THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

Cite This: J. Phys. Chem. B 2018, 122, 11030−11038

pubs.acs.org/JPCB

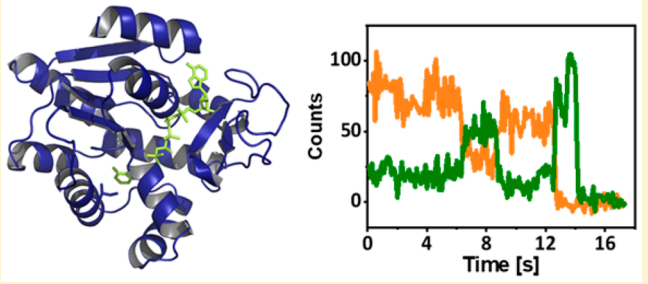
Manipulating the Folding Landscape of a Multidomain Protein

Raisa Kantaev,[†] Inbal Riven,[†] Adi Goldenzweig,[‡] Yoav Barak,[§] Orly Dym,[∥] Yoav Peleg,[∥] Shira Albeck,[∥] Sarel J. Fleishman,[‡] and Gilad Haran*[,†]

[†]Department of Chemical Physics, [‡]Department of Biomolecular Sciences, [§]Department of Chemical Research Support, and [∥]Center for Structural Proteomics, Weizmann Institute of Science, Rehovot 76100, Israel

Supporting Information

ABSTRACT: Folding of proteins to their functional conformation is paramount to life. Though 75% of the proteome consists of multidomain proteins, our knowledge of folding has been based primarily on studies conducted on single-domain and fast-folding proteins. Nonetheless, the complexity of folding landscapes exhibited by multidomain proteins has received increased scrutiny in recent years. We study the three-domain protein adenylate kinase from *E. coli* (AK), which has been shown to fold through a series of pathways involving several intermediate states. We use a protein design method to manipulate the folding landscape of AK, and single-molecule FRET spectroscopy to study the effects on the folding process. Mutations introduced in the NMP binding (NMPbind) domain of the protein are found to have unexpected effects on the folding landscape. Thus, while stabilizing mutations in the core of the NMPbind domain retain the main folding pathways of wild-type AK, a destabilizing mutation at the interface between the NMPbind and the CORE domains causes a significant repartition of the flux between the folding pathways. Our results demonstrate the outstanding plasticity of the folding landscape of AK and reveal how specific mutations in the primary structure are translated into changes in folding dynamics. The combination of methodologies introduced in this work should prove useful for deepening our understanding of the folding process of multidomain proteins.



INTRODUCTION

How the amino acid sequence of a protein dictates its three-dimensional atomic structure has been one of the fundamental and challenging questions in protein science for more than 50 years.[1,2] The protein folding field has gained biomedical relevance with the discovery of a growing number of human protein misfolding diseases.[3,4] A major guiding concept in the protein folding field in recent years has been the funnel-shaped multidimensional rugged folding energy landscape, which was introduced on the basis of a combination of experimental data, statistical physical theory, and simulations.[5,6] Identifying the major pathways on the folding landscape of a protein and the nature of intermediate states it samples has therefore been of major interest.[7] Sophisticated single-molecule experiments, using fluorescence and force spectroscopy, have provided much insight into these issues in recent years.[8−13]

Despite the fact that over 75% of eukaryotic proteomes consist of multidomain proteins with variable topological complexity, much of the experimental and theoretical work on protein folding has involved small single-domain proteins or isolated domains, in which the folding landscape is rather smooth.[14,15] In recent years there has been a growing effort to develop methodologies for studying the folding principles of proteins with more than a single domain. Such proteins are often found to fold by a hierarchical assembly of substructures involving multiple time-scales. The variability in domain connectivity in different proteins and the nature of the interface between them are important factors in determining their folding process.[16−18]

Adenylate kinase (AK) is a three-domain protein that consists of a central large CORE domain, which is split into three-discontinuous segments by two inserted domains (Figure 1): the LID domain on one side (residues 118−160) and the NMPbind domain on the other side (residues 30−76). The folding process of AK has been studied experimentally[19−26] and theoretically.[27,28] The complex folding of AK, which can be attributed to its intricate topology, was demonstrated early on by Zhang et al.[21] Extensive work by Haas and co-workers demonstrated differences in folding dynamics of selected subdomains and secondary structure elements.[20,22−24] Single-molecule FRET (smFRET) experiments from our own lab revealed that AK folding involves multiple parallel pathways and a series of intermediate states.[25,26,29] Li et al.[27] used coarse-grained simulations to generate folding trajectories of AK that were consistent with the experimental results and also provided detailed structural information for the intermediates. Interestingly, the simulations showed that the NMPbind domain remained unfolded along the main folding pathways

Special Issue: William A. Eaton Festschrift

Received: May 21, 2018
Revised: August 7, 2018
Published: August 8, 2018

DOI: 10.1021/acs.jpcb.8b04834
J. Phys. Chem. B 2018, 122, 11030−11038

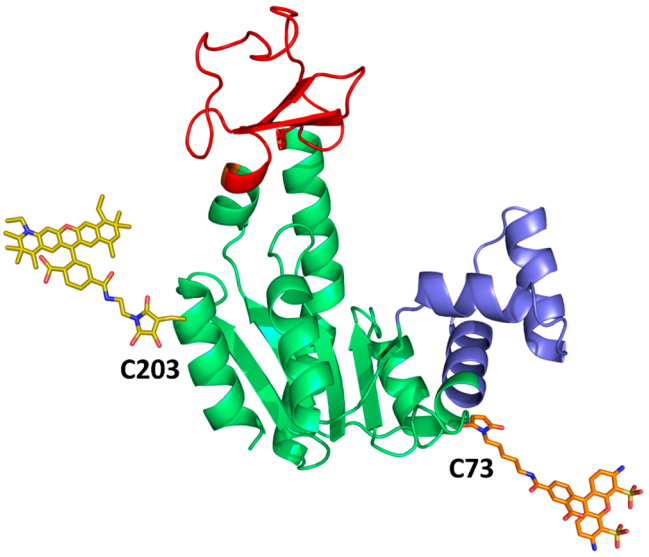


Figure 1. Structure of adenylate kinase from *E. coli*. CORE domain in green, LID domain in red, and NMPbind domain in blue. Two fluorescent dyes, Alexa488 and Atto590, are attached at positions 73 and 203, respectively, for measuring FRET. Based on Protein Data Bank entry 4AKE.

and attained its native fold at a very late stage of the folding process. Another computational work, by Giri Rao et al.,[28] indicated a similar structural instability of the NMPbind domain.

On the basis of these observations, we set out to investigate how the interaction between AK's inserted NMPbind domain and its CORE domain affects the folding landscape of the protein. Using Rosetta design,[30] we identified a set of mutations that affect the stability of the NMPbind domain. We selected two variants for further study, and used smFRET spectroscopy on immobilized molecules, combined with hidden Markov model (HMM) analysis,[31] to reconstruct their main folding pathways. We found that a significant repartition of folding pathways can be induced by a single mutation at the interface between the NMPbind and CORE domains.

RESULTS

1. AK Variants: Engineering and Selection. To stabilize the NMPbind domain of AK, we subjected the protein to atomistic design in Rosetta, guided by phylogenetic constraints derived from AK homologues (Supporting Information Methods, section 1). The design protocol comprised iterations of sequence optimization followed by backbone minimization.[32] At each amino acid position the energy function was augmented with a biasing term to favor mutations that appeared more frequently in homologous sequences at that position, similar to consensus-design approaches.[32]

The designed mutations were predicted to optimize different aspects of the NMPbind stability, including core packing, improved compatibility between the amino acid sequence and the secondary structure, and electrostatic interactions. We selected four designs comprising 1−3 mutations each for experimental testing (Figure S1A and Supporting Information Methods, sections 2 and 3). We first tested their relative thermal stability (Figure S1B) and activity (Figure S1C), and selected two variants that showed the largest change in thermal stability (Figure 2 and Figure S1B): (1) a surface mutation, K69R, showed a reduction of 2.2 °C in the thermal stability compared to wild-type (designated AK_{ds}, for destabilized); and (2) a triple mutant, V64I/L63I/V39I (AK_s, for stabilized), exhibited an increase of 3.5 °C compared to wild-type (WT), the highest thermal stability with respect to other tested variants (Table S1). Chemical denaturation studies (Figure 2B, see also SI Methods, section 6) revealed that AK_s exhibited an increase of the folding free energy by over 50%. This pronounced thermodynamic stabilization of 1.9 ± 0.3 kcal·mol^{-1} (Table S1) is attributed to the enhanced hydrophobic core packing in the NMPbind domain interior. In contrast, AK_{ds} is destabilized by 0.8 ± 0.3 kcal mol^{-1} (Table S1), a decrease of ∼20%. The preservation of native structure and function of the selected variants was verified by CD spectroscopy, X-ray crystallography and enzymatic activity assays (Supporting Information Methods sections 4, 5, and 7). In particular, crystal structures in the closed (holo) forms of the two mutants (PDB codes: AK_{ds}- 6HAP, AK_s- 6HAM) were found to be very similar to the equivalent WT crystal structure. Nevertheless, we could observe some difference in the loop region (residues 39−49) of AK_{ds} with respect to WT (Figure S2).

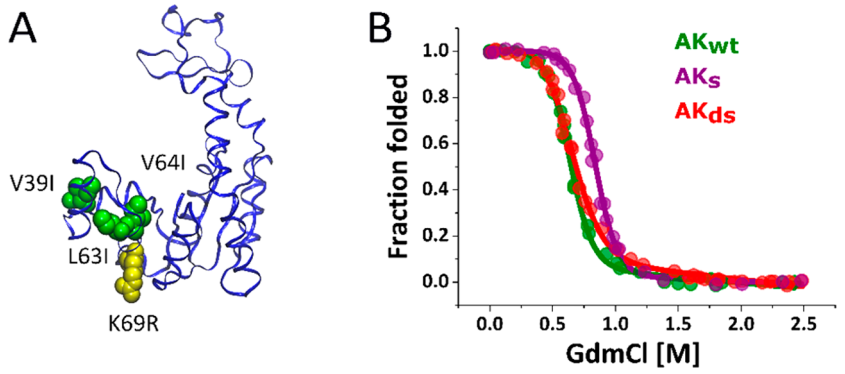


Figure 2. Effect of mutations on AK stability. (A) A view of AK showing mutations in the NMPbind domain that were selected for folding experiments: the single-point design AK_{ds} (K69R, yellow spheres) and the triple-mutation design AK_s (V64I/L63I/V39I, green spheres). (B) Bulk chemical denaturation curves for three AK variants were obtained by measuring fluorescence spectra of double-labeled protein molecules, from which FRET efficiency values and folded fractions were calculated.

DOI: 10.1021/acs.jpcb.8b04834
J. Phys. Chem. B 2018, 122, 11030−11038

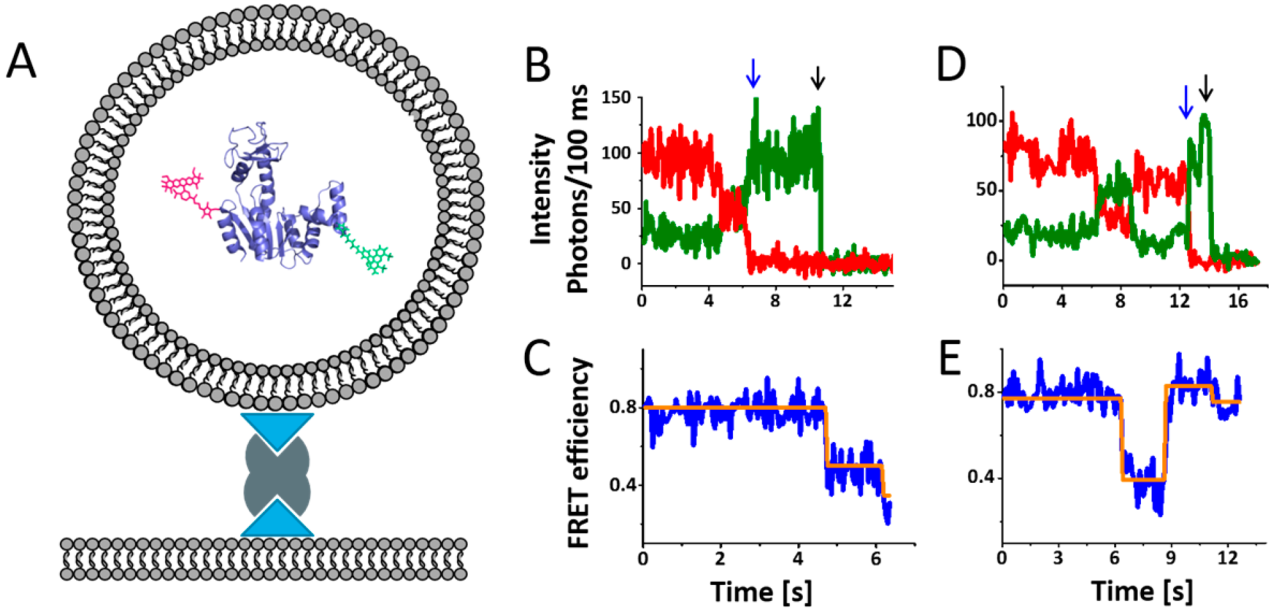


Figure 3. Single-molecule FRET experiment on immobilized AK molecules. (A) Individual AK molecules were encapsulated in vesicles tethered to a glass-supported bilayer via biotin−streptavidin chemistry. Note that the size of an AK molecule (~4 nm) is much smaller than the diameter of the vesicle (100 nm). (B−E) Two examples of single-molecule trajectories of AK molecules, showing one or more transitions between different conformational states. (B, D) Experimental intensity traces from the donor and acceptor channels (black arrows indicate the point at which both dyes become inactive). (C, E) Corresponding FRET efficiency traces, calculated until the photobleaching point of the acceptor (marked by the blue arrows). The orange curves in these graphs are state assignments based on application of the Viterbi algorithm following HMM analysis.

To clarify the microscopic reasons for the changes in stability of the two variants, we performed molecular dynamics (MD) simulations, starting from the open conformations of the proteins (details in section 8 of Suppoting Information Methods). From inspecting a few thousand snapshots from the MD trajectories, we found that in WT AK ~31% of the sampled conformations enabled formation of a hydrogen bond between the ϵ-amino group of Lys 69 and the backbone carbonyl group of Ala 99 at the C-cap of helix 5 (residues 90−99) in the CORE domain (Figure S3A). However, in the mutant K69R the fraction of conformations that could potentially form the hydrogen bond dropped by a factor of 2.5 (Figure S3B). This is most likely due to the bulkiness of the guanidinium group of the arginine side chain, favoring side-chain conformations that are incompatible with hydrogen bonding. On the basis of these observations, we suggest that the decrease in stability of AK_{ds} is primarily due to the disruption of a polar interaction at the interface between the CORE and the NMPbind domains. Both in the crystal structure of AK_s and in the average structure obtained from the MD simulations we observed that Ile64 faced into the core of the NMPbind domain, thereby better filling a cavity that existed in WT AK (Figures S2 and S4). Consistent with this interpretation, the packing density of AK_s (calculated from the MD-based structure using the software tool Voronoia[33]) increased by ~14% with respect to the WT (Table S2).

2. Single-Molecule FRET Experiments on AK Variants. To study the folding dynamics of AK variants, we encapsulated individual double-labeled AK molecules in surface-tethered vesicles (Figure 3A) under partially denaturing conditions and measured their fluorescence as a function of time (details in section 9 of SI Methods). This immobilization method, developed in our lab,[25,26,34−36] enables extending the acquisition time in smFRET experiments while avoiding artifacts related to interactions of protein molecules with the surface.[37] Using an automated home-built single-molecule setup (details in section 10 of the Supporting Information Methods and Figure S6), large data sets of single-molecule trajectories were collected for each AK variant at several denaturant concentrations around the denaturation transition midpoint (C_m). At this range of GdmCl concentrations, protein molecules were expected to exhibit rich folding dynamics. Each trajectory, consisting of the photon arrival times of both donor and acceptor fluorophores, was binned in 50 ms bins and subjected to rigorous computational filtering as described in section 11.2 of the Supporting Information Methods section. Statistical information on the single-molecule data sets can be found in Table S3. Mean FRET efficiency values calculated at each denaturant concentration matched very well the FRET efficiency values measured in ensemble experiments, thereby validating the quality of the data sets (Figure S7A). Moreover, we compared FRET efficiency probability distributions generated from the binned single-molecule traces with FRET efficiency histograms obtained from experiments on freely diffusing molecules. We found a satisfactory agreement in the peak positions and their widths, demonstrating the validity of both the vesicle encapsulation method and the applied computational filters (Figure S7B).

Examples of single-molecule intensity and FRET trajectories are shown in Figure 3B (for more trajectories see Figure S8). Folding and unfolding conformational transitions appeared in the trajectories as anticorrelated steps in donor and acceptor fluorescence intensities. Using a model-free change-point algorithm,[38] we analyzed each individual trajectory to automatically identify transitions between adjacent FRET efficiency levels. To obtain a global picture of the folding and unfolding transitions between the states, we generated two-dimensional transition density (TD) maps, in which the density of transitions for each pair of initial and final FRET efficiency values was plotted. (For details on TD map

DOI: 10.1021/acs.jpcb.8b04834
J. Phys. Chem. B 2018, 122, 11030−11038

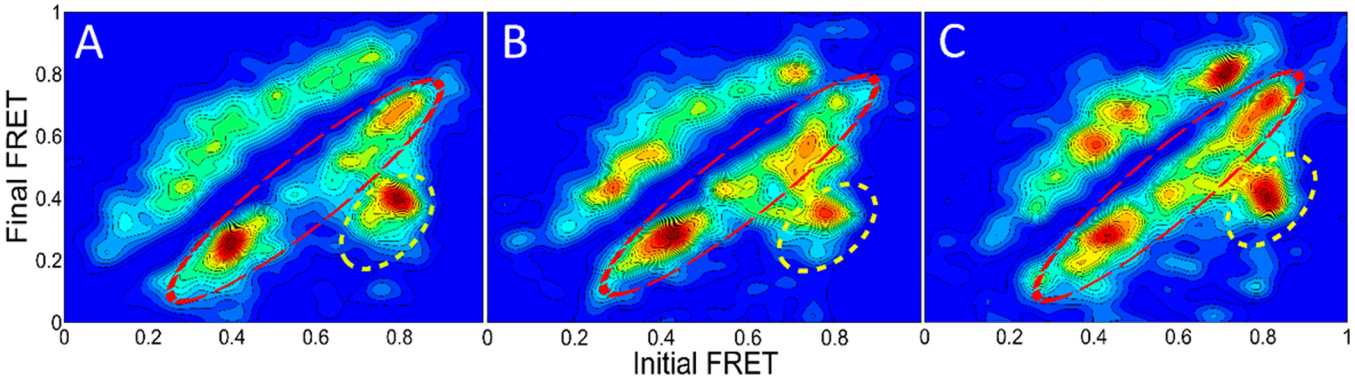


Figure 4. Transition density maps. The maps were calculated from experimental data sets acquired at the corresponding midpoint GdmCl concentrations of AK_s (A), AK_{wt} (B), and AK_{ds} (C). They were constructed from multiple transitions, which were detected in FRET efficiency trajectories using the change-point method. Close-to-the-diagonal peaks (marked by red dashed circles) and far-from-the-diagonal peaks (yellow dashed circles) indicate the existence of various connectivities between intermediate states (i.e., both sequential and nonsequential transitions).

construction, see section 11.3 of Supporting Information Methods.)

Figure 4 shows TD maps obtained at the C_m of each AK variant. The occurrence of multiple peaks on the maps clearly implies deviation from two-state folding behavior, which should involve only two symmetrical peaks.[36] The maps show a very similar general pattern for all AK variants. Some peaks (marked in red dashed circles in Figure 4) are dispersed close to the diagonal, and can be attributed to transitions between states that are close in FRET efficiency. Other peaks (marked in yellow dashed circles) are found far from the diagonal, and are attributed to transitions between states that largly differ in their FRET efficiencies. The asymmetry of the maps with respect to the diagonal is due to due to the larger photobleaching rate of the donor probe compared with the acceptor probe.[26] We also found that for all AK variants the overall shape of the TD maps changes in a similar manner with denaturant concentration (Figure S9). In particular, at high GdmCl concentration folding is dominated by sequential transitions. However, as conditions become more native-like, folding involves more and more nonsequential transitions, implying that there are more possible pathways for folding.[26]

As noted above, a similar response of the folding landscape to denaturant concentration variation was observed in all AK variants. However, a major difference between the three AK variants is the range of denaturant concentrations over which this equivalent folding behavior persists. Indeed, while for AK_{wt} this range was 0.57−0.77 M, for AK_{ds} it shifted to lower denaturant concentrations, 0.5−0.7 M, and for AK_s it moved to 0.65−0.85 M.

3. HMM Analysis of Single-Molecule Trajectories. Although the TD maps provide important information on the connectivity between states, it is not trivial to extract from them the population of states, or to quantify kinetic and thermodynamic properties. To obtain a quantitative description of the conformational states visited by AK during its folding process, we applied HMM analysis.[26,31,39] In this analysis one assumes Markovian dynamics involving N discrete hidden states. The FRET efficiency of each state is normally distributed and described by a mean value and a width. At each time step, AK could adopt any one of the N possible states, with a probability that is conditionally dependent on its state at the previous time step via a transition probability matrix whose entries are proportional to the rates of interconversion between the states. Our analysis was based on the same version of HMM developed in our previous work,[26] which included the detailed-balance assumption and treated photobleaching explicitly (details in section 11.4 in the Supporting Information Methods). To obtain a maximum-likelihood estimate for the HMM parameters, we implemented the Baum−Welch algorithm.[31] For the current analysis we introduced a modification according to which the mean FRET efficiency and the width values assigned to each molecular state were computed globally as shared variables throughout all data sets of a particular AK variant. On the other hand, the transition rates between states and their relative populations were optimized for each individual data set (details in section 11.4 in Supporting Information Methods). This modification was motivated by the assumption that, in the narrow range of denaturant concentrations used in the current experiments, we do not expect the structure of the intermediate states, and the corresponding FRET efficiency values, to vary significantly between different data sets acquired for the same AK variant. Importantly, global analysis both strengthened the statistical robustness by reducing sensitivity to noise and facilitated comparisons between data sets.

HMM analysis does not determine the optimal number of metastable states involved in the observed dynamics, but rather infers optimal parameters under a fixed model complexity (number of parameters). Various criteria have been used in order to obtain the optimal number of states.[40−42] In this work, we developed a new empirical criterion based on the consistency between HMM-optimized parameters and parameters retrieved from dwell-time distributions (details in section 11.5 in Supporting Information Methods). Values of this empirical criterion as a function of the number of states for the wild-type and the two variants are plotted in the graphs in Figure S10. Although the plots do not reveal a sharp single minimum, we do observe a broad minimum spanning four to seven states. Importantly, this range overlaps with the range obtained in our previous work,[26] in which the same positions for the donor and acceptor labels and the same pair of fluorophores were used, though with a different criterion. A model including six states thus appears as a reasonable choice for all variants, and we therefore decided to use six states in all further analysis.

Mean FRET efficiency values for these states and values of the widths of their distributions were obtained from a global HMM analysis of each protein variant and are listed in Table S4. The introduction of global analysis in this work led to a

DOI: 10.1021/acs.jpcb.8b04834
J. Phys. Chem. B 2018, 122, 11030−11038

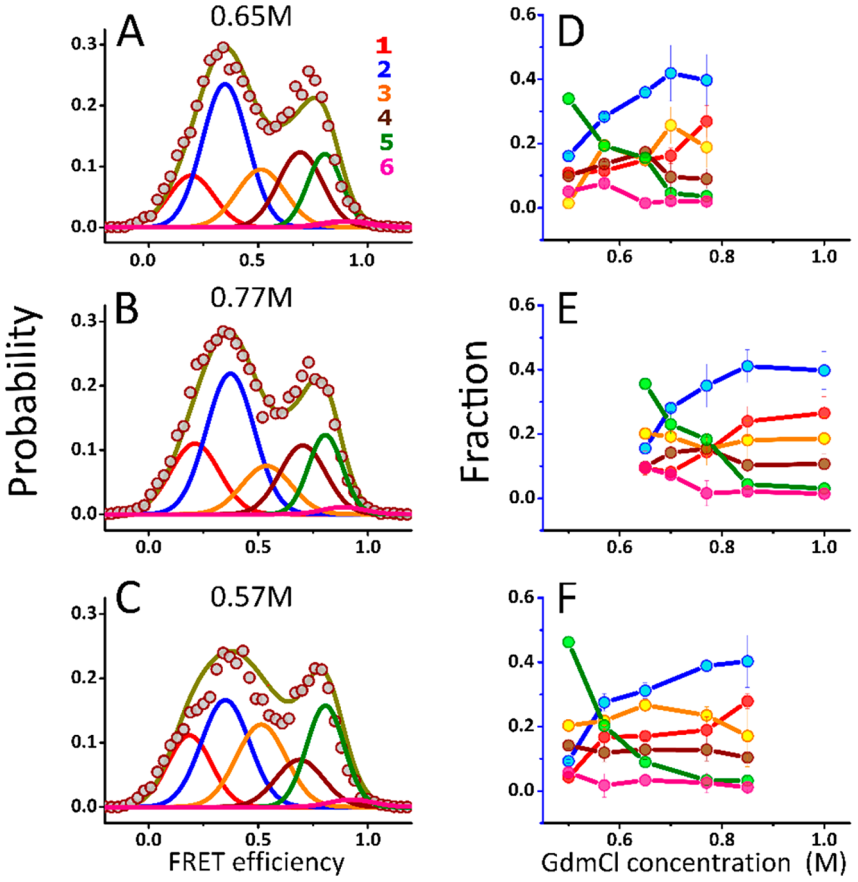


Figure 5. State-probability distributions and evolution of state propensities with denaturant concentration. (A−C) State-probability distributions for the three variants (AK_{wt}, A; AK_s, B; AK_{ds}, C). The distributions were constructed from model parameters obtained from HMM analysis on the data sets acquired at the C_m concentrations, shown at the top of each panel. Each of the six states is represented by a different color. Olive-colored lines show the total probability distributions, which fit well with the histograms retrieved directly from the experimental data sets (plotted as circles). (D−F) The change of the relative state population in response to varying denaturant concentration (AK_{wt}, D; AK_s, E; AK_{ds}, F). The value of each point is the relative population of the corresponding state, i.e., the area under its peak in the state-probability distribution.

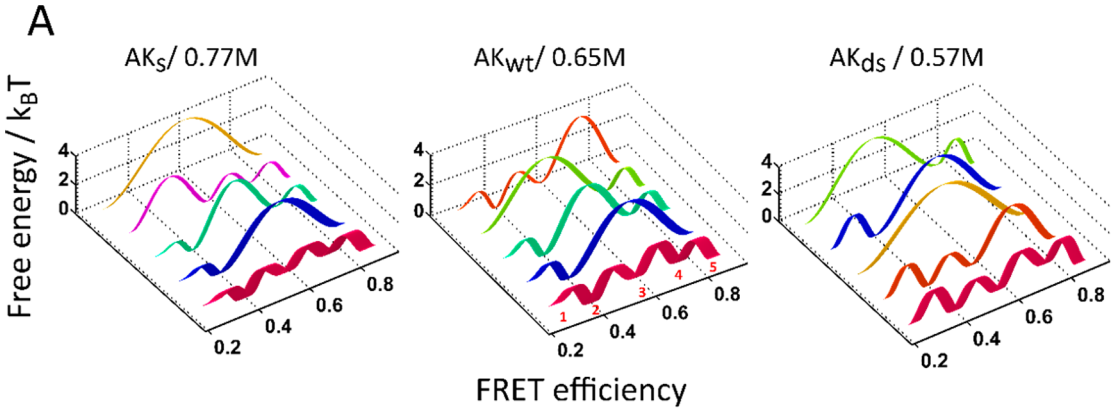
slightly different set of FRET efficiency values and state populations compared to those reported in our previous work.[26] However, FRET efficiency values were actually found to be similar for all three AK variants, validating the new approach and suggesting that the structure of the metastable states populated on the folding landscape did not change significantly upon introduction of mutations. Below, we concentrate on analysis of the state-probability distributions and the folding pathways partition, as obtained from the HMM analysis.

Equilibrium State-Probability Distributions and Their Variation with Denaturant Concentration. Stationary FRET efficiency distributions of the six conformational states at C_m and the evolution of state propensities with increasing denaturant concentrations are shown in Figure 5. The overall FRET efficiency probability distribution matches well the experimental FRET efficiency histograms, plotted in circles in Figure 5A−C, pointing to the excellent fitting between the model and the experimental data. The relative populations of the six states are quite similar for AK_{wt} and AK_s but differ substantially for AK_{ds}. In particular, the relative population of the abundant intermediate state depicted in blue is decreased in AK_{ds}, compared to the other two variants. The relative population of the state in brown is also substantially decreased in AK_{ds}. A bar graph in Figure S11 delineates these variations in state propensity by plotting the relative free energy of states with respect to the most unfolded state (colored in red), which is a reference state in all AK variants.

The plots in Figure 5D−F demonstrate the dependence of state propensities on denaturant concentration. Interestingly, in all variants, the populations of only two states vary significantly, while the populations of the four other states change rather moderately. In particular, the population of the folded state (colored in green) decreases, and the population of the blue-colored intermediate state rises monotonically. Thus, it is mostly the balance between populations of these two states that determines the change in the mean FRET efficiency as the denaturant concentration increases. Interestingly, at the highest denaturant concentration all states reach similar relative populations in all AK variants.

Distributions of Folding Pathways at C_m Reveal Folding Landscape Differences between Variants. In order to shed light on differences in folding dynamics between the variants, we computed the probabilities of various productive folding pathways, namely, pathways that start with the most unfolded state and end with the most folded state. To that end we performed stochastic simulations of AK folding using the state-to-state transition rates obtained from the HMM analysis (details in section 11.6 in Supporting Information Methods). The five most sampled productive pathways at C_m are presented in Figure 6A. State 6, which is also the least populated, does not emerge in these pathways due to its poor

DOI: 10.1021/acs.jpcb.8b04834
J. Phys. Chem. B 2018, 122, 11030−11038



B

Pathways / Variants	AK_{wt}	AK_s	AK_{ds}
1→2→3→4→5	0.26	0.29	0.24
1→2→5	0.24	0.25	0.14
1→2→4→5	0.16	0.14	0.04
1→4→5	0.12	0.03	0.08
1→2→3→5	0.09	0.04	0.16
1→5	0.02	0.08	0.14
1→3→4→5	0.02	0.09	0.06

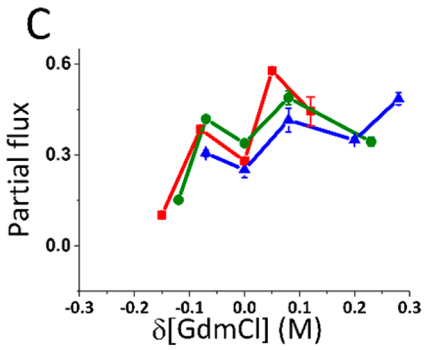


Figure 6. Folding pathways of AK. (A) Free-energy profiles of the five most probable folding pathways for each AK variant at its C_m concentration (indicated on top of each panel). The relative free energy of each state (in kT) was calculated from its relative population with respect to the most populated state. The heights of the free-energy barriers between pairs of states were evaluated from the HMM transition probability matrices using the Arrhenius equation, in which the value of the pre-exponential factor was arbitrarily set to 1 s^{-1}. (The correct value of the pre-exponential factor would only scale all barriers by the same factor.) The width of the lines depicts the partial flux of each pathway. State 6, which is the least populated and is poorly connected to other states, does not contribute significantly and hence does not emerge in these pathways. (B) A list of the sequences of states visited in each pathway, where the same color code is used as in part A. Numbers are the fractions of the total folding flux for each of the listed pathways of each AK variant. (C) A plot of partial flux of the sequential pathway as a function of the difference in denaturant concentration with respect to C_m. AK_{wt} in red, AK_s in green, and AK_{ds} in blue.

kinetic connectivity to other states. This finding and the fact that it exhibits the highest FRET efficiency suggest that state 6 might be a misfolded, highly compact state of the protein. Both the sequential pathway (involving only transitions between neighboring states) and nonsequential pathways (involving also transitions between non-neighboring states) coexist in all AK variants, in agreement with the TD maps. Indeed, in all variants the sequential pathway is dominant. Interestingly, the three main folding pathways in AK_s are the same as in the wild-type protein, showing comparable flux contributions (Figure 6B). By contrast, for AK_{ds} the folding pathways are significantly repartitioned. Particularly, while the sequential pathway is still the most sampled one, the pathways 1→ 2 → 5 and 1 → 2 → 4 → 5, which carry significant flux in AK_{wt} and AK_s, are are much less frequented in this mutant, and instead, the folding process involves other dominant pathways. Further, the probabilities of pathways 1→ 2 → 3 → 5 and 1 → 5 are increased by ∼80% (16% vs 9% in AK_{wt}) and by 7-fold (14% vs <2.0% in AK_{wt}), respectively. Overall, most of the folding flux in AK_{ds} still flows through multistate pathways involving 3−5 states. The partial flux of the fully sequential pathway rises monotonically with denaturant concentration in all AK variants (Figure 6C). At high denaturant concentrations, the sequential pathway is the most favorable one in all variants, carrying ∼50% of the folding flux, whereas at low denaturant concentrations this pathway is much suppressed, and parallel intersecting pathways involving nonsequential transitions prevail. This observation is in agreement with previous results on the relation between the denaturant concentration and state connectivity.[26]

DISCUSSION

Multidomain proteins constitute over 75% of the eukaryotic proteome. Therefore, understanding how their complex topologies influence their folding behavior is of great interest. Although some aspects of folding of multidomain proteins were deciphered,[43,44] many of these studies focused on two-domain proteins that are singly linked and our understanding of more complex topologies is still limited. The inherent heterogeneity and complexity of such proteins calls for studies at the single-molecule level.[9,45,46] In this work, we used smFRET to further explore the folding of AK, whose structure involves two domains inserted within a larger one. We hypothesized that since the NMPbind domain of AK is inserted within the CORE domain, stabilizing the former will significantly affect the latter, and by necessity the folding of the whole protein. We based our narrative on observations from the simulations of Li et al.,[27] according to which the folding process of AK proceeds mainly through pathways in which the NMPbind domain folds very late, probably due to its low stability. Accordingly, we altered the stability of the NMPbind domain by introducing selected mutations using atomistic

DOI: 10.1021/acs.jpcb.8b04834
J. Phys. Chem. B 2018, 122, 11030−11038

modeling in Rosetta. We then investigated the variations in the folding course of the two variants that showed the most significant changes in thermal stability.

Interestingly, both mutants showed a general dependence of the partition between sequential and nonsequential pathways on the denaturant concentration, as observed for WT AK in our previous study.[26] In fact, as the denaturant concentration increased, a larger fraction of the folding flux went through the fully sequential pathway, while at lower concentrations folding was more likely to proceed via multiple parallel pathways involving nonsequential transitions. The simulations of Li et al. supported this observation: As the temperature increased, a larger fraction of the folding flux went through one specific sequence of states.[27] A similar dependence on denaturant concentration was demonstrated for a PDZ domain in a recent paper by Liu et al.[47] How can we explain this behavior?

In general, the progress in folding of a protein is dictated by the balance between the tendency to form stabilizing interactions resulting in compaction of the protein chain and the tendency to minimize loss in conformational entropy as the protein becomes more and more compact.[48] In the case of AK, it is likely that intermediate states that differ significantly in their structure (i.e., nonsequential states) may interconvert at low denaturant concentrations, as the formation of enthalpically favored intrachain interactions may compensate for the steep decrease in conformational entropy during such a transition. However, as the denaturant concentration increases, the protein chain is stabilized in more expanded conformations and the contribution of intrachain interactions decreases. Therefore, the folding reaction might tend to proceed through sequential intermediate states, with only mild variations in structure and moderate changes in conformational entropy between each pair of two adjacent intermediate states.

On the basis of the HMM parameters we were able to annotate the different folding pathways. We found that, at the transition midpoint concentration, the stabilized AK_s mutant preserved a nearly wild-type-like partition of the main folding pathways, and only minor changes were observed in the relative stabilities of the intermediate states (Figure S11). From the simulations of Li et al.[27] we expected that stabilization of the NMPbind domain would actually depress folding through the sequential pathway and facilitate folding through more direct pathways. Our findings did not verify this prediction. Instead it seemed that the stabilization of the NMPbind domain globally stabilized CORE domain intermediate states to the same extent, allowing the main folding course to proceed via a similar scenario to the WT, though shifted to a higher denaturant concentration. In particular, the three most probable folding pathways, accounting for more than 65% of the folding flux, appeared with similar fluxes in both AK_{wt} and AK_s. Nevertheless, we could still point out some changes rendered in other folding pathways of AK_s. For example, the flux of the direct transition from the most unfolded to the native state ($1 \rightarrow 5$) became quite significant ($9.1 \pm 1.5\%$) in this mutant, as opposed to the wild-type protein where it contributes only 2% of the total flux.

The analysis of the folding of AK_{ds} revealed quite a different picture. The mutation K69R, which destabilized the protein, also repartitioned folding pathways significantly. A significant reshaping of the folding landscape is a direct result of the changes in the relative stabilities of the intermediate states, as observed in Figure S11. Thus, CORE-NMPbind interdomain interactions, which are disrupted in AK_{ds}, seem to have a more profound effect on AK's folding landscape than NMPbind intradomain interactions. Furthermore, AK_{ds} exhibited a reduction in folding cooperativity, as discerned both from denaturant *m*-values (Figure 6, Table S1) and from a simulation of changes in the solvent accessible surface area during folding based on the HMM results (Figure S12, section 11.7 in Supporting Information methods). These findings are in agreement with other studies pointing at the importance of interdomain interactions in shaping the folding dynamics of multidomain proteins.[16,49,50]

In conclusion, our results reveal the folding plasticity of AK, and the adaptive potential of this protein to accomplish folding in response to variations in sequence or stability. The tunability of the folding routes by minimal sequence changes (e.g., a single, surface, conservative mutation) makes this system an excellent prototype for studying subtle features of the folding free-energy landscape. Future studies of additional AK variants would certainly shed more light on the relation of structure and dynamics in the folding of this multidomain protein.

■ ASSOCIATED CONTENT

Supporting Information

See Supporting Information for additional figures and for detailed methods. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcb.8b04834.

Additional figures, including structures, plots of rates, probability distributions, and domain cavities, as well as detailed methods (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: gilad.haran@weizmann.ac.il.

ORCID

Sarel J. Fleishman: 0000-0003-3177-7560

Gilad Haran: 0000-0003-1837-9779

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by Grant 686/14 from the Israel Science Foundation. We thank Menachem Pirchi for help with statistical analysis and Harry Greenblatt from Weizmann Institute of Science for help with molecular dynamics simulations.

■ REFERENCES

(1) Anfinsen, C. B. Principles That Govern the Folding of Protein Chains. *Science (Washington, DC, U. S.)* **1973**, *181*, 223−230.

(2) Dill, K. a; Chan, H. S. From Levinthal to Pathways to Funnels. *Nat. Struct. Mol. Biol.* **1997**, *4*, 10−19.

(3) King, J.; Haase-Pettingell, C.; Gossard, D. Protein Folding and Misfolding. *Am. Sci.* **2002**, *90*, 445−453.

(4) Soto, C. Unfolding the Role of Protein Misfolding in Neurodegenerative Diseases. *Nat. Rev. Neurosci.* **2003**, *4*, 49−60.

(5) 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.

(6) Leopold, P. E.; Montal, M.; Onuchic, J. N. Protein Folding Funnels: A Kinetic Approach to the Sequence-Structure Relationship. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 8721−8725.

(7) Dill, K. A.; Ozkan, S. B.; Shell, M. S.; Weikl, T. R. The Protein Folding Problem. *Annu. Rev. Biophys.* **2008**, *37*, 289−316.

DOI: 10.1021/acs.jpcb.8b04834
J. Phys. Chem. B 2018, 122, 11030−11038

(8) Orte, A.; Craggs, T. D.; White, S. S.; Jackson, S. E.; Klenerman, D. Evidence of an Intermediate and Parallel Pathways in Protein Unfolding from Single-Molecule Fluorescence. *J. Am. Chem. Soc.* **2008**, *130*, 7898−7907.
(9) Stigler, J.; Ziegler, F.; Gieseke, A.; Gebhardt, J. C. M.; Rief, M. The Complex Folding Network of Single Calmodulin Molecules. *Science (Washington, DC, U. S.)* **2011**, *334*, 512−516.
(10) Chung, H. S.; McHale, K.; Louis, J. M.; Eaton, W. a. Single-Molecule Fluorescence Experiments Determine Protein Folding Transition Path Times. *Science (Washington, DC, U. S.)* **2012**, *335*, 981−984.
(11) König, I.; Zarrine-Afsar, A.; Aznauryan, M.; Soranno, A.; Wunderlich, B.; Dingfelder, F.; Stüber, J. C.; Plückthun, A.; Nettels, D.; Schuler, B. Single-Molecule Spectroscopy of Protein Conformational Dynamics in Live Eukaryotic Cells. *Nat. Methods* **2015**, *12*, 773−779.
(12) Guinn, E. J.; Jagannathan, B.; Marqusee, S. Single-Molecule Chemo-Mechanical Unfolding Reveals Multiple Transition State Barriers in a Small Single-Domain Protein. *Nat. Commun.* **2015**, *6*, 6861.
(13) Benke, S.; Nettels, D.; Hofmann, H.; Schuler, B. Quantifying Kinetics from Time Series of Single-Molecule Förster Resonance Energy Transfer Efficiency Histograms. *Nanotechnology* **2017**, *28*, 114002.
(14) Eaton, W. A.; Munoz, V.; Hagen, S. J.; Jas, G. S.; Lapidus, L. J.; Henry, E. R.; Hofrichter, J. Fast Kinetics and Mechanisms in Protein Folding. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 327−359.
(15) Gelman, H.; Gruebele, M. Fast Protein Folding Kinetics. *Q. Rev. Biophys.* **2014**, *47*, 95−142.
(16) Han, J.-H.; Batey, S.; Nickson, A.; Teichmann, S.; Clarke, J. The Folding and Evolution of Multidomain Proteins. *Nat. Rev. Mol. Cell Biol.* **2007**, *8*, 319−330.
(17) Bhaskara, R. M.; Srinivasan, N. Stability of Domain Structures in Multi-Domain Proteins. *Sci. Rep.* **2011**, *1*. DOI: 10.1038/srep00040
(18) Levy, Y. Protein Assembly and Building Blocks: Beyond the Limits of the LEGO Brick Metaphor. *Biochemistry* **2017**, *56*, 5040−5048.
(19) Lerner, E.; Orevi, T.; Ben Ishay, E.; Amir, D.; Haas, E. Kinetics of Fast Changing Intramolecular Distance Distributions Obtained by Combined Analysis of FRET Efficiency Kinetics and Time-Resolved FRET Equilibrium Measurements. *Biophys. J.* **2014**, *106*, 667−676.
(20) Ben Ishay, E.; Rahamim, G.; Orevi, T.; Hazan, G.; Amir, D.; Haas, E. Fast Subdomain Folding Prior to the Global Refolding Transition of E. Coli Adenylate Kinase: A Double Kinetics Study. *J. Mol. Biol.* **2012**, *423*, 613−623.
(21) Zhang, H. J.; Sheng, X. R.; Pan, X. M.; Zhou, J. M. Refolding of Urea-Denatured Adenylate Kinase. *Biochem. J.* **1998**, *333* (Pt 2), 401−405.
(22) Ratner, V.; Kahana, E.; Haas, E. The Natively Helical Chain Segment 169−188 of Escherichia Coli Adenylate Kinase Is Formed in the Latest Phase of the Refolding Transition. *J. Mol. Biol.* **2002**, *320*, 1135−1145.
(23) Ratner, V.; Amir, D.; Kahana, E.; Haas, E. Fast Collapse but Slow Formation of Secondary Structure Elements in the Refolding Transition of E. Coli Adenylate Kinase. *J. Mol. Biol.* **2005**, *352*, 683−699.
(24) Orevi, T.; Ben Ishay, E.; Pirchi, M.; Jacob, M. H.; Amir, D.; Haas, E. Early Closure of a Long Loop in the Refolding of Adenylate Kinase: A Possible Key Role of Non-Local Interactions in the Initial Folding Steps. *J. Mol. Biol.* **2009**, *385*, 1230−1242.
(25) Rhoades, E.; Gussakovsky, E.; Haran, G. Watching Proteins Fold One Molecule at a Time. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 3197−3202.
(26) Pirchi, M.; Ziv, G.; Riven, I.; Cohen, S. S.; Zohar, N.; Barak, Y.; Haran, G. Single-Molecule Fluorescence Spectroscopy Maps the Folding Landscape of a Large Protein. *Nat. Commun.* **2011**, *2*, 493.
(27) Li, W. F.; Terakawa, T.; Wang, W.; Takada, S. Energy Landscape and Multiroute Folding of Topologically Complex Proteins Adenylate Kinase and 2ouf-Knot. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 17789−17794.
(28) Giri Rao, V. V. H.; Gosavi, S. In the Multi-Domain Protein Adenylate Kinase, Domain Insertion Facilitates Cooperative Folding While Accommodating Function at Domain Interfaces. *PLoS Comput. Biol.* **2014**, *10*, e1003938.
(29) Taylor, J. N.; Pirchi, M.; Haran, G.; Komatsuzaki, T. Deciphering Hierarchical Features in the Energy Landscape of Adenylate Kinase Folding/Unfolding. *J. Chem. Phys.* **2018**, *148*, 123325.
(30) Das, R.; Baker, D. Macromolecular Modeling with Rosetta. *Annu. Rev. Biochem.* **2008**, *77*, 363−382.
(31) Rabiner, L. R. R. A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. *Proc. IEEE* **1989**, *77*, 257−286.
(32) Goldenzweig, A.; Fleishman, S. Principles of Protein Stability and Their Application in Computational Design. *Annu. Rev. Biochem.* **2018**, *87*, 105−129.
(33) Rother, K.; Hildebrand, P. W.; Goede, A.; Gruening, B.; Preissner, R. Voronoia: Analyzing Packing in Protein Structures. *Nucleic Acids Res.* **2009**, *37*, D393.
(34) Boukobza, E.; Sonnenfeld, A.; Haran, G. Immobilization in Surface-Tethered Lipid Vesicles as a New Tool for Single Biomolecule Spectroscopy. *J. Phys. Chem. B* **2001**, *105*, 12165−12170.
(35) Piwonski, H. M.; Goomanovsky, M.; Bensimon, D.; Horovitz, A.; Haran, G. Allosteric Inhibition of Individual Enzyme Molecules Trapped in Lipid Vesicles. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, E1437−E1443.
(36) Aviram, H. Y.; Pirchi, M.; Barak, Y.; Riven, I.; Haran, G. Two States or Not Two States: Single-Molecule Folding Studies of Protein L Two States or Not Two States: Single-Molecule Folding Studies of Protein L. *J. Chem. Phys.* **2018**, *148*, 123303.
(37) Friedel, M.; Baumketner, A.; Shea, J.-E. Effects of Surface Tethering on Protein Folding Mechanisms. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 8396−8401.
(38) Taylor, W. A. Change-Point Analysis: A Powerful New Tool For Detecting Changes. 2000. http://www.variation.com/cpa/tech/changepoint.html.
(39) Andrec, M.; Levy, R. M.; Talaga, D. S. Direct Determination of Kinetic Rates from Single-Molecule Photon Arrival Trajectories Using Hidden Markov Models. *J. Phys. Chem. A* **2003**, *107*, 7454−7464.
(40) Hu, S. Akaike Information Criterion Statistics. *Math. Comput. Simul.* **1987**, *29*, 452.
(41) McKinney, S. A.; Joo, C.; Ha, T. Analysis of Single-Molecule FRET Trajectories Using Hidden Markov Modeling. *Biophys. J.* **2006**, *91*, 1941−1951.
(42) Rissanen, J. A Universal Prior for Integers and Estimation by Minimum Description Length. *Ann. Stat.* **1983**, *11*, 416−431.
(43) Batey, S.; Nickson, A. A.; Clarke, J. Studying the Folding of Multidomain Proteins. *HFSP J.* **2008**, *2*, 365−377.
(44) Arviv, O.; Levy, Y. Folding of Multidomain Proteins: Biophysical Consequences of Tethering Even in Apparently Independent Folding. *Proteins: Struct., Funct., Genet.* **2012**, *80*, 2780−2798.
(45) Kotamarthi, H. C.; Sharma, R.; Narayan, S.; Ray, S.; Ainavarapu, S. R. K. Multiple Unfolding Pathways of Leucine Binding Protein (LBP) Probed by Single-Molecule Force Spectroscopy (SMFS). *J. Am. Chem. Soc.* **2013**, *135*, 14768−14774.
(46) Borgia, A.; Kemplen, K. R.; Borgia, M. B.; Soranno, A.; Shammas, S.; Wunderlich, B.; Nettels, D.; Best, R. B.; Clarke, J.; Schuler, B. Transient Misfolding Dominates Multidomain Protein Folding. *Nat. Commun.* **2015**, *6*, 8861.
(47) Liu, Z.; Thirumalai, D. Denaturants Alter the Flux through Multiple Pathways in the Folding of PDZ Domain. *J. Phys. Chem. B* **2018**, *122*, 1408−1416.
(48) Bicout, D. J.; Szabo, A. Entropic Barriers, Transition States, Funnels, and Exponential Protein Folding Kinetics: A Simple Model. *Protein Sci.* **2000**, *9*, 452−465.

DOI: 10.1021/acs.jpcb.8b04834
J. Phys. Chem. B 2018, 122, 11030−11038

(49) Freire, E.; Murphy, K. P.; Sanchez-Ruiz, J. M.; Galisteo, M. L.; Privalov, P. L. The Molecular Basis of Cooperativity in Protein Folding. Thermodynamic Dissection of Interdomain Interactions in Phosphoglycerate Kinase+. *Biochemistry* **1992**, *31*, 250−256.
(50) Bhaskara, R. M.; Srinivasan, N. Stability of Domain Structures in Multi-Domain Proteins. *Sci. Rep.* **2011**, *1*, 40.

DOI: 10.1021/acs.jpcb.8b04834
J. Phys. Chem. B 2018, 122, 11030−11038