Published in final edited form as:
J Am Chem Soc. 2009 June 3; 131(21): 7470–7476. doi:10.1021/ja901860f.

Probing the Folding Transition State Structure of the Villin Headpiece Subdomain via Sidechain and Backbone Mutagenesis

Michelle R. Bunagan[†], **Jianmin Gao**[‡], **Jeffery W. Kelly**[‡,*], and **Feng Gai**[†,*]

[†]Department of Chemistry, University of Pennsylvania, 231 S. 34th Street, Philadelphia, PA 19104

[‡]Department of Chemistry and The Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, BCC265, La Jolla, CA 92037

Abstract

Backbone-backbone hydrogen bonds are a common feature of native protein structures, yet their thermodynamic and kinetic influence on folding has long been debated. This is reflected by the disparity between current protein folding models, which place hydrogen bond formation at different stages along the folding trajectory. For example, previous studies have suggested that the denatured-state of the villin headpiece subdomain contains residual helical structure that may provide a bias toward the folded state by confining the conformational search associated with its folding. Although helical hydrogen bonds clearly stabilize the folded state, here we show, using an amide-to-ester mutation strategy, that the formation of backbone hydrogen bonds within helices is not rate-limiting in the folding of the subdomain, thereby suggesting that such hydrogen bonds are unlikely to be formed en route from the denatured to the transition state. On the other hand, elimination of hydrogen bonds within the turn region elicits a slower folding rate, consistent with the hypothesis that these residues are involved in the formation of a folding nucleus. While illustrating a potentially conserved aspect of helix-turn-helix folding, our results further underscore the inherent importance of turns in protein supersecondary structure formation.

Keywords

Protein folding; amide-to-ester mutation; villin headpiece subdomain; *T*-jump

Introduction

Backbone-backbone hydrogen bonds (H-bond) are common and apparently important features of native protein structures,[1–4] yet their role in folding transition state structure remains unclear. One possibility is that H-bonds result as a consequence of the formation of native sidechain contacts during folding.[5] Another possibility is that they play a more deciding role by providing a framework that directs the flow of folding toward the native state.[6] To provide a better understanding of the kinetic influence of backbone-backbone H-bonds in protein folding, we investigate the role of helical H-bonding in the folding kinetics of the villin headpiece subdomain (HP-35) using the Φ-value analysis method in conjunction with amide-to-ester (A-to-E) backbone mutagenesis.[7–15]

*To whom correspondence should be addressed. Email: gai@sas.upenn.edu, jkelly@scripps.edu.

Supporting Information Available. Thermodynamic unfolding parameters, ^{1}H-NMR spectra, and CD melting curves for all of the mutants, as well as FT-IR data for wild-type HP-35 and P21A and representative *T*-jump relaxation kinetics for wild-type HP-35 and A18α. This material is available free of charge via the Internet at http://pubs.acs.org

Φ-value analysis[16] is commonly used to examine the critical features of the folding transition state. The Φ-value arising from a specific mutation, which is calculated according to $\Phi_F = \Delta\Delta G^{\ddagger}/ \Delta\Delta G_f$, where $\Delta\Delta G^{\ddagger}$ is the change of the folding free energy barrier and $\Delta\Delta G_f$ is the change of the folding free energy, provides information regarding the extent to which the native contacts contribute to the stability of the folding transition state (Figure 1). However, using traditional sidechain mutagenesis-based Φ-value analysis to definitively assess the kinetic role of backbone H-bonding in protein folding is difficult, if not impossible. Thus, the recently established A-to-E backbone mutagenesis approach[7–15] was applied to eliminate specific backbone-backbone hydrogen bonds. While an ester lacks a H-bond donor, it is an ideal steric and electronic mimic of the amide group, exhibiting nearly identical bond lengths and angles, and strongly favoring a planar, trans-conformation, although the cis-trans barrier is significantly lower.[12]

An A-to-E mutation within a residue that donated an NH to create a backbone-backbone H-bond is expected to increase the free energy of the folded state by 3 to 30 kJ/mol, the extent of destabilization increasing with the degree of desolvation of the perturbed H-bond in the folded state.[12] While an A-to-E mutation introduces a lone pair-lone pair repulsion between the non-carbonyl oxygen of the ester and what was the H-bond acceptor carbonyl oxygen of the amide, we have shown that this destabilization only amounts to ≈1 kJ/mol.[15] The energetic preference for the transfer of an ester bond over an amide bond from the solvated unfolded ensemble to the folded state is determined by the extent to which the bond is desolvated in the folding process and the extent to which the proximal sidechains contribute, although this difference is generally small, ≈2 kJ/mol.[15] Therefore, the $\Delta\Delta G_f$ resulting from an A-to-E mutation largely arises from removal of a native backbone-backbone H-bond that is absent or largely absent in the denatured state and the majority of this perturbation free energy does not arise from desolvation and O-O repulsion contributions, which are small in comparison. Thus, deletion of a H-bond that is formed in both the transition and folded states would be expected to have a large influence on the folding free energy barrier ($\Delta G^{\ddagger}$) (Figure 1A). On the other hand, deletion of a H-bond that is only formed in the folded state would be expected to have little or no influence on $\Delta G^{\ddagger}$ (Figure 1B). Similarly, deletion of a H-bond that is also formed in the unfolded state and the transition state would be expected to have little or no influence on $\Delta G^{\ddagger}$ (Figure 1C). Therefore, A-to-E backbone mutagenesis in conjunction with Φ-value analysis should allow estimation of the extent of the formation of specific H-bonds within the folding transition state, thus providing a more comprehensive understanding of the folding mechanism of the HP-35, as has been demonstrated for the Pin1 WW domain,[10–11] the B domain of protein L and the P22 Arc repressor.[13]

There has been extensive study of the folding mechanism of HP-35 (Figure 2), a small helical protein[17] whose folding is exceedingly fast.[18–46] Several studies have suggested that the unfolded state retains residual native-like interactions and structures, including tertiary contacts between residues forming the hydrophobic core and native helical segments.[27,33,38,39] Although such features have been invoked to explain the ultrafast folding of HP-35 and other proteins,[47–49] a recent kinetic study demonstrated that hydrophobic core formation is not rate-limiting for folding of HP-35.[28] In order to provide further insight into H-bond mediated helix and turn formation, herein we examine the folding thermodynamics and kinetics of the wild-type HP-35 and five single site A-to-E mutants (Figure 2). Mutations were chosen to perturb specific native-state H-bonds. Of the A-to-E mutations investigated, three eliminate helical hydrogen bonds, V9ϖ in helix I, A18α in helix II, and L28λ in helix III, while G11γ and L20λ eliminate H-bonds formed in the loop between helices I and II and the turn between helices II and III, respectively. Additionally, sidechain mutants T13A and Q25nL (norleucine) remove sidechain to backbone H-bonds, while mutant G11A reduces the flexibility of the loop and P21A reduces the rigidity of the turn in order to test the kinetic role of the conformational entropy in these two regions. Our results show no evidence of H-bond-mediated helix formation

in the transition state, as determined via Φ-value analysis. Instead, our results suggest the loop and turn may act as nuclei for folding through a nucleation-condensation mechanism.

Materials and Methods

Peptide synthesis and purification

The wild-type and sidechain variants of HP-35 were synthesized on an ABI 433A peptide synthesizer employing Fmoc/tBu chemistry. The ester mutants were prepared manually using Boc/benzyl strategy, during which the α-hydroxy acids were incorporated individually in place of the corresponding amino acids according to the previously published protocol.[10–12] The α-hydroxy acids used in this study were purchased from commercial sources and used without further purification. All peptides were purified by reverse-phase HPLC, and the A-to-E HP-35 variants were further purified by size-exclusion chromatography employing a Superdex 30 column. Multiple rounds of lyophilization were performed against 0.1 M DCl to remove the trifluoroacetic acid from peptide purification and to allow for proton-deuterium exchange. Exchanged A-to-E samples were again characterized by mass spectroscopy to ensure that the acid had not hydrolyzed the ester bond. For circular dichroism (CD) and IR experiments, the samples were prepared by directly dissolving the lyophilized peptides in either 20 mM (CD) or 40 mM (IR) 2-(*N*-morpholino)ethanesulfonic acid (MES) D_2O buffers (pH* 5.5). The concentration of the sample was determined by absorbance at 280 nm, using an extinction coefficient of 5500 cm^{-1} M^{-1}. The final concentration was ~40 μM and ~1.5 mM for CD and IR, respectively.

CD spectroscopy

Both CD wavelength scans and thermal melting curves were collected on an Aviv model 202SF circular dichroism spectrometer equipped with a cell holder with a Peltier temperature controller (Hellma, Forest Hills, NY). Far-UV CD spectra were recorded from 200 to 250 nm at 2 °C. The peptide sample was dissolved in 20 mM MES (pH 5.5). Thermal denaturation was monitored at 222 nm over the temperature range of 2 to 98 °C with a 2 °C step size and a 90 s equilibration time. The signal was averaged for 30 s at each temperature. After the highest temperature was reached, the sample was cooled to 2 °C and another full CD spectrum was measured to ensure that folding was reversible. The folding-unfolding thermodynamics of each peptide were obtained by fitting its CD thermal melting transition obtained at 222 nm to the following two-state model:

$$\theta(T)=\frac{\theta_F(T)+K_{eq}(T)\times\theta_U(T)}{1+K_{eq}(T)} \quad (1)$$

$$K_{eq}(T)=\exp(-\Delta G(T)/RT) \quad (2)$$

$$\Delta G(T)=\Delta H_m+\Delta C_p\cdot(T-T_m)-T\cdot[\Delta S_m+\Delta C_p\cdot\ln(T/T_m)] \quad (3)$$

where $K_{eq}(T)$ is the equilibrium constant for unfolding, $T_m = \Delta H_m/\Delta S_m$ is the thermal melting temperature, ΔH_m and ΔS_m are the enthalpy and entropy changes at T_m, respectively, and ΔC_p is the heat capacity change associated with unfolding, which has been assumed here to be temperature independent. Since all backbone and sidechain mutations studied here are not expected to significantly change the hydrophobicity of the protein, ΔC_p was treated as a global

fitting parameter. In addition, the folded and unfolded CD baselines $\theta_F(T)$ and $\theta_U(T)$ are assumed to be linear functions of temperature, as follows:

$$\theta_F(T)=m+n\times T \tag{4}$$

$$\theta_U(T)=p+q\times T \tag{5}$$

During fitting, n and q were treated as global parameters.

FT-IR Measurements

FT-IR spectra were collected on a Magna-IR 860 spectrometer (Nicolet, WI) equipped with a HgCdTe detector using a spectral resolution of 2 cm^{-1} . A home made, twocompartment, 52 μm CaF_2 sample cell mounted on a programmable translation stage was used to allow separate and alternate measurements of the single-beam spectra of the sample and reference under identical conditions. Temperature control with ±0.2 °C precision was achieved through a thermostated copper block and a temperature bath. The reported spectra correspond to the average of 256 scans.

Infrared *T*-jump Kinetic Measurements

The time-resolved *T*-jump IR apparatus used in the current study has been described in detail elsewhere.[50] Briefly, a 3 ns laser pulse centered at about 1.9 μm was used to generate a *T*-jump in the laser interaction volume, and the *T*-jump induced transient absorbance change of the sample was measured by a continuous wave (CW) IR diode laser in conjunction with a 50 MHz MCT detector. In all cases, the observed relaxation kinetics were well described by a double exponential. The fast component, which was too fast to be resolved, was attributed to imperfect background subtraction and temperature-dependent spectral shifting.[51] The slower component was attributed to the conformational redistribution kinetics of the peptide. By assuming that the *T*-jump induced relaxation of HP-35 and the mutants occurs between two states, the corresponding folding (k_f) and unfolding (k_u) rate constants were further evaluated using the relaxation rate constant (k_R) and the equilibrium constant (K_{eq}):

$$k_R=k_f+k_u \tag{6}$$

$$K_{eq}=k_f/k_u \tag{7}$$

Φ-value Analysis

The Φ-values were calculated using the following equation:

$$\Phi_F=\frac{RT\ln\left(\frac{\tau_{f,mut}}{\tau_{f,wt}}\right)}{\Delta\Delta G_{mut-wt}} \tag{8}$$

where τ_f is the folding time constant, the subscripts ‘mut’ and ‘wt’ represent the mutant and wild-type, respectively, R is the gas constant, T is the temperature, and $\Delta\Delta G_{mut\text{-}wt} = \Delta G_{mut} - \Delta G_{wt}$.

Results and Discussion

Wild-type HP-35 and nine single site variants were synthesized using standard solid-phase peptide synthesis protocols. In the case of the A-to-E mutants, specific α-amino acid residues were replaced by their α-hydroxy acid counterparts, which are commercially available.[10–12] Following previously used nomenclature,[10–12] all α-hydroxy acid residues are denoted by lowercase Greek letters that correspond to the single letter code of the corresponding α-amino acid residues. A 1D ^{1}H NMR spectrum was collected at 10 °C for each HP-35 sequence studied. In every case, the ^{1}H amide NH and aromatic regions of the NMR spectra exhibit well dispersed resonances similar to those of wild-type HP-35 (Figure S1 in Supporting Information), providing evidence that A-to-E mutations do not alter the integrity of the HP-35 fold.

Equilibrium study

The far-UV CD spectra of wild type HP-35 and its mutants all exhibit double minima at 208 and 222 nm, characteristic of α-helical proteins.[17] Temperature dependent CD measurements further indicate that all of the HP-35 sequences undergo a cooperative thermal unfolding transition, as determined by monitoring the increase in ellipticity at 222 nm (Figure S2 in Supporting Information). The thermodynamic folding parameters were extracted by globally fitting all CD thermal denaturation curves to a two-state model (Table S1 in Supporting Information). All of the A-to-E mutations are destabilizing relative to wild-type HP-35, which is consistent with these H-bonds being formed in the folded state, but not the denatured state. For example, L28λ eliminates a H-bond and weakens an additional H-bond owing to the ester carbonyl being a weaker acceptor. The L20λ HP-35 variant, which eliminates a H-bond between the amide NH of L20 and the backbone carbonyl of F17 in wild-type HP-35, is the least stable A-to-E variant (Table S1 in Supporting Information), underscoring the thermodynamic importance of the H-bond-mediated loop sandwiched between these two residues. A collective look at the perturbation free energies reveals that all H-bonds do not contribute equally to the native state stability, as observed previously.[9–11,14]

The thermal unfolding transition of HP-35 and its single site variants was also studied using FT-IR spectroscopy. The amide Ι band of folded HP-35, which arises mainly from the backbone C=O stretching vibrations, is centered at ~1645 cm^{-1} at 5.6 °C (Figure S3 in Supporting Information). Increasing temperature shifts the peak position towards higher wavenumbers. Consequently, the FT-IR difference spectra of HP-35, calculated by subtracting the 5.6 °C spectrum from those obtained at higher temperatures, exhibit a negative feature centered around 1632 cm^{-1} and a broad positive feature centered at ~1675 cm^{-1}. The former results primarily from loss of α-helical conformations with increasing temperature, whereas the latter is due to the concurrent formation of a non-helical conformational ensemble.[52] Since temperature-induced variations in backbone solvation also shift the amide Ι band towards higher frequency and thus contribute to the aforementioned spectral changes,[51] we did not attempt to quantify the thermal denaturation of HP-35 or the mutants using the amide Ι band. However, these spectral features are useful conformational markers and have been used in the subsequent kinetic studies to monitor the conformational relaxation kinetics of HP-35 and its mutants in response to a *T*-jump. In addition, the temperature dependent FT-IR spectra of all the peptides except P21A, indicate that under the current experimental conditions (for both equilibrium and kinetic studies), no detectable aggregates are formed. Anti-parallel β-sheet aggregates are known to give rise to a pair of characteristic sharp bands, centered at ~1620 and ~1680 cm^{-1}.[53]

T-jump kinetics

The folding-unfolding (or relaxation) kinetics of these peptides were studied by the *T*-jump IR technique, which has been described in detail elsewhere.[50] As shown (Figure S4 in Supporting

Information), the relaxation (or global folding-unfolding) kinetics of HP-35 in response to a *T*-jump of 54.5 – 59.9 °C, probed at 1632 cm^{-1}, indicate that the single-exponential conformational re-equilibration of HP-35 occurs on the microsecond time scale, in agreement with observations of previous NMR lineshape experiments,[25] *T*-jump IR,[28,40] *T*-jump fluorescence,[23,31,37] and fluorescence quenching[29] studies of this system. While a three-state model has been invoked to describe the folding dynamics of HP-35,[40,45] Eaton and coworkers have shown that the intermediate state is located on the folded side of the major free-energy barrier and its relaxation takes places on a much faster time scale than the global folding-unfolding transition, and thus a simple two-state treatment is adequate for Φ-value analysis.[45] As shown (Figure 3), the folding rate constant of HP-35, obtained using equation 6 and equation 7 (methods section), is relatively insensitive to temperature over the temperature range investigated (48 – 85 °C), which is also consistent with previous studies.[23,28,31,37,40] The *T*-jump induced relaxation kinetics of the A-to-E backbone and sidechain mutants are similar to those of the wild-type. In spite of their decreased conformational stability, most of these mutants exhibit a folding rate that is either similar to or only slightly slower than that of wild-type HP-35 (Figure 3). Interestingly, the slowest folder is mutant L20λ, which is also the most destabilized relative to the wild-type, and it has a folding time constant that is more than three times that of HP-35 at 50 °C (i.e., 23.9 versus 7.4 μs).

Implications for the folding mechanism of HP-35

To assess whether specific native-state interactions are also formed in the folding transition state, we have further determined the Φ-value for each mutant at 50 °C. As shown (Table 1), none of the mutations lead to a Φ-value >0.5, suggesting that most of the native interactions perturbed by these mutations do not contribute significantly to the transition state ensemble structure of HP-35. In particular, each of the A-to-E mutations made in the helical regions, namely V9ϖ, A18α, and L28λ, exhibit a Φ-value of essentially zero, indicating that the helical H-bonds perturbed by these mutations are most likely formed after the folding transition state of HP-35. A zero or near-zero Φ-value could also result from A-to-E elimination of a helical H-bond that exists in both the unfolded and transition states (Figure 1C). However, such a scenario would not elicit the observed destabilization of the folded state (removing the H-bond donor would destabilize both the unfolded and folded states leading to little change in ΔG_f). In fact, the magnitude of $\Delta\Delta G_f$ for each A-to-E mutation is consistent with previous estimates of H-bond energies (Table 1).[7–12]

Previous studies have shown that the unfolded or denatured state of HP-35 contains native-like interactions and structural elements,[27,33,38,39] including hydrophobic contacts (27) between Phe residues that are critical to the native stability and helical structures.[19] As has been suggested in the ultrafast folding of other small proteins,[47–49] the presence of such residual structure might prompt one to assume that the denatured state of HP-35 is structurally biased towards the native state, thus leading to very fast folding. However, the low Φ-values of the A-to-E variants indicate that residual H-bonded helical content in the denatured state of HP-35, which has been estimated to be as high as 25% for helix II,[33] has little, if any, effect on its folding rate. Consistent with this picture, the recent isotope-edited *T*-jump IR study of Brewer *et al*.[40] indicates that helix II may unfold prior to the global unfolding of the protein. In other words, this helix is formed on the native side of the major folding free energy barrier. Their findings are consistent with our data, which suggest that most of the helical H-bonds of HP-35 have not yet been formed in the transition state. Brewer *et al*.[28] also demonstrate that modulating hydrophobic interactions in the denatured state of HP-35 also yields little change in its folding kinetics.

Taken together, these results suggest a folding mechanism for HP-35 wherein the formation of the secondary structures and the hydrophobic core is not rate-limiting. Instead, folding is

likely initiated by assembling one of the two loops, as observed in the folding of other proteins.[11,54–57] The strongest experimental evidence supporting this hypothesis derives from the folding energetics exhibited by the L20λ variant, which lacks the native (i, i+3) H-bond formed between the backbone carbonyl of F17 and the backbone NH of L20.[18] This A-to-E variant folds more than three times slower than the wild-type. It is highly unlikely that the i, i+3 H-bond, which aids in defining the turn between helices II and III in the wild-type protein, is formed in the denatured state, as it is very important to the stability of the native state, as is evident from the large perturbation free energy of the L20λ variant (Table 1). If this turn is indeed important in initiating folding of HP-35, one would expect P21 to also play a critical role by restricting movement of the inter-helical region, as suggested by a recent folding simulation study.[41] However, the FT-IR data of P21A (Figure S3 in Supporting Information) show that it is prone to aggregate, indicating that its folding free energy landscape is drastically different from that of the wild-type, as well as the other mutants of HP-35. While the aggregation propensity of P21A renders Φ-value analysis inconclusive, that P21 is an important residue in the folding of chicken HP-35 demonstrates the importance of this turn. In fact, NMR studies[58,59] have shown that the human villin headpiece subdomain is unable to fold to its native structure when P21, which is conserved in the headpiece sequences of many F-actin binding cytoskeletal proteins,[60] is substituted by alanine.

The Q25nL variant exhibits a negative Φ-value (Table 1). In the native state of HP-35, Q25 forms sidechain to backbone H-bonds with the carbonyls of L20 and A18, which facilitate the productive docking of helices II and III. Thus, it is not surprising that eliminating these native H-bonds decreases the stability of the folded state. What is interesting, however, is that the resultant mutant folds faster than the parent protein (Table 1), in consequence yielding a negative Φ-value. Although the interpretation of negative Φ-values is often not straightforward and is in fact debated,[61] we interpret the faster folding rate of Q25nL as most likely arising from mutation and hence disruption of certain nonnative H-bonds in the unfolded state. Destabilization of the unfolded state would consequently decrease the folding free energy barrier of this mutant. Eaton and coworkers[31,37] have similarly shown that mutants K24nL and K24nL/K29nL also fold faster than the wild-type protein.

Excluding G11A, mutations in the first loop (i.e., G11γ, and T13A) do not change the folding rate significantly (the Φ-value of T13A is expected to have a large uncertainty because of the small $\Delta\Delta G_f$), indicating that this region is relatively unstructured in the folding transition state. The exception is G11A which exhibits a folding time which is twice that of the wild-type peptide. Position 11 is largely solvent exposed, and the mutation from Gly to Ala is unlikely to remove or modulate any interactions of the sidechain. However, since Gly is the most flexible amino acid, it is likely that mutation to Ala does reduce the conformational freedom of the loop, as indicated by the disparity between these two residues in their preferences for the native dihedral angles at position 11.[62] However, it may not be the difficulty of attaining the native state conformation which restricts and slows folding, but rather the loss of loop flexibility throughout the folding process. The Gly flexibility may be important in allowing the loop to be dynamic, giving helix I the ability to dock against helices II and III without conformational restraints from the backbone. This idea is consistent with the results of *ab initio* folding of HP-35 by Lei and Duan.[41] A similar scenario has also been suggested by Fersht and coworkers[63] in their description of the mechanistic role of the turn in the folding of the WW domain.

Taken together, our results suggest that HP-35 folds via an early transition state wherein very few native contacts have been formed. While the current study alone does not rule out other possibilities, for example a small Φ-value could arise from a heterogeneous transition state ensemble the population of which shifts in response to a destabilizing mutation, the above picture is consistent with the fast folding behavior of HP-35 as well as the results of other

mutational studies.[45] For example, Dyer and coworkers[28] have shown that the folding rate of mutant F6L/F10L is similar to that of the wild-type, although these mutations destabilize the hydrophobic core. Similarly, Eaton and coworkers[23,31] have shown that the destabilized mutants F35A, A18S, and A18V also fold with a rate that is similar to that of the wild-type and the Φ-values are unusually low at 310 K.[45] Besides L20λ and G11A, other known mutants exhibiting significantly different folding times relative to wild-type HP-35 are K24nL and K24nL/K29nL.[31,37] Thus, these results indicate that neither hydrophobic core nor secondary structure formation provide a framework for fast folding of HP-35. Instead, folding is most likely initiated by the formation of a folding nucleus involving a turn between helices II and III.

The small size and fast folding kinetics of HP-35 have also made this subdomain an attractive model system for computational studies.[18,20–22,24,26,30,32,34–36,39,41–43,46] Interestingly, however, those computer simulations have revealed somewhat different folding mechanisms. For instance, some have identified specific tertiary interactions as being critical for folding, while others have observed early helix formation. While our experimental results are in part consistent with the molecular dynamics simulations of Duan and coworkers,[41,42] which identified formation of the turn between helices II and III as rate-limiting for HP-35 folding, additional mutational studies, especially those involving residues in the turn or loop regions, as well as experiments capable of revealing the role of the solvent, are warranted in order to provide a more complete view of the folding mechanism of HP-35. Moreover, analysis of the relaxation kinetics of wild-type HP-35 and HP-35 variants using other theoretical models, such as those based on Langevin dynamics simulations,[64,65] may provide further insights into the folding energy landscape of this protein. Finally, in light of the backbone-based theories of protein folding,[66] it also would be interesting to investigate whether the formation of helical H-bonds in other proteins are rate-limiting for folding.

Conclusions

Using a small helical protein, the villin headpiece subdomain (HP-35), as a model, we demonstrated that amide-to-ester backbone mutagenesis is capable of revealing the role of individual backbone H-bonds in the folding transition state of a helical protein. This approach, when used in conjunction with traditional Φ-value analysis, is capable of providing new insight into the microscopic details of protein folding. For HP-35, our backbone mutagenesis studies show that removal of native helical H-bonds does not appreciably change the folding rate, but does significantly alter stability, suggesting that H-bond-mediated helix formation largely occurs after the transition state is surmounted. Interestingly, backbone mutagenesis of L20, eliminating an (i, i+3) H-bond located in the second loop, dramatically slows the folding rate, which implies that the turn between helices II and III is involved in the formation of a folding nucleus. Integrating our data with that of others indicates that the folding transition state of HP-35 does not contain a significant amount of native tertiary contacts and/or H-bonded helices, but may contain a H-bonded loop, which is consistent with the relatively small folding free energy barrier[44] and thus the ultrafast folding kinetics of HP-35.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgement

We gratefully acknowledge financial support from the National Institutes of Health (GM-065978, GM-051105 and RR-01348).

References

1. Pauling L, Corey RB, Branson HR. Proc. Natl. Acad. Sci. U.S.A 1951;37:205–211. [PubMed: 14816373]
2. Pauling L, Corey RB. Proc. Natl. Acad. Sci. U.S.A 1951;37:729–740. [PubMed: 16578412]
3. Dill KA. Biochemistry 1990;29:7133–7155. [PubMed: 2207096]
4. Baldwin RL. J. Biol. Chem 2003;278:17581–17588. [PubMed: 12582164]
5. Fersht AR. Proc. Natl. Acad. Sci. U.S.A 1995;92:10869–10873. [PubMed: 7479900]
6. Kim PS, Baldwin RL. Ann. Rev. Biochem 1982;51:459–489. [PubMed: 6287919]
7. Koh JT, Cornish VW, Schultz PG. Biochemistry 1997;36:11314–11322. [PubMed: 9298950]
8. Beligere GS, Dawson PE. J. Am. Chem. Soc 2000;122:12079–12082.
9. Blankenship JW, Balambika R, Dawson PE. Biochemistry 2002;41:15676–15684. [PubMed: 12501196]
10. Deechongkit S, Dawson PE, Kelly JW. J. Am. Chem. Soc 2004;126:16762–16771. [PubMed: 15612714]
11. Deechongkit S, Nguyen H, Powers ET, Dawson PE, Gruebele M, Kelly JW. Nature 2004;430:101–105. [PubMed: 15229605]
12. Powers ET, Deechongkit S, Kelly JW. Adv. Pro. Chem 2006;72:39–78.
13. Yang XY, Wang M, Fitzgerald MC. J. Mol. Biol 2006;363:506–519. [PubMed: 16963082]
14. Wang M, Wales TE, Fitzgerald MC. Proc. Natl. Acad. Sci. U.S.A 2006;103:2600–2604. [PubMed: 16473949]
15. Gao J, Kelly JW. Protein Sci 2008;17:1096–1101. [PubMed: 18434500]
16. Matouschek A, Kellis JT Jr, Serrano L, Fersht AR. Nature 1989;340:122–126. [PubMed: 2739734]
17. McKnight CJ, Doering DS, Matsudaira PT, Kim PS. J. Mol. Biol 1996;260:126–134. [PubMed: 8764395]
18. Duan Y, Wang L, Kollman PA. Proc. Natl. Acad. Sci. U.S.A 1998;95:9897–9902. [PubMed: 9707572]
19. Frank BS, Vardar D, Buckley DA, McKnight CJ. Protein Sci 2002;11:680–687. [PubMed: 11847290]
20. Islam SA, Karplus M, Weaver DL. J. Mol. Biol 2002;318:199–215. [PubMed: 12054779]
21. Fernandez A, Shen M-Y, Colubri A, Sosnick TR, Berry RS, Freed KF. Biochemistry 2003;42:664–671. [PubMed: 12534278]
22. Jang SM, Kim E, Shin S, Pak Y. J. Am. Chem. Soc 2003;125:14841–14846. [PubMed: 14640661]
23. Kubelka J, Eaton WA, Hofrichter J. J. Mol. Biol 2003;329:625–630. [PubMed: 12787664]
24. Lin CY, Hu CK, Hansmann UH. Proteins 2003;52:436–445. [PubMed: 12866054]
25. Wang M, Tang Y, Sato S, Vugmeyster L, McKnight CJ, Raleigh DP. J. Am. Chem. Soc 2003;125:6032–6033. [PubMed: 12785814]
26. Ripoll DR, Vila JA, Scheraga HA. J. Mol. Biol 2004;339:915–925. [PubMed: 15165859]
27. Tang Y, Rigotti DJ, Fairman R, Raleigh DP. Biochemistry 2004;43:3264–3272. [PubMed: 15023077]
28. Brewer SH, Vu DM, Tang Y, Li Y, Franzen S, Raleigh DP, Dyer RB. Proc. Natl. Acad. Sci. U.S.A 2005;102:16662–16667. [PubMed: 16269546]
29. Buscaglia M, Kubelka J, Eaton WA, Hofrichter J. J. Mol. Biol 2005;347:657–664. [PubMed: 15755457]
30. Carr JM, Wales DJ. J. Chem. Phys 2005;123:234901. [PubMed: 16392943]
31. Chiu TK, Kubelka J, Herbst-Irmer R, Eaton WA, Hofrichter J, Davies DR. Proc. Natl. Acad. Sci. U.S.A 2005;102:7517–7522. [PubMed: 15894611]
32. DeMori G, Colombo G, Micheletti C. Proteins 2005;58:459–471. [PubMed: 15521059]
33. Havlin RH, Tycko R. Proc. Natl. Acad. Sci. U.S.A 2005;102:3284–3289. [PubMed: 15718283]
34. Herges T, Wenzel W. Structure 2005;13:661–668. [PubMed: 15837204]
35. Bandyopadhyay S, Chakraborty S, Bagchi B. J. Chem. Phys 2006;125:084912. [PubMed: 16965062]
36. Ensign DL, Kasson PM, Pande VS. J. Mol. Biol 2006;374:806–816. [PubMed: 17950314]

37. Kubelka J, Chiu TK, Davies DR, Eaton WA, Hofrichter J. J. Mol. Biol 2006;359:546–553. [PubMed: 16643946]
38. Tang Y, Goger MJ, Raleigh DP. Biochemistry 2006;45:6940–6946. [PubMed: 16734429]
39. Wickstrom L, Okur A, Song K, Hornak V, Raleigh DP, Simmerling CL. J. Mol. Biol 2006;360:1094–1107. [PubMed: 16797585]
40. Brewer SH, Song B, Raleigh DP, Dyer RB. Biochemistry 2007;46:3279–3285. [PubMed: 17305369]
41. Lei H, Duan Y. J. Mol. Biol 2007;370:196–206. [PubMed: 17512537]
42. Lei H, Wu C, Liu H, Duan Y. Proc. Natl. Acad. Sci. U.S.A 2007;104:4925–4930. [PubMed: 17360390]
43. Chen J, Brooks CL. Phys. Chem. Chem. Phys 2008;10:471–481. [PubMed: 18183310]
44. Godoy-Ruiz R, Henry ER, Kubelka J, Hofrichter J, Munoz V, Sanchez-Ruiz JM, Eaton WA. J. Phys. Chem. B 2008;112:5938–5949. [PubMed: 18278894]
45. Kubelka J, Henry ER, Cellmer T, Hofrichter J, Eaton WA. Proc. Natl. Acad. Sci. U.S.A 2008;105:18655–18662. [PubMed: 19033473]
46. Yang JS, Wallin S, Shakhnovich EI. Proc. Natl. Acad. Sci. U.S.A 2008;105:895–900. [PubMed: 18195374]
47. Mayor U, Guydosh NR, Johnson CM, Grossman JG, Sato S, Jas GS, Freund SMV, Alonso DOV, Daggett V, Fersht AR. Nature 2003;421:863–867. [PubMed: 12594518]
48. Zhu Y, Alonso DOV, Maki K, Huang C-Y, Lahr SJ, Daggett V, Roder H, DeGrado WF, Gai F. Proc. Natl. Acad. Sci. U.S.A 2003;100:15486–15491. [PubMed: 14671331]
49. Mok KH, Kuhn LT, Goez M, Day IJ, Lin JC, Andersen NH, Hore PJ. Nature 2007;447:106–109. [PubMed: 17429353]
50. Du DG, Bunagan MR, Gai F. Biophys. J 2007;93:4076–4082. [PubMed: 17704172]
51. Mukherjee S, Chowdhury P, Gai F. J. Phys. Chem. B 2007;111:4596–4602. [PubMed: 17419612]
52. Barth A, Zscherp C. Q. Rev. Biophys 2002;35:369–430. [PubMed: 12621861]
53. Miyazawa T, Blout ER. J. Am. Chem. Soc 1961;83:712–719.
54. Du D, Zhu Y, Huang C-Y, Gai F. Proc. Natl. Acad. Sci. U.S.A 2004;101:15915–15920. [PubMed: 15520391]
55. Du D, Gai F. Biochemistry 2006;45:13131–13139. [PubMed: 17073435]
56. Du D, Tucker MJ, Gai F. Biochemistry 2006;45:2668–2678. [PubMed: 16489760]
57. Li H, Frieden C. Proc. Natl. Acad. Sci. U.S.A 2007;104:11993–11998. [PubMed: 17615232]
58. Vermeulen W, Van Troys M, Bourry D, Dewitte D, Rossenu S, Goethals M, Borremans FAM, Vanderkerckhove J, Martins JC, Ampe C. J. Mol. Biol 2006;359:1277–1292. [PubMed: 16697408]
59. Piana S, Laio A, Marinelli F, Van Troys M, Bourry D, Ampe C, Martins JC. J. Mol. Biol 2008;375:460–470. [PubMed: 18022635]
60. Vermeulen W, Vanhaesebrouck P, Van Troys M, Verschueren M, Fant F, Goethals M, Ampe C, Martins CJ, Borremans FAM. Protein Sci 2004;13:1276–1287. [PubMed: 15096633]
61. Merlo C, Dill KA, Weikl TR. Proc. Natl. Acad. Sci. U.S.A 2005;102:10171–10175. [PubMed: 16009941]
62. Scott KA, Alonso DOV, Sato S, Fersht AR, Daggett V. Proc. Natl. Acad. Sci. U.S.A 2007;104:2661–2666. [PubMed: 17307875]
63. Sharpe T, Jonsson AL, Rutherford TJ, Daggett V, Fersht AR. Protein Sci 2007;16:2233–2239. [PubMed: 17766370]
64. Xu Y, Purkayastha P, Gai F. J. Am. Chem. Soc 2006;128:15836–15842. [PubMed: 17147395]
65. Liu F, Dumont C, Zhu YJ, DeGrado WF, Gai F, Gruebele M. J. Chem. Phys 2009;130:061101. [PubMed: 19222256]
66. Rose GD, Fleming PJ, Banavar JR, Maritan A. Proc. Natl. Acad. Sci. U.S.A 2006;103:16623–16633. [PubMed: 17075053]

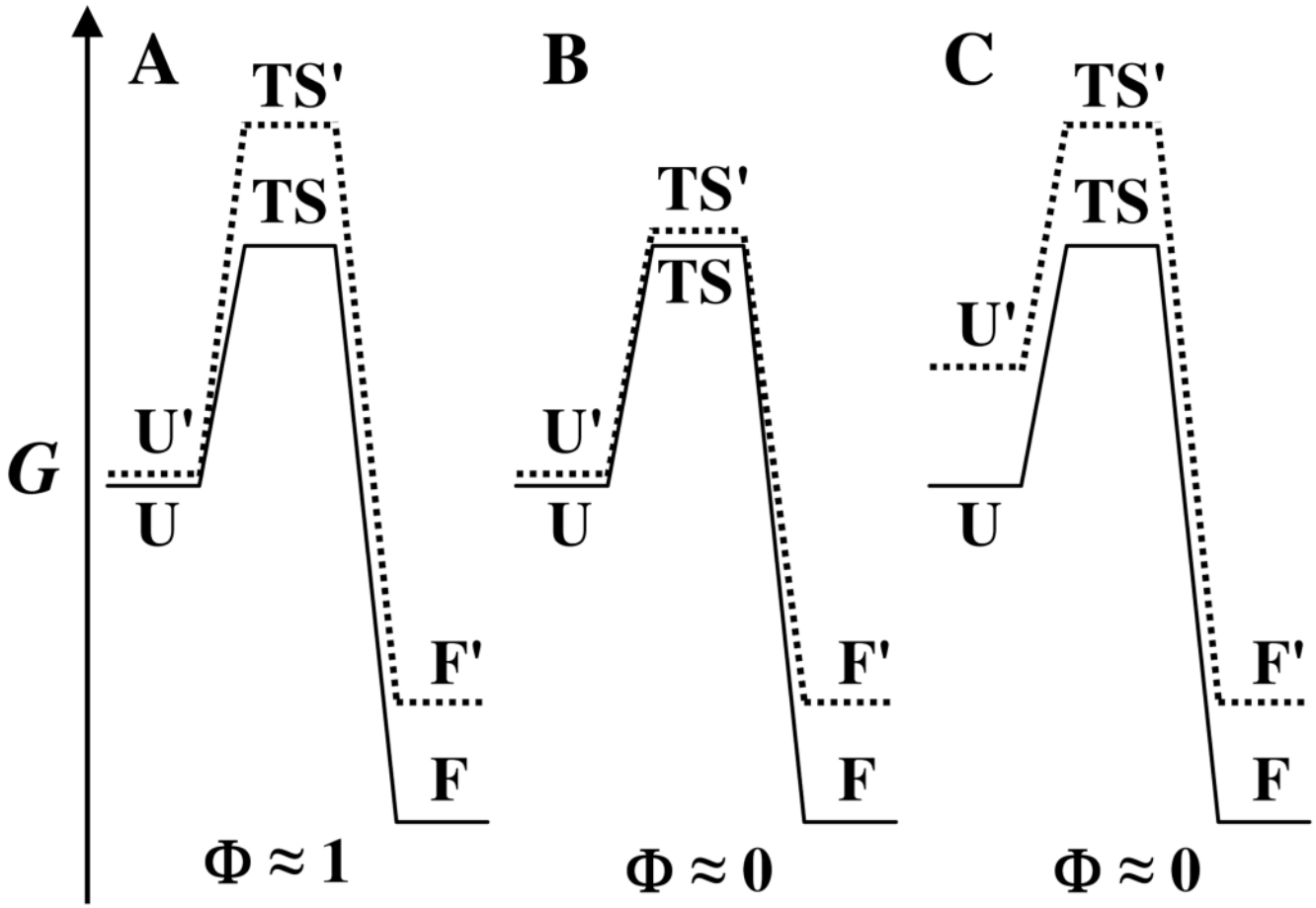

Figure 1.
Schematic illustration of how an A-to-E mutation might change the free energy of the unfolded (U), transition (TS), and folded (F) states (solid lines - wild-type, dashed lines - mutants) in Φ-value analysis. (A) Elimination of a native H-bond that is also formed in the transition state, (B) elimination of a native H-bond that is not formed in the transition state, and (C) elimination of a native H-bond that is also formed in the unfolded state.

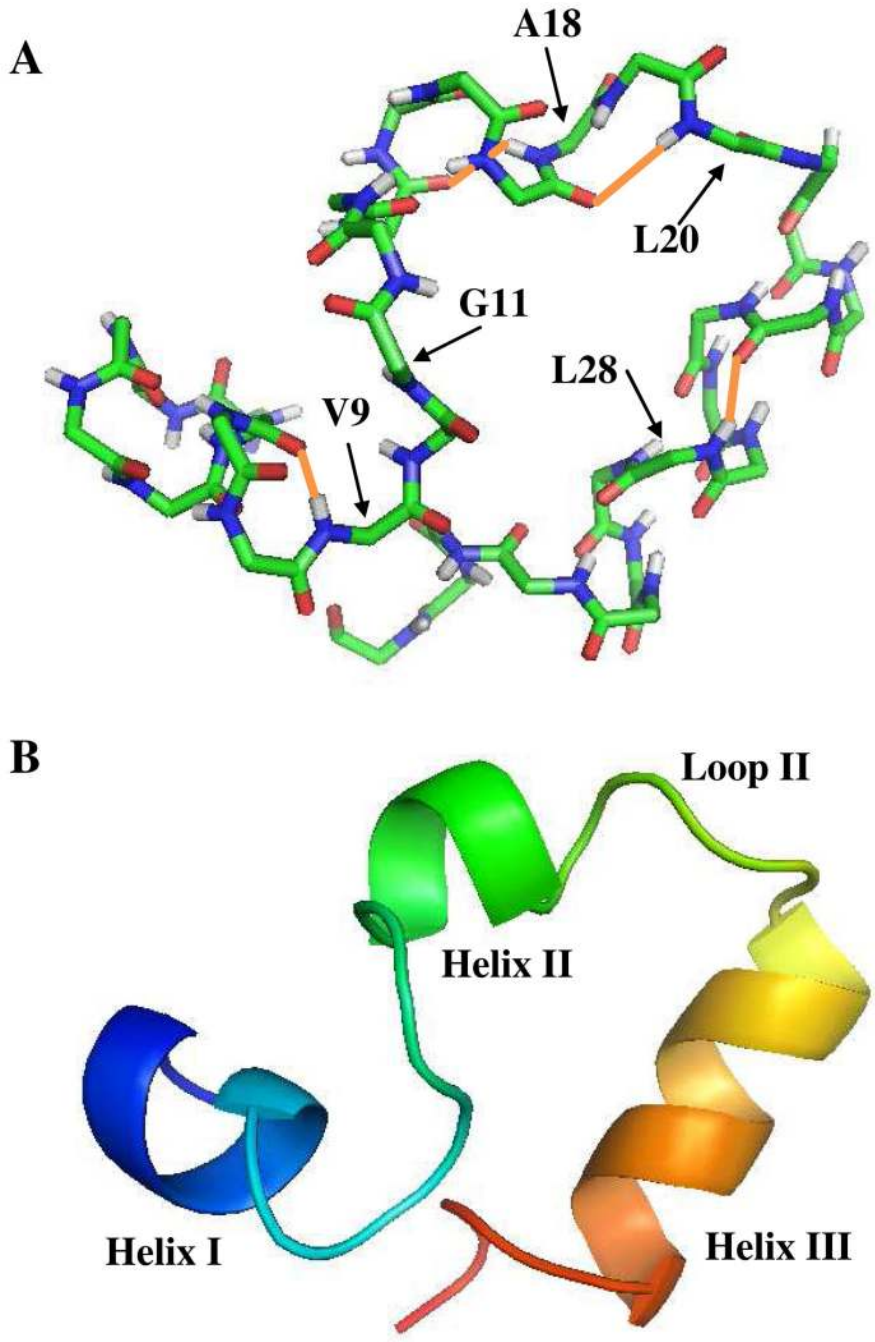

Figure 2.
NMR structure of HP-35 (1VII). The figure was generated using the program PyMol (http://pymol.sourceforge.net/).

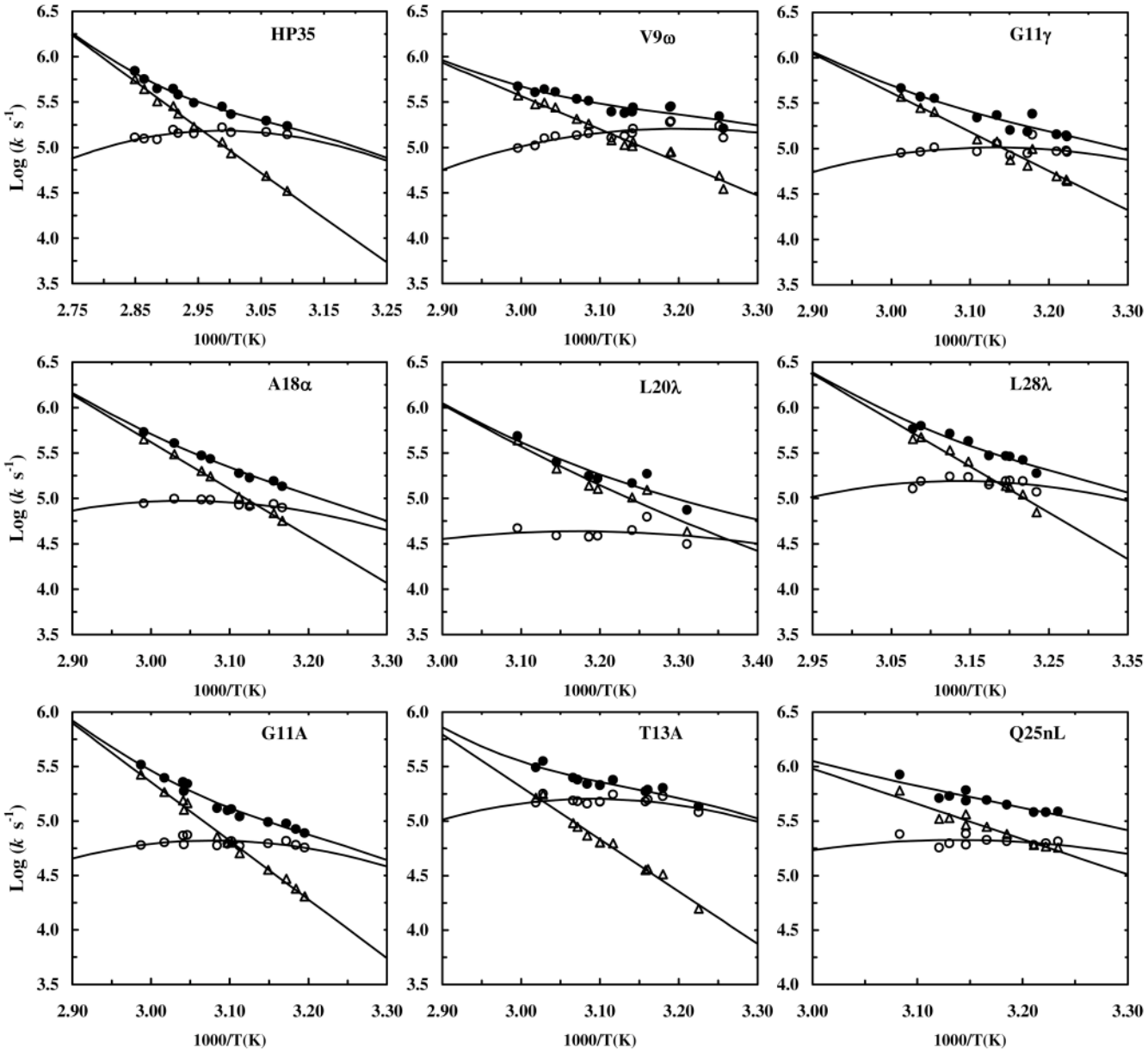

Figure 3.
Arrhenius plots of the observed relaxation rate constant (solid circles) as well as the folding (open circles) and unfolding (open triangles) rate constants for wild-type HP-35 and mutants, as indicated. Lines are fits to the Eyring equation.

Table 1

Summary of thermodynamic and kinetic data for wild-type HP-35 and mutants obtained at 50 °C.

	HP35	V9ω	G11γ	A18α	L20λ	L28λ	G11A	T13A	Q25nL
ΔG_f (kJ/mol)	−3.9	0.4	1.2	1.0	6.0	2.8	0.2	−2.1	2.2
$\Delta\Delta G_f$ (kJ/mol)*	0	4.3	5.1	4.9	9.9	6.7	4.1	1.8	6.1
τ_f (μs)	7.4	7.0	9.8	10.9	23.9	6.6	15.1	6.2	4.8
Φ_f		−0.04	0.15	0.20	0.32	−0.05	0.47	−0.26	−0.19

*$\Delta\Delta G_f = \Delta G_{f,\text{ mutant}} - \Delta G_{f,\text{ HP-35}}$