

# Slow Folding of a Helical Protein: Large Barriers, Strong Internal Friction, or a Shallow, Bumpy Landscape?

Sandhya Subramanian,<sup>1</sup> Hemashree Golla,<sup>1</sup> Kalivarathan Divakar, Adithi Kannan, David de Sancho,\* and Athi N. Naganathan\*

Cite This: *J. Phys. Chem. B* 2020, 124, 8973–8983

Read Online

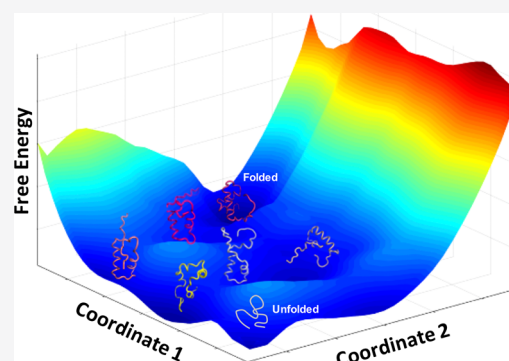
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

**ABSTRACT:** The rate at which a protein molecule folds is determined by opposing energetic and entropic contributions to the free energy that shape the folding landscape. Delineating the extent to which they impact the diffusional barrier-crossing events, including the magnitude of internal friction and barrier height, has largely been a challenging task. In this work, we extract the underlying thermodynamic and dynamic contributions to the folding rate of an unusually slow-folding helical DNA-binding domain, PurR, which shares the characteristics of ultrafast downhill-folding proteins but nonetheless appears to exhibit an apparent two-state equilibrium. We combine equilibrium spectroscopy, temperature-viscosity-dependent kinetics, statistical mechanical modeling, and coarse-grained simulations to show that the conformational behavior of PurR is highly heterogeneous characterized by a large spread in melting temperatures, marginal thermodynamic barriers, and populated partially structured states. PurR appears to be at the threshold of disorder arising from frustrated electrostatics and weak packing that in turn slows down folding due to a shallow, bumpy landscape and not due to large thermodynamic barriers or strong internal friction. Our work highlights how a strong temperature dependence on the pre-exponential could signal a shallow landscape and not necessarily a slow-folding diffusion coefficient, thus determining the folding timescales of even millisecond folding proteins and hints at possible structural origins for the shallow landscape.



## INTRODUCTION

The marginal stability of most proteins has its origins in the strength and nature of interactions including hydrophobic packing in the protein interior and surface electrostatics that can be both favorable and unfavorable. The requirements for marginal stability are manifold and can range from regulatory needs of the organism to efficient folding and function.<sup>1–9</sup> These conflicting requirements contribute to the evolution of protein sequences and determine not just the identity of the amino acid at a specific location in the sequence but also subtly influence the immediate environment around it, contributing to the coevolution of sites far in the protein sequence.<sup>10–14</sup> Such features and constraints are driven to an extreme in DNA-binding domains (DBDs) whose entire nucleic acid-binding face exhibits a net positive electrostatic potential to bind the polyanionic partner. In fact, many of the DBDs display complex thermodynamics in the apo form, which has been attributed to the population of partially structured states in solution.<sup>15–26</sup> The conformational complexity need not be restricted to DBDs but likely applies to any macromolecule that binds an oppositely charged counterpart. “Frustration,” even if minimized by evolutionary forces to guarantee efficient folding, is a universal feature of proteins and enzymes and is

predicted to influence or in some cases even drive folding-function behaviors.<sup>27,28</sup>

An outstanding question is the extent to which such frustration (from geometrical constraints, packing interactions, or electrostatics) influences the folding landscape of DBDs and thereby the folding speed. Assuming a simple one-dimensional free-energy profile, the rate of diffusional barrier crossing ( $k$ ) according to Kramers rate theory<sup>29</sup> can be written as

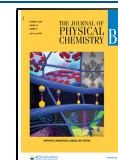
$$k = \frac{\omega_U \omega^* D}{2\pi k_B T} e^{(-\Delta G^*/k_B T)} \quad (1)$$

where  $\omega_U^2$  and  $\omega^{*2}$  are the curvatures of the unfolded well and barrier top (transition state), respectively, while  $D$  is the folding diffusion coefficient and  $\Delta G^*$  is the free-energy barrier to folding at temperature  $T$ . Thus, the rate of folding is determined not just by the thermodynamic barrier height ( $\Delta G^*$ ) but also the curvatures of the reactant and barrier top

Received: June 30, 2020

Revised: August 27, 2020

Published: September 21, 2020

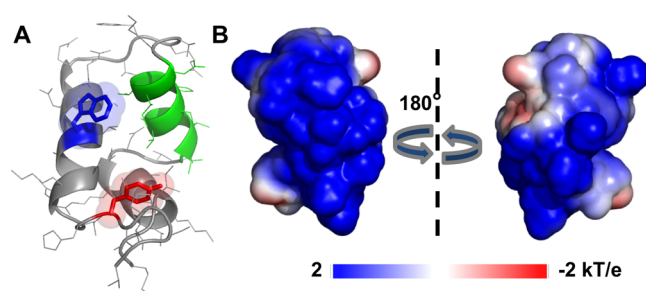


wells ( $\omega^2$ ) and the folding diffusion coefficient ( $D$ ), with latter terms being part of the pre-exponential factor,  $k_0$ . The large residual frustration in DBDs can have an impact on the magnitude of thermodynamic free-energy barriers or on the pre-exponential factor to protein folding or both, thus governing the folding speed. The effect on the diffusion constant in the pre-exponential has variously been termed as internal friction or landscape roughness or ruggedness.<sup>27,30–33</sup>

It arises from microscopic barriers to peptide bond rotations<sup>34,35</sup> and non-native interactions<sup>36–38</sup> that need to be broken to form native contacts, with the latter likely dominating the dynamics of single-domain proteins.<sup>39</sup> Such a feature can manifest as roughness both along and orthogonal to the folding coordinate, thus slowing down the folding rate.<sup>40</sup>

Interestingly, many fast-folding proteins exhibit a dramatic slowing down of folding relaxation rates at low temperatures ( $\sim 5000$ – $30,000$  s<sup>-1</sup> between 280–290 K) despite matching with theoretical speed limit expectations of  $\sim 10^6$  s<sup>-1</sup> at 333 K.<sup>41–45</sup> Given that they exhibit downhill folding profiles under these conditions, the rates directly report on the pre-exponential factor to protein folding and are suggestive of a large temperature dependence on the pre-exponential term.<sup>27</sup> The large temperature dependence is conventionally interpreted as large internal friction that in turn slows down the folding diffusion coefficient ( $D$  in eq 1). However, it is also possible that a shallow landscape or a weak curvature in the unfolded well or barrier top ( $\omega_U^2$  and  $\omega^{*2}$  in eq 1) could equally slow down the folding, with the degree of curvature changing with temperature. In this regard, millisecond folders are generally thought to fold slow because of large thermodynamic barriers ( $>3$  RT). While this expectation holds true for many systems,<sup>46</sup> it is possible that some proteins exhibit enhanced frustration due to functional or regulatory constraints that slow down the pre-exponential factor, similar to the extent observed in downhill-folding proteins. Recent simulations highlight that the dynamics of  $\alpha$ -helical systems are more sensitive to frustration than those of  $\beta$ -sheet proteins,<sup>47</sup> and this is borne out in the studies of R17 and  $\alpha_3D$ .<sup>48,49</sup> Interestingly, even  $\beta$ -sheet proteins exhibit an order of magnitude difference in the pre-exponential factor within members of the same family.<sup>50</sup>

These observations raise questions on the extent to which the magnitude of the pre-exponential factor affects the folding speed of slow-folding  $\alpha$ -helical proteins and whether one can extract the thermodynamic and dynamic contributions to the rate equation by studying them. We extend this question further and ask if the slow folding is due to large free-energy barriers separating the various substrates, strong internal friction, or the manifestation of broad unfolded wells and barrier tops (indicative of shallow landscapes with multiple minima). We answer these questions by probing the folding conformational landscape of the DBD of PurR (purine repressor), a 56-residue helical domain belonging to the LacR family of transcription regulators. PurR exhibits large electrostatic frustration throughout its solvent-exposed surface including the DNA-binding face (Figure 1A,B). This structural feature is surprisingly similar to its homologue CytR that is disordered in solution and folds in the presence of DNA<sup>51</sup> and unlike other related proteins of the same family (Figure S1). Such strong frustration is expected to not only destabilize the protein but also contribute to pockets of local structure that can impede access to the folded state during folding of the protein. Here, we combine experiments, statistical mechanical



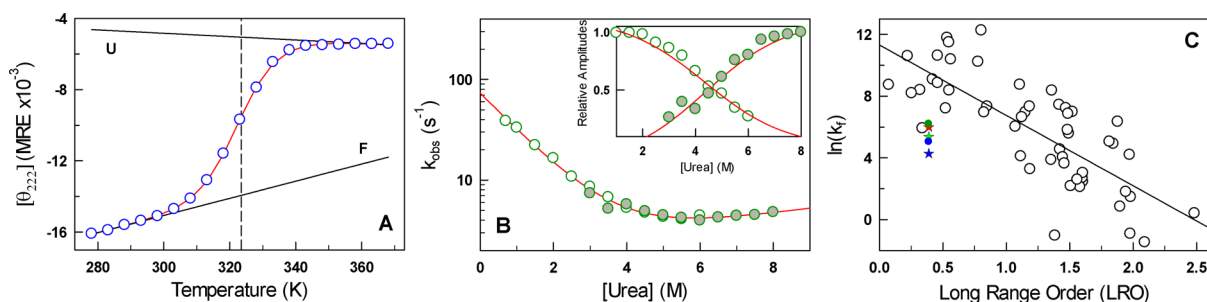
**Figure 1.** (A) Cartoon representation of PurR (PDB id 1pru) highlighting the DNA-binding helix (green) and the intrinsic fluorescence probes W37 (blue) and Y45 (red). The helical sequence boundaries are 5–11, 15–23, and 31–43 for helices 1, 2, and 3, respectively. (B) Electrostatic potential maps of PurR DNA-binding face (left) and the opposite face (right). The large positive potential on the DNA-binding face enables PurR to bind DNA.

modeling, and coarse-grained simulations to show that PurR exhibits little apparent internal friction, thermodynamically uncoupled folding of structural elements, and downhill-like folding profiles with multiple partially structured states. Despite this, PurR folds slowly, hinting that the slow folding has its likely origins in a shallow landscape.

## METHODS

The DBDs of the transcriptional repressor PurR (corresponding to residues 1–57 of the protein with the sequence: MATIKDVAKRANVSTTTVSHVINKTRFVAEETRNAVWAAIKELHYSPSAVARSLKVN) and its truncated variant (lacking the disordered C-terminal tail highlighted in bold) were overexpressed and purified as previously described.<sup>26</sup> Far-UV and near-UV circular dichroism (CD) and fluorescence experiments on PurR and its variants were performed as described before.<sup>52</sup> All experiments were carried out in pH 7.0, 20 mM sodium phosphate buffer [effective ionic strength (IS) of 43 mM], unless otherwise mentioned. Chemical denaturation experiments monitored by far-UV CD and fluorescence were performed at 285, 298, and 310 K at protein concentrations of  $\sim 25$  and  $\sim 10$   $\mu$ M, respectively. The samples were incubated for 2 h in increasing concentrations of urea (0–8 M, in intervals of 0.5 M) in 20 mM sodium phosphate buffer, pH 7.0 before measurements.

**Stopped-Flow Kinetics.** The folding and unfolding traces of PurR and its mutants were recorded at 285, 298, and 310 K by fluorescence using a Chirascan SF3 Stopped Flow instrument (deadtime of  $\sim 1$ – $2$  ms; Applied Photophysics Ltd.) coupled to a thermostated water bath as described earlier.<sup>52</sup> For temperature-dependent folding kinetics, denatured PurR in 6 M urea was refolded at 0.55 M urea (final protein concentration  $\sim 10$   $\mu$ M) at temperatures 285–305 K in steps of 2.5 K. The folding traces were recorded by exciting the protein with a 280 nm light-emitting diode (LED). A thousand data points were collected for every scan, and six scans were recorded at an interval of every 1 min at every temperature. The unfolding and refolding kinetics of PurR at different glucose (0.25–1.5 M) and urea concentrations (0–8 M) were performed in a similar manner to determine the rate dependence on solvent viscosity. A rolling-ball microviscometer (Lovis 2000 ME, Anton Paar) with a built-in temperature controller was employed to measure the dynamic viscosity of urea-glucose mixtures at 285 K.



**Figure 2.** Slow Folding of PurR. (A) Thermal unfolding profile of PurR as monitored by far-UV CD at 222 nm plotted in mean residue ellipticity units of  $\text{deg cm}^2 \text{dmol}^{-1}$ . The vertical dashed line signals the melting temperature from a two-state fit (red curve), while U and F represent unfolded and folded baselines, respectively. (B) Folding kinetics from stopped-flow experiments with open and filled circles representing the measured refolding and unfolding relaxation rates, respectively. Inset: relative kinetic amplitudes (circles) and equilibrium populations derived from a two-state model (red curve). (C) Correlation between LRO and folding rates for a database of proteins<sup>65</sup> (black circles). Blue, green, and red stars signal the PurR folding rate constants at 285, 298, and 310 K, respectively, at 43 mM IS, pH 7.0. Blue and green circles are the folding rates of PurR at 285 and 298 K at 500 mM IS, pH 7.0.

**Differential Scanning Calorimetry.** A Microcal VP-DSC microcalorimeter (Malvern Ltd.) coupled to an automated sample injector was employed to measure heat capacity profiles. All samples were degassed at room temperature prior to calorimetric measurements. Desalted protein solutions of the wild-type (WT) PurR and the truncated variant (concentrations  $\sim 25$  to  $\sim 100 \mu\text{M}$ ) and buffers were scanned at a rate of 1 K/min. Calorimetric cells were maintained under an excess pressure of 60 psi to prevent boiling at high temperatures. Buffer–buffer baselines before and after the protein scans were routinely acquired to ensure there was little thermal drift. The resulting apparent heat capacities were converted into absolute units following the method of Sanchez-Ruiz and co-workers.<sup>53</sup>

**Fluorescence Lifetime Measurements.** Time-dependent fluorescence intensity decays of W37 in PurR were recorded in a ChronosBH (ISS Inc.) spectrometer coupled to a Peltier temperature controller. The excitation pulse (from a 300 nm LED) and emitted photons were passed through UV grade Glan-Thompson polarizers set at 0 and  $54.7^\circ$ , respectively, from the vertical  $z$ -axis. The instrument response function was measured using LUDOX solution. The emitted photons were passed through a 345 nm long-pass filter (SCHOTT) to minimize scattering artifacts. All decay curves were recorded until the peak count reached  $10^4$  or the total count approached  $10^8$ . The traces were fitted to biexponential functions with the  $\chi^2$  values being  $< 1.5$  at all temperatures.

**WSME Model.** The native-centric Wako-Saitô-Muñoz-Eaton (WSME) model, coarse-grained at the residue level,<sup>54,55</sup> was employed to derive the thermodynamic landscape of PurR. Briefly, the WSME model assigns a residue conformational status 0 for unfolded and 1 for folded residues, enabling every possible microstate to be represented as strings of 1s and 0s. Instead of employing the version that accounts  $2^N$  states (where  $N$  is the protein length), we employ an advanced version of the WSME model<sup>56,57</sup> that includes contributions from single and two stretches of folded residues (single and double sequence approximations) while also allowing for interactions between the two folded stretches, thus accounting for a total of 791,617 microstates. The statistical weight of each of the microstates includes contributions from van der Waals interactions (heavy-atom neighbors identified with a 5 Å distance cutoff from the PDB file 1pru), electrostatics (all-to-all native electrostatics with an effective dielectric constant of 29 in the Debye–Hückel formalism), implicit solvation, and

excess conformational entropy of  $-6.1 \text{ J mol}^{-1} \text{ K}^{-1}$  per residue for nonhelical disordered residues identified by STRIDE.<sup>58</sup> Proline at position 47 was assigned an entropic penalty of 0, given its limited backbone flexibility. Heat capacity profiles were generated from derivatives of the total partition function ( $Z$ ), while free-energy profiles and conformational landscapes were obtained by algorithmically grouping microstates with a specific number of structured residues. The final parameters, obtained by quantitatively reproducing the differential scanning calorimetry (DSC) curve, are  $\xi = -106.5 \pm 0.62 \text{ J mol}^{-1}$  (van der Waals interaction energy per native contact),  $\Delta S_{\text{conf}} = -16.5 \pm 0.12 \text{ J mol}^{-1} \text{ K}^{-1}$  per residue for all residues except proline (entropic penalty for fixing a residue in the native conformation), and  $a = 1.91 \pm 0.01 \text{ kJ mol}^{-1} \text{ K}^{-1}$ . The parameter  $a$  determines the intercept of the native heat capacity baseline in the equation  $N = (a + 0.0067 \times (T - 273.15))M_w/1000$  where  $M_w$  is the molecular weight of PurR ( $6282.2 \text{ g mol}^{-1}$ ).

**Coarse-Grained Simulations.** Molecular simulations were run using the Karanicolas and Brooks structure-based model that is coarse-grained to the level of  $C_\alpha$ s.<sup>59</sup> In this model, the potential energy is defined as the sum of terms for bonds, angles, torsions, and nonbonded interactions. The terms for bonds and angles are harmonic with the equilibrium values corresponding to those between pairs of  $C_\alpha$  beads. Propensities for the dihedral energy terms are derived from a statistical analysis of structures in the PDB. Finally, favorable nonbonded terms are included for pairs of beads corresponding to amino acids whose heavy atoms are “in contact” (i.e., their distance is under a cutoff of 4.5 Å) in the reference native structures. In addition to the prescription by Karanicolas and Brooks, we incorporate the effects of electrostatic and non-native interactions following Kim and Hummer (KH), as before.<sup>60,61</sup> Interactions between charged residues are defined using a Debye–Hückel potential

$$V_{\text{elec}} = q_i q_j^* \exp(-r_{ij}/\xi) / (4\pi D r_{ij}) \quad (2)$$

where  $q_i$  and  $q_j$  are the net amino acid charges,  $D$  is the dielectric constant, and  $\xi$  is the screening length. Non-native interactions are described using a sequence-dependent Lennard-Jones potential that replaces the excluded volume term in the Karanicolas and Brooks model. Further details on the models can be found elsewhere.<sup>60,61</sup>

We generated simulation models using the PDB structure for PurR (1pru). Simulations were run at temperatures ranging between 270 and 360 K using a Langevin integrator with a friction coefficient of  $0.2 \text{ ps}^{-1}$  and a time step of 10 fs using Gromacs 4.0.5.<sup>62</sup> To analyze the data, we project the resulting trajectories on the fraction of native contacts,  $Q$ , and combine the information from multiple temperatures using the weighted histogram analysis method.<sup>63</sup>

## RESULTS AND DISCUSSION

**Equilibrium Thermodynamics and Slow-Folding Kinetics of PurR.** The equilibrium thermal unfolding of WT PurR DBD (hereon referred to as PurR) as monitored by far-UV CD at 222 nm exhibits the characteristic sigmoidal profile expected of two-state systems with well-defined pre- and post-transition regions (Figure 2A). The data can be fit to a two-state model with a  $T_m$  of  $323.5 \pm 0.5 \text{ K}$  and  $\Delta H_m$  of  $149.7 \pm 9.5 \text{ kJ mol}^{-1}$  (error bars represent 95% confidence here and throughout the text), similar to most mesophilic proteins. PurR being a small helical domain has a long-range order (LRO<sup>64</sup>) of just  $\sim 0.39$ ; this is indicative of local interactions dominating the contact energetics and predicts that PurR should fold relatively fast in the microsecond timescale. However, stopped-flow kinetics points to well-defined kinetic phases described by single exponential functions with the kinetic amplitudes matching the populations from equilibrium measurements (inset to Figures 2B and S2C). A chevronlike behavior is observed at all experimental temperatures with an extrapolated folding rate constant in the absence of urea of just  $\sim 71 \text{ s}^{-1}$  at 285 K that increases to  $\sim 347 \text{ s}^{-1}$  at 310 K (Figures 2B and S2A). PurR, therefore, falls well below the expected range in the plot of LRO versus folding rate (Figure 2C). Moreover, changing the conditions (increasing temperatures or IS to 0.5 M) affects the folding rate only marginally (Figures 2C and S2B). Thus, though PurR exhibits an apparent two-state equilibrium, it folds significantly slower than the expectation from LRO predictions, arguing for specific sequence effects determining its folding behavior. In the next few sections, we systematically explore the likely reasons for slow folding combining experiments and simulations.

**Minimal Apparent Internal Friction.** One of the primary origins of slow folding is internal friction. Although a rigorous definition of internal friction has been elusive,<sup>34,35,39</sup> it is broadly accepted that its effects are encapsulated within the folding diffusion coefficient  $D$  in Kramers rate theory (eq 1). For practical purposes, internal friction effects are quantified experimentally by measuring the kinetic rate constants at varying concentrations of viscosogens that increase the solvent viscosity ( $\eta$ ) and assuming that the folding free energy surface is unaltered.<sup>66</sup> If the dynamics are enslaved to friction from the solvent, because the friction coefficient varies in proportion to the viscosity, we can rewrite eq 1 as<sup>67</sup>

$$k = \frac{A}{\eta} e^{(-\Delta G^*/k_B T)} \quad (3)$$

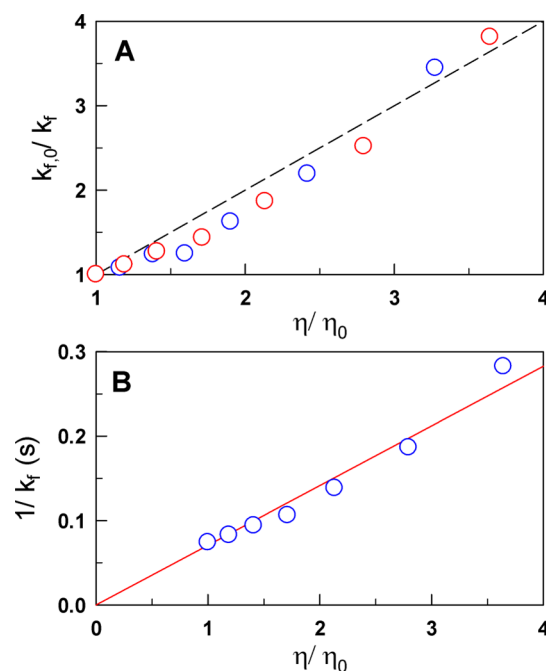
where  $A$  is a constant that includes the contributions of the energy landscape to the pre-exponential (i.e., the curvatures at the bottom of the well and the barrier top). Under iso-stability conditions, if we divide the folding rate at a reference viscosity  $\eta_0$  by that at a working viscosity  $\eta$ , that is,  $k_{f0}/k_f$ , the exponential term and  $A$  cancel out. Hence, the plot of  $k_{f0}/k_f$  versus the ratio of the viscosities  $\eta/\eta_0$  would follow a straight

line with a slope of 1 and 0 offset.<sup>67</sup> If instead, internal friction ( $\sigma$ ) needs to be invoked, then eq 3 reduces to<sup>67</sup>

$$k = \frac{A}{\eta + \sigma} e^{(-\Delta G^*/k_B T)} \quad (4)$$

and an insensitivity to the viscosity that translates to a non-zero offset and a slope lower than 1 in the plot of  $k_{f0}/k_f$  versus  $\eta/\eta_0$ .

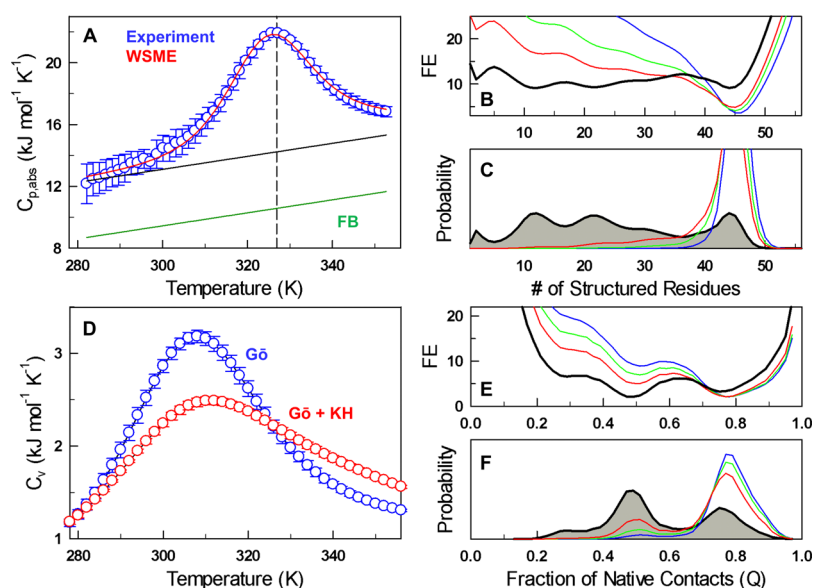
To explore this, we measured the folding relaxation rates of PurR at different glucose (0–1.5 M, the viscogen) and urea concentrations. A chevronlike behavior is evident at various glucose concentrations (Figures S3 and S4) from which the rates are extracted at two different iso-stability conditions of 8.4 and 4 kJ mol<sup>-1</sup>. The measured relative rates scale directly with the relative solvent viscosity for PurR (Figure 3A), and



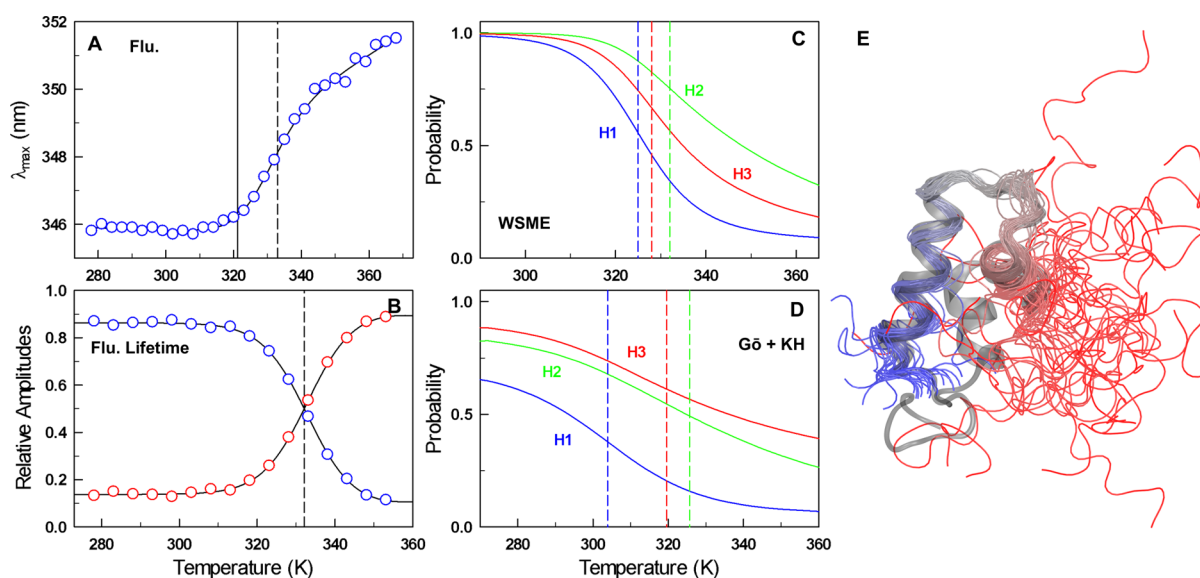
**Figure 3.** Minimal apparent internal friction. (A) Plots of relative folding rates (ordinate) versus relative viscosity at 285 K and at two different iso-stability conditions of 8.4 and 4 kJ mol<sup>-1</sup>, respectively (blue and red). For the iso-stability condition of 8.4 kJ mol<sup>-1</sup>, the urea concentration spans the range 0–3.25 M at different glucose concentrations, while for 4 kJ mol<sup>-1</sup> iso-stability, the urea concentration range is 1.95–5.55 M. (B) Plot of the folding time (ordinate) against relative viscosity at iso-stability conditions of 8.4 kJ mol<sup>-1</sup>. Red line is the expected linear dependence.

the folding times ( $\tau$ ) exhibit a near-linear dependence on the relative viscosity (Figure 3B). These results indicate that the slow folding of PurR is not a consequence of internal friction, at least as conventionally interpreted.

**Marginal Thermodynamic Barriers.** An alternate possibility is that PurR folds over large free-energy barriers that in turn contribute to its unusually slow folding. DSC is an ideal avenue for extracting the thermodynamic barriers,<sup>68</sup> given the fundamental connection between heat capacity and partition function,<sup>69,70</sup> which has been validated in several small single-domain proteins.<sup>71,72</sup> To probe if the folding thermodynamics of PurR is characterized by large thermodynamic barriers, we measured the absolute heat capacity profile of PurR from the dependence of the apparent heat capacity on protein concentrations (Figure S5A). The resulting DSC



**Figure 4.** Marginal thermodynamic barriers. (A) Absolute heat capacity of PurR (circles) fit to the WSME model (red) together with the predicted native baseline (black). The thermodynamic fluctuations of well-folded proteins follow a specific trend with temperature and molecular weight that is represented by the Freire baseline (green). The observed upward shift of the experimental data is generally seen as evidence for large thermodynamic fluctuations in the native ensemble. The vertical dashed line signals the melting temperature from a two-state fit. (B,C) Predicted free-energy profiles (panel B, in  $\text{kJ mol}^{-1}$ ) and probability distributions (panel C) at 278 K (blue), 298 K (green), 310 K (red), and 329 K (black). Note the downhill-folding profiles at low temperatures and the marginal barrier with multiple partially structured states at 329 K. (D) Heat capacity curves derived from coarse-grained simulations with native ( $G\ddot{o}$ ) and native plus non-native potentials ( $G\ddot{o} + KH$ ). (E,F) Free-energy profiles (panel E, in  $\text{kJ mol}^{-1}$ ) and the corresponding density distributions (panel F) as a function of the fraction of native contacts for the  $G\ddot{o} + KH$  potential at 270 K (blue), 280 K (green), 290 K (red), and 310 K (black). Note that 310 K represents the apparent melting temperature in the relative energy scale of the coarse-grained simulations.



**Figure 5.** Non-cooperative unfolding thermodynamics from experiments and simulations. (A) Changes in fluorescence emission maximum of PurR as a function of temperature and the corresponding inflection point of 332 K (dashed line). The melting temperature of 321 K from changes in fluorescence intensity at 295 nm from the same experiment is shown as a continuous black line for reference. (B) FLT amplitudes of PurR for the long (blue) and short (red) lifetime components, together with the crossover temperature of 332 K. (C) Predicted unfolding curves for the different secondary structure elements (helix 1/H1, helix 2/H2, and helix 3/H3) of PurR from the WSME model. The corresponding inflection points from first-derivative analysis are marked at 325, 332, and 328 K in blue, green, and red, respectively. (D) Temperature dependence of the fraction of native contacts ( $Q$ ) for contacts formed by the different protein helices from coarse-grained simulations. (E) Overlay of multiple snapshots corresponding to the partially structured state observed in the simulations at a reaction coordinate value of  $Q \sim 0.5$  (colored tube). We show a transparent cartoon representation of the native state for reference.

profile is broad, upshifted from the expected Freire baseline, exhibits little pretransition, cannot be characterized by a chemical two-state model (Figure S5B), and therefore displays

the features of marginal barrier systems (Figure 4A). The broad DSC curve is not a consequence of the disordered C-terminal tail as the PurR variant without the disordered tail

does not change the overall broadness of the DSC curve and the cooperativity as monitored by far-UV CD or the relaxation rates (Figure S6). The WSME model fits with detailed energetics reproduce the experimental heat capacity profile of the WT PurR very well (see Methods for parameters). The one-dimensional free-energy profiles as a function of the number of structured residues (the reaction coordinate) reveal a downhill-like folding gradient at low temperatures and a maximal thermodynamic barrier of just  $\sim 3$  kJ mol<sup>-1</sup> at 329 K (Figure 4B). The probability distribution at 329 K is broad and without well-defined unfolded or partially structured states (Figure 4C).

To gain further insight into the folding of PurR, we have run coarse-grained simulations using the model by Karanicolas and Brooks.<sup>59</sup> To test for the role of non-native interactions in modulating barrier height magnitudes, we supplemented the original model with the transferable Kim–Hummer potential, which incorporates non-native hydrophobic and electrostatic interactions. Interestingly, the resulting thermogram is broader when non-native states are considered compared to the pure Gō-like potential (Figure 4D). A similar broad thermogram is observed on different realizations of the energy function, where we separately consider the influence of non-native interactions and charges (Figure S7). In all cases, we find little evidence for large thermodynamic barriers, and the resulting barrier height magnitudes (Figures 4E and S7) are consistent with the predictions from the WSME model. This agreement across models with different degrees of structural resolution and energy functions is strong evidence that PurR folds over only small thermodynamic barriers ( $\sim 1$  to  $2 k_B T$ ) and that the slow-folding rate has other molecular origins.

**Non-Cooperative Unfolding and Partially Structured States.** One striking aspect of the free-energy profiles is that they point to three partially structured states (as the barriers are small, we cannot classify them as true intermediates) that are populated *en route* to folded state in the WSME model and at least one partially structured state in the coarse-grained simulations. To further test for this feature from experiments, we performed multiprobe spectroscopic measurements. We identify considerable differences in melting temperatures of  $\sim 5$  K in PurR when analyzed by a two-state model:  $T_m$  of  $321.2 \pm 1.2$  from quantum yield (QY) measurements on excitation at 295 nm (QY<sub>295</sub>) (Figure S8A),  $323.2 \pm 1.0$  from QY measurements at 274 nm excitation (QY<sub>274</sub>) (Figure S8A),  $323.5 \pm 0.5$  K from far-UV CD (Figure 2A), and  $326.9 \pm 0.3$  K from DSC (Figure 4A). Remarkably, the wavelength of maximum fluorescence emission ( $\lambda_{\text{max}}$ ) of the sole tryptophan (W37) displays a transition point of 333 K (from a first-derivative analysis, as the baseline for the unfolded state is not well defined) which is nearly 12 K higher than the melting temperature obtained from QY measurements, that is, from the same data set (Figure 5A). This is more apparent when one observes that  $\lambda_{\text{max}}$  starts changing only at about 320 K, while the heat capacity profile already displays large changes at this temperature.

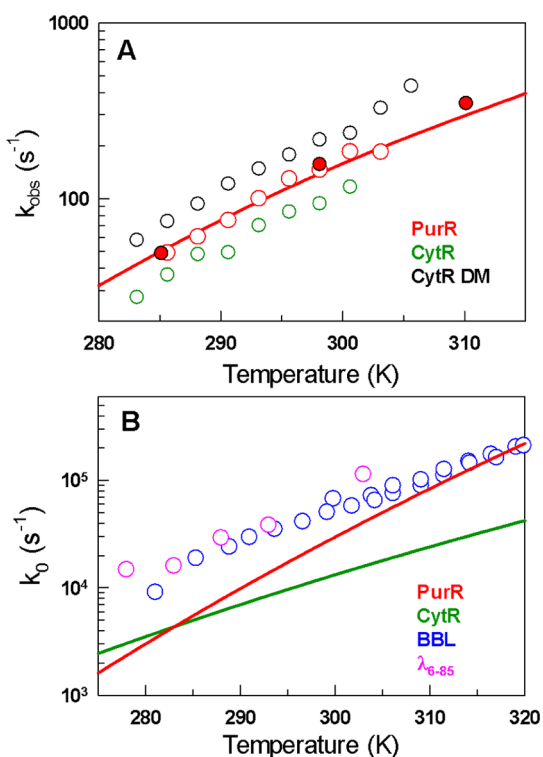
To validate this observation further, we performed fluorescence lifetime (FLT) analysis of W37 located in helix 3 making long-range interactions with helix 2; the advantage of this technique is that the signals and the species population can be directly decoupled without resorting to baselines. FLT measurements reveal two lifetimes for W37, the longer one ( $\sim 7$  ns) corresponding to the main-chain conformation or side-chain orientation sensitive to the folded state and the

shorter one ( $\sim 1$  ns) representing the unfolded conformation (Figure S8B,C). The corresponding amplitudes follow a clear sigmoidal pattern with the amplitude crossover at 332 K (Figure 5B), very similar to the inflection point of fluorescence  $\lambda_{\text{max}}$  changes. The large difference in apparent melting temperatures that is consistent across different experimental probes thus suggests that W37 is differentially sensitive to the folding environment, with the fluorescence intensity (from which QY is calculated) and  $\lambda_{\text{max}}$  sensitive to global and local unfolding events around the tryptophan, respectively.

These experimental observations are strikingly captured by both the WSME model and coarse-grained simulations. The WSME model predicts that the individual helical elements exhibit differences in the overall stability that translates to differences in the melting temperature ranging from 325–332 K, with 332 K being the melting temperature of helix 2 (Figure 5C). Coarse-grained molecular simulations paint a picture of the folding–unfolding equilibrium qualitatively consistent with that of the WSME model. In Figure 5D, we show the average values of  $Q$  calculated for contacts involving helix 1, 2, or 3 (we note that these may include some overlapping pairs of interactions). Clearly, the melting of the interactions formed by the helices is decoupled, with helix 1 being first to unfold, helix 3 being the closest to the average unfolding, and helix 2 being third. We note that the rank order in the melting temperatures is exactly the same as that from the WSME model. The decoupled unfolding manifests as a partially structured state during (un)folding at a reaction coordinate value of  $\sim 0.5$  (Figure 4E) characterized by a detached helix 1 that samples varied conformations (Figure 5E). The emergence of this partially structured state results in different conformational changes being probed preferentially at different temperatures ( $\Phi$ -values for these barrier crossing events calculated from the simulations are shown in Figure S9). At the same time, the unique peak in the heat capacity curve (Figure 4) does not warrant a description in terms of separate thermodynamic transitions.

**Strong Temperature Dependence on the Pre-Exponential to Folding.** In this section, we explore the extent to which the pre-exponential factor described in eq 1 needs to change with temperature to account for the observed slow folding. Since the thermodynamic free-energy profiles are available, it should be possible to extract this pre-exponential factor to protein folding ( $k_0$ ) by measuring the folding relaxation rates at a range of temperatures, similar to earlier work on fast-folding proteins.<sup>73</sup> Note that this approach does not disentangle the different terms in  $k_0$  but serves to only identify its dependence on temperature. We measured the folding relaxation rates of PurR at a range of temperatures from 285–305 K and at a final urea concentration of 0.55 M (near-native conditions; Figure 6A). The measured rates are surprisingly similar to the sampling rate of an excited state in the disordered ensemble of CytR<sup>74</sup> and the folding kinetics of an engineered folded variant CytR A28V/A48M (double mutant, DM) that exhibits an equally complex thermodynamic behavior with equilibrium melting temperature differences of  $\sim 10$  K.<sup>52</sup>

We extract the pre-exponential term by performing diffusive calculations on the free-energy profiles generated by the WSME model (thus accounting for barrier effects) by solving the one-dimensional diffusion equation using the rata-matrix approach of Hofrichter and co-workers.<sup>75</sup> We employ a phenomenological Arrhenius dependence on  $k_0$  as  $Ae^{-E_a/RT}$ ,



**Figure 6.** Strong temperature dependence on the pre-exponential factor to folding. (A) Observed relaxation rates of PurR from stopped-flow experiments at near-native conditions of 0.55 M urea (open red circles) starting from protein denatured in 6 M urea. The corresponding rates from chevron plots (Figures 2B and S2) are shown in filled red circles highlighting the internal consistency across different experiments. The fit from diffusive calculations on the free-energy profiles (Figure 4B) is shown as a red curve. The observed relaxation rates for the sampling of an excited folded conformation in the disordered ensemble of CytR is shown in green. Black circles represent the relaxation rates for the folded variant of CytR, termed the CytR DM (double mutant, A29V/A48M).<sup>52</sup> (B) Measurements of the pre-exponential factor from experiments (blue and magenta) versus that estimated for the folded PurR (red) and the disordered CytR (green).

where  $E_a$  is the activation energy that approximates the landscape roughness including solvent effects and changes in shape of the free-energy profile with temperature.<sup>73</sup> The temperature dependence is captured well only when  $k_0$  varies from  $\sim 3000$  s<sup>-1</sup> at 285 K to  $\sim 200,000$  s<sup>-1</sup> at 320 K (Figure 6B). It is pertinent to note that the former is just an order of magnitude slower than the relaxation rates measured for downhill-folding proteins while agreeing well with the downhill estimates at higher temperatures.

The activation energy is found to be 80 or 1.43 kJ mol<sup>-1</sup> per residue, nearly 40% higher than the average per residue dependence of fast-folding proteins.<sup>73</sup> Interestingly, the pre-exponential terms match with those of the disordered CytR at low temperatures that displays a shallow, bumpy landscape from experiments and simulations.<sup>74</sup>

## DISCUSSION

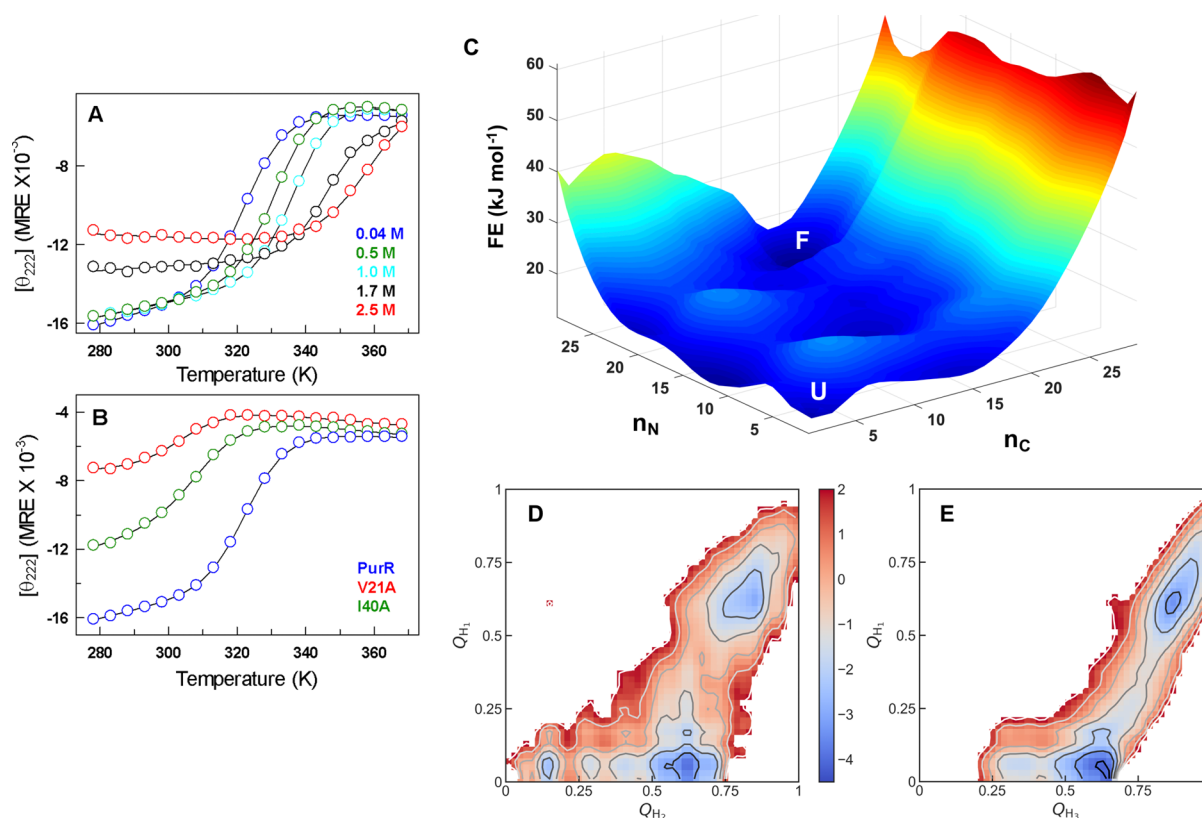
The conformational behavior of PurR is thus observed to be complex, with some attributes from each of downhill (spread in melting temperatures, marginal barriers, and downhill-folding profiles from simulations), two-state folding (chevron kinetics and sigmoidal unfolding curves), and multistate

folding (partially structured states from simulations). These features are captured well at a (semi-)quantitative level by the WSME model and variants of coarse-grained simulations, with the latter including non-native energetic effects. The simulations paint a dynamic picture in which the protein molecule struggles to fit in the helices together as it folds. Accordingly, the three helices exhibit graded thermodynamic stability that in turn lowers the thermodynamic cooperativity and barriers but promotes pockets of partial structure all along the folding coordinate that manifests as differences in melting temperatures and broad heat capacity profiles. In other words, the different structural elements are only weakly coupled, that is, exhibiting large sensitivity to perturbations, contributing to complex underlying landscape. Interestingly, the homologue CytR is disordered in solution, while the engineered folded variant of CytR with two hydrophobic substitutions (A29V/A48M) exhibits very similar features: broad DSC profile, differences in melting temperatures of 10 K, heterogeneous, and slow folding.<sup>52</sup>

What is unique about the structure of PurR that contributes to the conformationally heterogeneous unfolding, despite exhibiting slow two-state like chevron kinetics? In this regard, it is known that DBDs function in a highly complex environment around DNA arising from counterion condensation. It is therefore likely that the structure of PurR is evolutionarily selected for folding and function in the vicinity of DNA and not in the conditions employed in the current set of experiments (43 mM IS). Increasing the solvent IS to mimic the environment around DNA ( $>2$  M IS<sup>76,77</sup>) results in an unusual feature where parts of the PurR structure are lost (less negative ellipticity at 222 nm) despite the increased stability (Figure 7A). This hints at a complex surface electrostatic feature involving a combination of both stabilizing and destabilizing effects and partitioning of local-nonlocal electrostatics, with one effect dominating over the other depending on the conditions. The folding relaxation rates increase on increasing the solvent IS but fold faster than the dead-time of the stopped flow-instrument even at 298 K at IS greater than 0.5 M (Figure S2). In fact, our observations are consistent with single-molecule experiments on  $\alpha_3D$  that point to conflicting electrostatic interactions as the primary source of internal friction or a slower pre-exponential to folding.<sup>49</sup> Recent statistical mechanical modeling of PurR folding in the presence of DNA demonstrates a progressive titling of its landscape toward the folded state in the vicinity of DNA, providing hints that the folding landscape could be fine-tuned by quinary interactions.<sup>78</sup>

Additionally, protein engineering experiments highlight an unusual packing thermodynamics with single-point mutations V21A and I40A strongly destabilizing the protein (Figure 7B).<sup>52</sup> In fact, the V21A mutation fully unfolds PurR, arguing for a weak hydrophobic effect driving folding in this system. This observation is borne out by the fact that a close family member, CytR, is disordered in solution and is a mere two hydrophobic substitutions from being folded at low temperatures.<sup>26,52</sup>

Taken together, the molecular origin of slow folding in PurR seems to be a combination of multiple sequence-structural features that in turn affect the pre-exponential to folding and not because of large thermodynamic barrier height. We find evidence for this from a steep temperature dependence of relaxation rates and the extracted pre-exponential terms that match with downhill folding proteins at higher temperatures



**Figure 7.** Electrostatic frustration and weak packing contribute to a shallow folding landscape. (A) Thermal unfolding curves of PurR at different IS conditions as monitored by far-UV CD at 222 nm and reported in mean-residue ellipticity units of  $\text{deg cm}^2 \text{dmol}^{-1}$ . Note the loss of secondary structure at high IS (black and red) despite an increase in the melting temperature. (B) Thermal unfolding curves of PurR mutants V21A (red) and I40A (green).<sup>52</sup> (C) Two-dimensional conformational landscape of PurR predicted by the WSME model at 329 K highlighting the lack of large thermodynamic barriers between folded (F), unfolded (U), and the numerous partially structured states (valleys in dark blue shade).  $n_N$  and  $n_C$  represent the number of structured residues in the N- and C-terminals, respectively.  $n_N$  includes the first two helices, while the rest of the sequence falls in  $n_C$ . (D,E) Two-dimensional free-energy landscape (in units of  $k_B T$ ) of PurR from coarse-grained simulations pointing to a complex conformational ensemble with multiple valleys (light and dark blue contours). The coordinates employed are the fraction of native contacts in helix 1 ( $Q_{H1}$ ), helix 2 ( $Q_{H2}$ ), and helix 3 ( $Q_{H3}$ ).

and are slower by an order of magnitude at lower temperatures. Additionally, the free-energy landscape of PurR is complex when assuming two conformational coordinates from the WSME model or coarse-grained simulations. In the former, it is seen that nearly all conformations are equally likely with a broad distribution of molten-globule-like conformations at the midpoint with no well-defined funnel toward the folded state (Figure 7C). This is also observed in two-dimensional probability density plots from coarse-grained simulations, wherein partially structured states are populated irrespective of the projections (Figure 7D,E; see also contact maps in Figure S10).

Experiments on unfolded protein L under folding conditions highlight a dramatic slowdown in folding diffusion coefficient compared to unfolding conditions,<sup>79</sup> while single-molecule experiments point to large contributions from internal friction on even unfolded and disordered proteins.<sup>80</sup> Given these observations and since internal friction effects arise likely from microscopic barriers to dihedral motions,<sup>34</sup> it is surprising to find that conventional measures of internal friction fail to reveal any landscape roughness or apparent internal friction in PurR. If the established techniques for “measuring” internal friction are robust (Figure 3), assuming a one-dimensional coordinate, it can therefore be concluded that the folding landscape of PurR is characterized by broad unfolded and

folded basins with no well-defined transition-state ensemble. In terms of Kramers theory of reaction rates, these features hint at small curvatures ( $\omega^2$ ) of unfolded well and barrier top in PurR, which in turn slow down the folding rate and not through a slow-folding diffusion coefficient. Experimental works narrowing the barrier increase the folding rate by an order of magnitude,<sup>81</sup> thus suggesting that broader barriers or unfolded wells would proportionately slow down folding. An alternate possibility is that denaturant-stabilizer mixtures (as in urea-glucose use in the current study) modulate both the intramolecular diffusion coefficient and barriers in compensatory fashions, resulting in a linear relative viscosity versus rate plot (Figure 3). It is important to note that the extent of trapping in the populated partially structured states could also be different that necessitates the use of a coordinate dependent diffusion coefficient or “heterogeneous friction.”<sup>67</sup> A two-dimensional landscape, on the other hand, could contribute to complex distribution of folding fluxes, again requiring additional slow diffusional terms. Experiments on mutants that populate one or more of the partially structured states could be specifically employed to probe for such coordinate-dependent effects.



## CONCLUSIONS

In summary, our experiments combined with simulations rule out a large thermodynamic barrier in PurR and point to the slaving of the dynamics to the shape of the underlying free-energy landscape that appears as a strong temperature dependence on the pre-exponential, slowing down folding. Structural, mutational, and thermodynamic analyses of packing and electrostatic effects indicate that PurR is at the threshold of disorder. We also find that studies combining scanning calorimetry experiments, multiprobe spectroscopy, viscosity-dependent kinetics, and thermodynamic modeling can provide an unparalleled view on the underlying folding landscape of proteins. It still remains to be seen if the slow folding determined by the slow pre-exponential is a conserved feature of LacR/PurR family members. The suitability of additional coordinates in explaining slow folding and the role of DNA in smoothening the folding landscape or speeding up the folding pre-exponential through its large negative electrostatic potential also remain to be seen.

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jpcb.0c05976>.

Electrostatic potential maps of the PurR family members, IS-urea-viscogen-dependent kinetics, apparent heat capacity of PurR WT and the truncated version, two-state fits to the DSC curves, coarse-grained simulation results, equilibrium fluorescence, and FLT measurements (PDF)

## AUTHOR INFORMATION

### Corresponding Authors

**Athi N. Naganathan** – Department of Biotechnology, Bhupat & Jyoti Mehta School of Biosciences, Indian Institute of Technology Madras, Chennai 600036, India; [orcid.org/0000-0002-1655-7802](https://orcid.org/0000-0002-1655-7802); Email: [athi@iitm.ac.in](mailto:athi@iitm.ac.in)

**David de Sancho** – Polímero eta Material Aurreratuak: Fisika, Kimika eta Teknologia, Euskal Herriko Unibertsitatea UPV/EHU, Donostia-San Sebastián 20080, Spain; Donostia International Physics Center (DIPC), Donostia-San Sebastián 20080, Spain; [orcid.org/0000-0002-8985-2685](https://orcid.org/0000-0002-8985-2685); Phone: +91-44-2257 4140; Email: [david.desancho@ehu.eus](mailto:david.desancho@ehu.eus)

### Authors

**Sandhya Subramanian** – Department of Biotechnology, Bhupat & Jyoti Mehta School of Biosciences, Indian Institute of Technology Madras, Chennai 600036, India

**Hemashree Golla** – Department of Biotechnology, Bhupat & Jyoti Mehta School of Biosciences, Indian Institute of Technology Madras, Chennai 600036, India

**Kalivarathan Divakar** – Department of Biotechnology, National Institute of Technology Warangal, Warangal 506004, India

**Adithi Kannan** – Department of Biotechnology, Bhupat & Jyoti Mehta School of Biosciences, Indian Institute of Technology Madras, Chennai 600036, India

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acs.jpcb.0c05976>

### Author Contributions

<sup>†</sup>S.S. and H.G. contributed equally.

## Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

This work was supported by the Wellcome Trust/DBT India Alliance Fellowship IA/I/15/1/501837 awarded to A.N.N. The authors acknowledge the FIST facility sponsored by the Department of Science and Technology (DST), India at the Department of Biotechnology, IITM for the instrumentation. The authors thank Dr. Ramesh L. Gardas for providing access to the viscometer and Somenath Pandey for help with viscosity measurements. Financial support to D.D.S. comes from Eusko Jaurlaritza (Basque Government) through Project IT588-13 and from Grants RYC-2016-19590 and PGC2018-099321-B-I00 from the Spanish Ministry of Science and Universities through the Office of Science Research (MINECO/FEDER).

## ABBREVIATIONS

CD circular dichroism  
DSC differential scanning calorimetry  
CG coarse grained  
WSME Wako-Saitô-Muñoz-Eaton

## REFERENCES

- (1) Schreiber, G.; Buckle, A. M.; Fersht, A. R. Stability and function: two constraints in the evolution of barstar and other proteins. *Structure* **1994**, *2*, 945–951.
- (2) Strop, P.; Mayo, S. L. Contribution of surface salt bridges to protein stability. *Biochemistry* **2000**, *39*, 1251–1255.
- (3) Guerois, R.; Nielsen, J. E.; Serrano, L. Predicting changes in the stability of proteins and protein complexes: A study of more than 1000 mutations. *J. Mol. Biol.* **2002**, *320*, 369–387.
- (4) Strickler, S. S.; Gribenko, A. V.; Gribenko, A. V.; Keiffer, T. R.; Tomlinson, J.; Reihle, T.; Loladze, V. V.; Makhatazde, G. I. Protein stability and surface electrostatics: A charged relationship. *Biochemistry* **2006**, *45*, 2761–2766.
- (5) Tokuriki, N.; Tawfik, D. S. Stability effects of mutations and protein evolvability. *Curr. Opin. Struct. Biol.* **2009**, *19*, 596–604.
- (6) Monteith, W. B.; Cohen, R. D.; Smith, A. E.; Guzman-Cisneros, E.; Pielak, G. J. Quinary structure modulates protein stability in cells. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112*, 1739–1742.
- (7) Gupta, K.; Varadarajan, R. Insights into protein structure, stability and function from saturation mutagenesis. *Curr. Opin. Struct. Biol.* **2018**, *50*, 117–125.
- (8) Modi, T.; Huihui, J.; Ghosh, K.; Ozkan, S. B. Ancient thioredoxins evolved to modern-day stability-function requirement by altering native state ensemble. *Philos. Trans. R. Soc. London, Ser. B* **2018**, *373*, 20170184.
- (9) Rajasekaran, N.; Suresh, S.; Gopi, S.; Raman, K.; Naganathan, A. N. A general mechanism for the propagation of mutational effects in proteins. *Biochemistry* **2017**, *56*, 294–305.
- (10) McLaughlin, R. N., Jr.; Poelwijk, F. J.; Raman, A.; Gosal, W. S.; Ranganathan, R. The spatial architecture of protein function and adaptation. *Nature* **2012**, *491*, 138–142.
- (11) Liberles, D. A.; Teichmann, S. A.; Bahar, I.; Bastolla, U.; Bloom, J.; Bornberg-Bauer, E.; Colwell, L. J.; de Koning, A. P. J.; Dokholyan, N. V.; Echave, J.; Elofsson, A.; Gerloff, D. L.; Goldstein, R. A.; Grahnen, J. A.; Holder, M. T.; Lakner, C.; Lartillot, N.; Lovell, S. C.; Naylor, G.; Perica, T.; Pollock, D. D.; Pupko, T.; Regan, L.; Roger, A.; Rubinstein, N.; Shakhnovich, E.; Sjölander, K.; Sunyaev, S.; Teufel, A. I.; Thorne, J. L.; Thornton, J. W.; Weinreich, D. M.; Whelan, S. The interface of protein structure, protein biophysics, and molecular evolution. *Protein Sci.* **2012**, *21*, 769–785.
- (12) Jack, B. R.; Meyer, A. G.; Echave, J.; Wilke, C. O. Functional Sites Induce Long-Range Evolutionary Constraints in Enzymes. *PLoS Biol.* **2016**, *14*, e1002452.

- (13) Bershtein, S.; Serohijos, A. W.; Shakhnovich, E. I. Bridging the physical scales in evolutionary biology: from protein sequence space to fitness of organisms and populations. *Curr. Opin. Struct. Biol.* **2017**, *42*, 31–40.
- (14) Naganathan, A. N. Modulation of allosteric coupling by mutations: from protein dynamics and packing to altered native ensembles and function. *Curr. Opin. Struct. Biol.* **2019**, *54*, 1–9.
- (15) Lohman, T. M.; Bujalowski, W. Negative cooperativity within individual tetramers of *Escherichia coli* single strand binding protein is responsible for the transition between the (SSB)35 and (SSB)56 DNA binding modes. *Biochemistry* **1988**, *27*, 2260–2265.
- (16) Crane-Robinson, C.; Read, C. M.; Cary, P. D.; Driscoll, P. C.; Dragan, A. I.; Privalov, P. L. The energetics of HMG box interactions with DNA. Thermodynamic description of the box from mouse Sox-5. *J. Mol. Biol.* **1998**, *281*, 705–717.
- (17) Jelesarov, I.; Crane-Robinson, C.; Privalov, P. L. The energetics of HMG box interactions with DNA: thermodynamic description of the target DNA duplexes. *J. Mol. Biol.* **1999**, *294*, 981–995.
- (18) Dragan, A. I.; Read, C. M.; Makeyeva, E. N.; Milgotina, E. I.; Churchill, M. E. A.; Crane-Robinson, C.; Privalov, P. L. DNA binding and bending by HMG boxes: energetic determinants of specificity. *J. Mol. Biol.* **2004**, *343*, 371–393.
- (19) Religa, T. L.; Markson, J. S.; Mayor, U.; Freund, S. M. V.; Fersht, A. R. Solution structure of a protein denatured state and folding intermediate. *Nature* **2005**, *437*, 1053–1056.
- (20) Roy, R.; Kozlov, A. G.; Lohman, T. M.; Ha, T. Dynamic structural rearrangements between DNA binding modes of *E. coli* SSB protein. *J. Mol. Biol.* **2007**, *369*, 1244–1257.
- (21) Privalov, P. L.; Dragan, A. I.; Crane-Robinson, C. Interpreting protein/DNA interactions: distinguishing specific from non-specific and electrostatic from non-electrostatic components. *Nucleic Acids Res.* **2011**, *39*, 2483–2491.
- (22) Zandarashvili, L.; Esadze, A.; Vuzman, D.; Kemme, C. A.; Levy, Y.; Iwahara, J. Balancing between affinity and speed in target DNA search by zinc-finger proteins via modulation of dynamic conformational ensemble. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112*, E5142–E5149.
- (23) Esadze, A.; Chen, C.; Zandarashvili, L.; Roy, S.; Pettitt, B. M.; Iwahara, J. Changes in conformational dynamics of basic side chains upon protein-DNA association. *Nucleic Acids Res.* **2016**, *44*, 6961–6970.
- (24) Nguyen, D.; Hoffpauir, Z. A.; Iwahara, J. Internal Motions of Basic Side Chains of the Antennapedia Homeodomain in the Free and DNA-Bound States. *Biochemistry* **2017**, *56*, 5866–5869.
- (25) Munshi, S.; Gopi, S.; Asampille, G.; Subramanian, S.; Campos, L. A.; Atreya, H. S.; Naganathan, A. N. Tunable order-disorder continuum in protein-DNA interactions. *Nucleic Acids Res.* **2018**, *46*, 8700–8709.
- (26) Munshi, S.; Gopi, S.; Subramanian, S.; Campos, L. A.; Naganathan, A. N. Protein plasticity driven by disorder and collapse governs the heterogeneous binding of CytR to DNA. *Nucleic Acids Res.* **2018**, *46*, 4044–4053.
- (27) Bryngelson, J. D.; Onuchic, J. N.; Socci, N. D.; Wolynes, P. G. Funnels, Pathways, and the Energy Landscape of Protein-Folding - a Synthesis. *Proteins: Struct., Funct., Genet.* **1995**, *21*, 167–195.
- (28) Ferreiro, D. U.; Komives, E. A.; Wolynes, P. G. Frustration in biomolecules. *Q. Rev. Biophys.* **2014**, *47*, 285–363.
- (29) Kramers, H. A. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* **1940**, *7*, 284–304.
- (30) Ansari, A.; Jones, C.; Henry, E.; Hofrichter, J.; Eaton, W. The role of solvent viscosity in the dynamics of protein conformational changes. *Science* **1992**, *256*, 1796–1798.
- (31) Chavez, L. L.; Onuchic, J. N.; Clementi, C. Quantifying the roughness on the free energy landscape: Entropic bottlenecks and protein folding rates. *J. Am. Chem. Soc.* **2004**, *126*, 8426–8432.
- (32) Chung, H. S.; Eaton, W. A. Single-molecule fluorescence probes dynamics of barrier crossing. *Nature* **2013**, *502*, 685–688.
- (33) Sashi, P.; Ramakrishna, D.; Bhuyan, A. K. Dispersion Forces and the Molecular Origin of Internal Friction in Protein. *Biochemistry* **2016**, *55*, 4595–4602.
- (34) de Sancho, D.; Sirur, A.; Best, R. B. Molecular origins of internal friction effects on protein-folding rates. *Nat. Commun.* **2014**, *5*, 4307.
- (35) Echeverria, I.; Makarov, D. E.; Papoian, G. A. Concerted Dihedral Rotations Give Rise to Internal Friction in Unfolded Proteins. *J. Am. Chem. Soc.* **2014**, *136*, 8708–8713.
- (36) Zarrine-Afsar, A.; Wallin, S.; Neculai, A. M.; Neudecker, P.; Howell, P. L.; Davidson, A. R.; Chan, H. S. Theoretical and experimental demonstration of the importance of specific nonnative interactions in protein folding. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 9999–10004.
- (37) Korzhnev, D. M.; Vernon, R. M.; Religa, T. L.; Hansen, A. L.; Baker, D.; Fersht, A. R.; Kay, L. E. Nonnative interactions in the FF domain folding pathway from an atomic resolution structure of a sparsely populated intermediate: an NMR relaxation dispersion study. *J. Am. Chem. Soc.* **2011**, *133*, 10974–10982.
- (38) Yoo, T. Y.; Adhikari, A.; Xia, Z.; Huynh, T.; Freed, K. F.; Zhou, R.; Sosnick, T. R. The folding transition state of protein L is extensive with nonnative interactions (and not small and polarized). *J. Mol. Biol.* **2012**, *420*, 220–234.
- (39) Zheng, W.; Hofmann, H.; Schuler, B.; Best, R. B. Origin of Internal Friction in Disordered Proteins Depends on Solvent Quality. *J. Phys. Chem. B* **2018**, *122*, 11478–11487.
- (40) Muñoz, V.; Sadqi, M.; Naganathan, A. N.; de Sancho, D. Exploiting the downhill folding regime via experiment. *HFSP J.* **2008**, *2*, 342–353.
- (41) Yang, W. Y.; Gruebele, M. Folding at the speed limit. *Nature* **2003**, *423*, 193–197.
- (42) Fung, A.; Li, P.; Godoy-Ruiz, R.; Sanchez-Ruiz, J. M.; Muñoz, V. Expanding the realm of ultrafast protein folding: gpW, a midsize natural single-domain with alpha+beta topology that folds downhill. *J. Am. Chem. Soc.* **2008**, *130*, 7489–7495.
- (43) Li, P.; Oliva, F. Y.; Naganathan, A. N.; Muñoz, V. Dynamics of one-state downhill protein folding. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 103–108.
- (44) DeCamp, S. J.; Naganathan, A. N.; Waldauer, S. A.; Bakajin, O.; Lapidus, L. J. Direct Observation of Downhill Folding of lambda-Repressor in a Microfluidic Mixer. *Biophys. J.* **2009**, *97*, 1772–1777.
- (45) Naganathan, A. N.; Li, P.; Perez-Jimenez, R.; Sanchez-Ruiz, J. M.; Muñoz, V. Navigating the downhill protein folding regime via structural homologues. *J. Am. Chem. Soc.* **2010**, *132*, 11183–11190.
- (46) Naganathan, A. N.; Sanchez-Ruiz, J. M.; Muñoz, V. Direct measurement of barrier heights in protein folding. *J. Am. Chem. Soc.* **2005**, *127*, 17970–17971.
- (47) Kluber, A.; Burt, T. A.; Clementi, C. Size and topology modulate the effects of frustration in protein folding. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115*, 9234–9239.
- (48) Wensley, B. G.; Batey, S.; Bone, F. A. C.; Chan, Z. M.; Tumelty, N. R.; Steward, A.; Kwa, L. G.; Borgia, A.; Clarke, J. Experimental evidence for a frustrated energy landscape in a three-helix-bundle protein family. *Nature* **2010**, *463*, 685–688.
- (49) Chung, H. S.; Piana-Agostinetti, S.; Shaw, D. E.; Eaton, W. A. Structural origin of slow diffusion in protein folding. *Science* **2015**, *349*, 1504–1510.
- (50) Szczepaniak, M.; Iglesias-Bexiga, M.; Cerminara, M.; Sadqi, M.; Sanchez de Medina, C.; Martinez, J. C.; Luque, I.; Muñoz, V. Ultrafast folding kinetics of WW domains reveal how the amino acid sequence determines the speed limit to protein folding. *Proc. Natl. Acad. Sci. U.S.A.* **2019**, *116*, 8137–8142.
- (51) Moody, C. L.; Tretyachenko-Ladokhina, V.; Laue, T. M.; Senear, D. F.; Cocco, M. J. Multiple Conformations of the Cytidine Repressor DNA-Binding Domain Coalesce to One upon Recognition of a Specific DNA Surface. *Biochemistry* **2011**, *50*, 6622–6632.
- (52) Munshi, S.; Subramanian, S.; Ramesh, S.; Golla, H.; Kalivarathan, D.; Kulkarni, M.; Campos, L. A.; Sekhar, A.;

Naganathan, A. N. Engineering Order and Cooperativity in a Disordered Protein. *Biochemistry* **2019**, *58*, 2389–2397.

(53) Guzman-Casado, M.; Parody-Morreale, A.; Robic, S.; Marqusee, S.; Sanchez-Ruiz, J. M. Energetic evidence for formation of a pH-dependent hydrophobic cluster in the denatured state of *Thermus thermophilus* ribonuclease H. *J. Mol. Biol.* **2003**, *329*, 731–743.

(54) Wako, H.; Saitô, N. Statistical Mechanical Theory of Protein Conformation. 2. Folding Pathway for Protein. *J. Phys. Soc. Jpn.* **1978**, *44*, 1939–1945.

(55) Muñoz, V.; Eaton, W. A. A simple model for calculating the kinetics of protein folding from three-dimensional structures. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 11311–11316.

(56) Naganathan, A. N. Predictions from an Ising-like Statistical Mechanical Model on the Dynamic and Thermodynamic Effects of Protein Surface Electrostatics. *J. Chem. Theory Comput.* **2012**, *8*, 4646–4656.

(57) Rajasekaran, N.; Gopi, S.; Narayan, A.; Naganathan, A. N. Quantifying Protein Disorder through Measures of Excess Conformational Entropy. *J. Phys. Chem. B* **2016**, *120*, 4341–4350.

(58) Heinig, M.; Frishman, D. STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins. *Nucleic Acids Res.* **2004**, *32*, W500–W502.

(59) Karanicolas, J.; Brooks, C. L., 3rd The origins of asymmetry in the folding transition states of protein L and protein G. *Protein Sci.* **2002**, *11*, 2351–2361.

(60) Kim, Y. C.; Hummer, G. Coarse-grained models for simulations of multiprotein complexes: application to ubiquitin binding. *J. Mol. Biol.* **2008**, *375*, 1416–1433.

(61) De Sancho, D.; Best, R. B. Modulation of an IDP binding mechanism and rates by helix propensity and non-native interactions: association of HIF1 alpha with CBP. *Mol. BioSyst.* **2012**, *8*, 256–267.

(62) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.

(63) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules. I. The Method. *J. Comput. Chem.* **1992**, *13*, 1011–1021.

(64) Gromiha, M. M.; Selvaraj, S. Comparison between long-range interactions and contact order in determining the folding rate of two-state proteins: application of long-range order to folding rate prediction. *J. Mol. Biol.* **2001**, *310*, 27–32.

(65) De Sancho, D.; Muñoz, V. Integrated prediction of protein folding and unfolding rates from only size and structural class. *Phys. Chem. Chem. Phys.* **2011**, *13*, 17030–17043.

(66) Cellmer, T.; Henry, E. R.; Hofrichter, J.; Eaton, W. A. Measuring internal friction of an ultrafast-folding protein. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 18320–18325.

(67) Hagen, S. J. Solvent viscosity and friction in protein folding dynamics. *Curr. Protein Pept. Sci.* **2010**, *11*, 385–395.

(68) Muñoz, V.; Sanchez-Ruiz, J. M. Exploring protein folding ensembles: a variable barrier model for the analysis of equilibrium unfolding experiments. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 17646–17651.

(69) Cooper, A. Thermodynamic Fluctuations in Protein Molecules. *Proc. Natl. Acad. Sci. U.S.A.* **1976**, *73*, 2740–2741.

(70) Chan, H. S. Modeling protein density of states: additive hydrophobic effects are insufficient for calorimetric two-state cooperativity. *Proteins* **2000**, *40*, 543–571.

(71) Sanchez-Ruiz, J. M. Probing free-energy surfaces with differential scanning calorimetry. *Annu. Rev. Phys. Chem.* **2011**, *62*, 231–255.

(72) Ibarra-Molero, B.; Naganathan, A. N.; Sanchez-Ruiz, J. M.; Muñoz, V. Modern Analysis of Protein Folding by Differential Scanning Calorimetry. *Methods Enzymol.* **2016**, *567*, 281–318.

(73) Naganathan, A. N.; Doshi, U.; Muñoz, V. Protein folding kinetics: Barrier effects in chemical and thermal denaturation experiments. *J. Am. Chem. Soc.* **2007**, *129*, 5673–5682.

(74) Munshi, S.; Rajendran, D.; Naganathan, A. N. Entropic Control of an Excited Folded-Like Conformation in a Disordered Protein Ensemble. *J. Mol. Biol.* **2018**, *430*, 2688–2694.

(75) Lapidus, L. J.; Steinbach, P. J.; Eaton, W. A.; Szabo, A.; Hofrichter, J. Effects of chain stiffness on the dynamics of loop formation in polypeptides. Appendix: Testing a 1-dimensional diffusion model for peptide dynamics. *J. Phys. Chem. B* **2002**, *106*, 11628–11640.

(76) Jayaram, B.; Sharp, K. A.; Honig, B. The electrostatic potential of B-DNA. *Biopolymers* **1989**, *28*, 975–993.

(77) Allred, B. E.; Gebala, M.; Herschlag, D. Determination of Ion Atmosphere Effects on the Nucleic Acid Electrostatic Potential and Ligand Association Using AH(+).C Wobble Formation in Double-Stranded DNA. *J. Am. Chem. Soc.* **2017**, *139*, 7540–7548.

(78) Gopi, S.; Naganathan, A. N. Non-specific DNA-driven quinary interactions promote structural transitions in proteins. *Phys. Chem. Chem. Phys.* **2020**, *22*, 12671–12677.

(79) Waldauer, S. A.; Bakajin, O.; Lapidus, L. J. Extremely slow intramolecular diffusion in unfolded protein L. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 13713–13717.

(80) Soranno, A.; Buchli, B.; Nettels, D.; Cheng, R. R.; Muller-Spath, S.; Pfeil, S. H.; Hoffmann, A.; Lipman, E. A.; Makarov, D. E.; Schuler, B. Quantifying internal friction in unfolded and intrinsically disordered proteins with single-molecule spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 17800–17806.

(81) Abaskharon, R. M.; Culik, R. M.; Woolley, G. A.; Gai, F. Tuning the Attempt Frequency of Protein Folding Dynamics via Transition-State Rigidification: Application to Trp-Cage. *J. Phys. Chem. Lett.* **2015**, *6*, 521–526.