-E.; Onuchic, J. N.; Brooks, C. L., III *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 12512−12517.
(7) Clementi, C.; Nymeyer, H.; Onuchic, J. N. *J. Mol. Biol.* **2000**, *298*, 937−953.
(8) Levy, Y.; Wolynes, P. G.; Onuchic, J. N. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 511−516.
(9) Klimov, D. K.; Thirumalai, D. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 7254−7259.
(10) Bryngelson, J. D.; Wolynes, P. G. *J. Phys. Chem.* **1989**, *93*, 6902−6915.
(11) Kramers, H. A. *Physica* **1940**, *7*, 284−304.
(12) Beece, D.; Eisenstein, L.; Frauenfelder, H.; Good, D.; Marden, M. C.; Reinisch, L.; Reynolds, A. H.; Sorensen, L. B.; Yue, K. T. *Biochemistry* **1980**, *19*, 5147−5157.
(13) Ansari, A.; Jones, C. M.; Henry, E. R.; Hofrichter, J.; Eaton, W. A. *Science* **1992**, *256*, 1796−1798.
(14) Soranno, A.; Buchli, B.; Nettels, D.; Cheng, R. R.; Müller-Späth, S.; Pfeil, S. H.; Hoffmann, A.; Lipman, E. A.; Makarov, D. E.; Schuler, B. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 17800−17806.
(15) Qiu, L.; Hagen, S. J. *Chem. Phys.* **2005**, *312*, 327−333.
(16) Cellmer, T.; Henry, E. R.; Hofrichter, J.; Eaton, W. A. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 18320−18325.
(17) Wensley, B. G.; Batey, S.; Bone, F. A. C.; Chan, Z. M.; Tumelty, N. R.; Steward, A.; Kwa, L. G.; Borgia, A.; Clarke, J. *Nature* **2010**, *463*, 685−688.
(18) Chung, H. S.; Eaton, W. A. *Nature* **2013**, *502*, 685−688.
(19) Liu, F.; Nakaema, M.; Gruebele, M. *J. Chem. Phys.* **2009**, *131*, 195101.
(20) Chrunyk, B. A.; Matthews, C. R. *Biochemistry* **1990**, *29*, 2149−2154.
(21) Jacob, M.; Geeves, M.; Holtermann, G.; Schmid, F. X. *Nat. Struct. Biol.* **1999**, *6*, 923−926.
(22) Plaxco, K. W.; Baker, D. *Proc. Natl. Acad. Sci. U. S. A* **1998**, *95*, 13591−13596.
(23) Piana, S.; Lindorff-Larsen, K.; Shaw, D. E. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 5915−5920.
(24) Jas, G. S.; Eaton, W. A.; Hofrichter, J. *J. Phys. Chem. B* **2001**, *105*, 261−272.
(25) Klimov, D. K.; Thirumalai, D. *Phys. Rev. Lett.* **1997**, *79*, 317−320.
(26) Zagrovic, B.; Pande, V. S. *J. Comput. Chem.* **2003**, *24*, 1432−1436.
(27) Walser, R.; Mark, A. E.; van Gunsteren, W. F. *Chem. Phys. Lett.* **1999**, *303*, 583−586.
(28) Schulz, J.; Schmidt, L.; Best, R. B.; Dzubiella, J.; Netz, R. *J. Am. Chem. Soc.* **2012**, *134*, 6273−6279.
(29) McQuarrie, D. A. *Statistical Mechanics*; University Science Books: Sausalito, CA, 2000.
(30) De Sancho, D.; Best, R. B. *Nat. Commun.* **2014**, *5*, 4307.
(31) Echeverria, I.; Makarov, D. E.; Papoian, G. A. *J. Am. Chem. Soc.* **2014**, *136*, 8708−8713.
(32) Portman, J. J.; Takada, S.; Wolynes, P. G. *J. Chem. Phys.* **2001**, *114*, 5082−5096.
(33) Kuhn, W.; Kuhn, H. *Helv. Chim. Acta* **1946**, *29*, 609−626.
(34) Chung, H. S.; McHale, K.; Louis, J. M.; Eaton, W. A. *Science* **2012**, *335*, 981−984.
(35) Liu, F.; Gruebele, M. *J. Mol. Biol.* **2007**, *370*, 574−584.
(36) Carter, J. W.; Baker, C. M.; Best, R. B.; De Sancho, D. *J. Phys. Chem. B* **2013**, *117*, 13435−13443.
(37) Capaldi, A. P.; Kleanthous, C.; Radford, S. E. *Nat. Struct. Mol. Biol.* **2002**, *9*, 209−216.
(38) Korzhnev, D. M.; Religa, T. L.; Banachewicz, W.; Fersht, A. R.; Kay, L. E. *Science* **2010**, *329*, 1312−1316.
(39) Gallagher, T.; Alexander, P.; Bryan, P.; Gilliland, G. L. *Biochemistry* **1994**, *33*, 4721−4729.
(40) Cregut, D.; Civera, C.; Macias, M.; Wallon, G.; Serrano, L. *J. Mol. Biol.* **1999**, *292*, 389−401.
(41) Borgia, A.; Wensley, B. G.; Soranno, A.; Nettels, D.; Borgia, M. B.; Hoffmann, A.; Pfeil, S. H.; Lipman, E. A.; Clarke, J.; Schuler, B. *Nat. Commun.* **2012**, *3*, 1195.
(42) Hess, B.; Kutzner, C.; Van Der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.
(43) Bussi, G.; Donadio, D.; Parrinello, M. *J. Chem. Phys.* **2007**, *126*, 014101.
(44) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **1999**, *314*, 141−151.
(45) Parrinello, M.; Rahman, A. *J. Appl. Phys.* **1981**, *52*, 7182−7190.
(46) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. A. *J. Comput. Chem.* **2003**, *24*, 1999−2012.
(47) Best, R. B.; Mittal, J. *J. Phys. Chem. B* **2010**, *114*, 14916−14923.
(48) Best, R. B.; Zheng, W.; Mittal, J. *J. Chem. Theory Comput.* **2014**, *10*, 5113−5124.
(49) Neidigh, J. W.; Fesinmeyer, R. M.; Andersen, N. H. *Nat. Struct. Mol. Biol.* **2002**, *9*, 425−430.
(50) Hummer, G. *J. Chem. Phys.* **2004**, *120*, 516−523.
(51) Chodera, J. D.; Noé, F. *Curr. Opin. Struct. Biol.* **2014**, *25*, 135−144.
(52) Buchete, N.-V.; Hummer, G. *J. Phys. Chem. B* **2008**, *112*, 6057−6069.
(53) Prinz, J.-H.; Wu, H.; Sarich, M.; Keller, B.; Senne, M.; Held, M.; Chodera, J. D.; Schutte, C.; Noe, F. *J. Chem. Phys.* **2011**, *134*, 174105.
(54) Yang, W. Y.; Gruebele, M. *Nature* **2003**, *423*, 193−197.
(55) Munoz, V.; Thompson, P. A.; Hofrichter, J.; Eaton, W. A. *Nature* **1997**, *390*, 196−199.
(56) De Sancho, D.; Mittal, J.; Best, R. B. *J. Chem. Theory Comput.* **2013**, *9*, 1743−1753.
(57) Muñoz, V.; Henry, E. R.; Hofrichter, J.; Eaton, W. A. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 5872−5879.
(58) Best, R. B.; Mittal, J. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 11087−11092.
(59) Dellago, C.; Bolhuis, P. G.; Csajka, F. S.; Chandler, D. *J. Chem. Phys.* **1998**, *108*, 1964−1977.
(60) Chung, H. S.; Louis, J. M.; Eaton, W. A. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 11837−11844.
(61) Kabsch, W.; Sander, C. *Biopolymers* **1983**, *22*, 2577−2637.
(62) Bowman, G. R.; Beauchamp, K. A.; Boxer, G.; Pande, V. S. *J. Chem. Phys.* **2009**, *131*, 124101.
(63) De Sancho, D.; Best, R. B. *J. Am. Chem. Soc.* **2011**, *133*, 6809−6816.
(64) Juraszek, J.; Bolhuis, P. G. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 15859−15864.
(65) Hagarman, A.; Measey, T. J.; Mathieu, D.; Schwalbe, H.; Schweitzer-Stenner, R. *J. Am. Chem. Soc.* **2010**, *103*, 540−551.
(66) Best, R. B.; De Sancho, D.; Mittal, J. *Biophys. J.* **2012**, *102*, 1897−1906.
(67) Moreau, R. J.; Schubert, C. R.; Nasr, K. A.; Torok, M.; Miller, J. S.; Kennedy, R. J.; Kemp, D. S. *J. Am. Chem. Soc.* **2009**, *131*, 13107−13116.

DOI: 10.1021/ja511609u
J. Am. Chem. Soc. 2015, 137, 3283−3290