Folding of A Three-Helix Bundle at the Folding Speed Limit

Ting Wang, Yongjin Zhu, and Feng Gai*

Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104

Received: January 24, 2004; In Final Form: February 13, 2004

We show in this Letter that a double mutant (K5I/K39V) of $1prb_{7-53}$, the GA module of an albumin binding domain, has a maximum folding rate constant of $\sim 1\ (\mu s)^{-1}$. This value is comparable to the estimated theoretical speed limit for protein folding. In addition, we found that the mean hydrophobicity of a given tertiary fold plays an important role in controlling its folding rate.

Introduction

There is growing interest in studying/identifying proteins that are capable of folding on ultrafast time scale,[1−6] e.g., microsecond. The goal is to determine the speed limit of protein folding and also provide model systems for computer simulation.[7] Though different estimates[8−12] have been put forward, the folding speed limit is probably between 10^6 and $10^7\ s^{-1}$ for single domain proteins. Here we report the folding kinetics of two mutants of $1prb_{7-53}$, the GA module of an albumin binding domain.[13] Laser-induced temperature-jump (*T*-jump) infrared (IR) studies show that both mutants fold extremely rapidly, and one of them even folds at the semiempirical $\sim 1\ (\mu s)^{-1}$ folding speed limit suggested by Hagen et al.[9]

$1prb_{7-53}$ adopts a simple three-helix bundle structure in solution.[13] Because of its small size (47 residues) and simple topology, $1prb_{7-53}$ constitutes a good candidate for exploring fast folding. In fact, a recent simulation study carried out by Takada,[14] who employed a coarse-grained protein model that takes into account solvent effects, suggests that this protein can fold within 1 μs from its unfolded state. Indeed, *T*-jump IR experiments confirm the fast folding nature of $1prb_{7-53}$ and reveal that its minimum folding time constant is $\sim 6\ \mu$s (data not shown). This renders $1prb_{7-53}$ one of the fastest folding proteins known to date.

Although the native sequence of $1prb_{7-53}$ is capable of ultrafast folding, its folding rate may be further optimized if frustrations or energetic components that hamper folding can be diminished or even eliminated. Herein, we focus on replacing buried polar side chains because it is well-known that the formation of buried polar interactions decreases the folding rate.[15] The NMR structures of 1prb (PDB code: 1prb) clearly show that the side chain of K5 is partially buried and that of K39 is more than 90% buried. Therefore, we sought to replace these polar side chains with hydrophobic groups. Specifically, two mutants, i.e., K5I and K5I/K39V, were studied.

Experimental Methods

The proteins used in the current study were synthesized by employing the standard Fmoc-protocol and purified by reverse phase HPLC. The identity of the samples was verified by electrospray-ionization mass spectroscopy. Residual TFA from the synthesis was removed by lyophilization against a 0.1 M DCl solution, which also allowed amide hydrogen−deuterium (H−D) exchange. A complete H−D exchange is important because the amide vibrational modes of deuterated protein samples are often different from those of hydrogenated protein samples. Static IR spectra were collected on a Nicolet Magna 860 spectrometer equipped with a MCT detector. A thermostated sample cell with an optical path length of 52 μm and two compartments allowed the measurements of both the sample and the buffer under the same conditions. The protein samples used in both equilibrium and time-resolved measurements were prepared by dissolving the lyophilized solids directly in 50 mM phosphate buffer solutions (D_2O, pH* = 7.0). The final concentration was 2−3 mM. Circular dichroism (CD) data were collected on an AVIV 62DS spectropolarimeter using a 1 mm quartz cell. The protein concentration used in the CD measurements was about 20−25 μM, which was determined optically by the single tyrosine absorbance at 276 nm using $\epsilon_{276} = 1450\ cm^{-1}\ M^{-1}$.

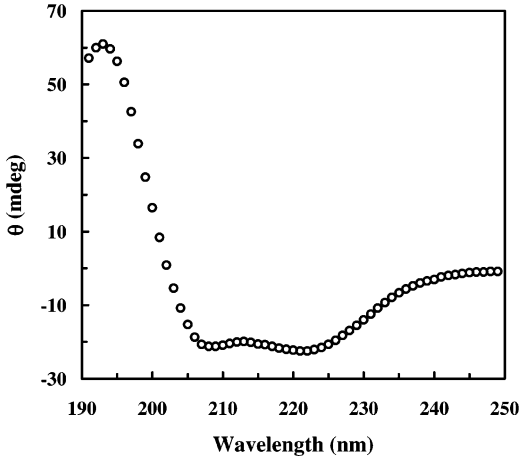

Figure 1. Far UV CD spectrum of K5I/K39V in water (20 mM phosphate buffer, pH 7).

The *T*-jump IR setup has been described in detail previously.[16] Briefly, the 1.9 μm *T*-jump pulse was generated via Raman shifting the fundamental output of a Nd:YAG laser in H_2. The *T*-jump induced transient absorbance changes were measured by a continuous wave (CW) IR diode laser in conjunction with a 50 MHz MCT (HgCdTe) detector and a digital oscilloscope. Similar to that used in the equilibrium experiments, a thermostated two-compartment sample holder was employed to allow the separate measurements of the transient signals of both the sample and the buffer under identical conditions. The buffer measurements provide information for both background subtraction and *T*-jump amplitude calibration.

Results and Discussion

As shown (Figure 1), the far UV CD spectrum of K5I/K39V

* To whom correspondence should be addressed. E-mail: gai@sas.upenn.edu.

10.1021/jp049652q CCC: $27.50
Published on Web 02/26/2004

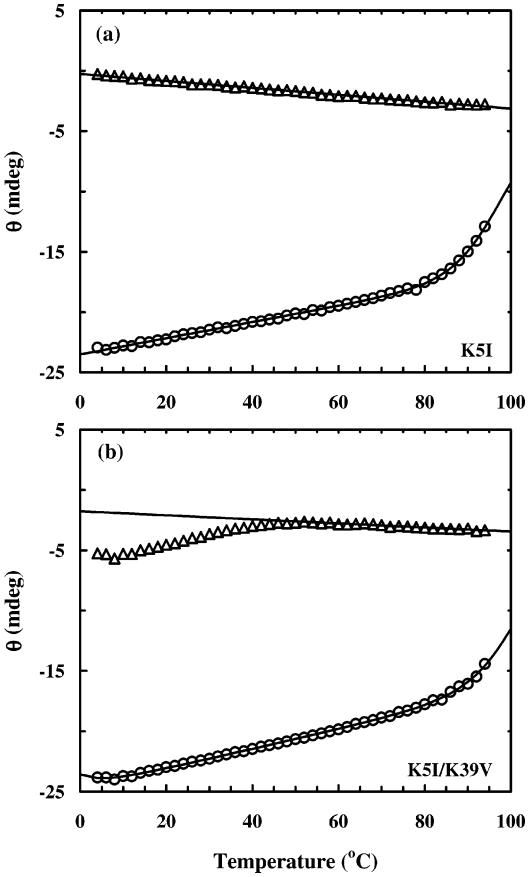

Figure 2. Ellipticity of (a) K5I and (b) K5I/K39V in water (○) and 6 M GuHCl solution (△) at 222 nm as a function of temperature. As indicated, the measurements in 6 M GuHCl solution allowed the determination of the unfolded CD baseline (i.e., $\theta_U(T)$). Solid lines are best fits using eqs 1−3, and the resulting thermodynamic parameters are given in the text.

at 25 °C exhibits characteristic features of helical proteins and is almost identical to that of the wild type (data not shown), indicating that the mutation did not change the helical content. Following common practice, we probed the equilibrium thermal unfolding transition of K5I and K5I/K39V by monitoring the CD signal at 222 nm. As indicated (Figure 2), at neutral pH both mutants exhibit very high thermal stability with a thermal melting temperature (T_m) > 90 °C. To further evaluate the unfolding thermodynamics of both proteins, we fit the temperature-dependent ellipticity (i.e., $\theta(T)$) to an apparent two-state model, namely,

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

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

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

Here, $\theta_F(T)$ is the pre-transition CD baseline, $\theta_U(T)$ is the post-transition CD baseline, $K_{eq}(T)$ is the equilibrium constant for unfolding, $T_m = \Delta H_m/\Delta S_m$ is the thermal melting temperature, ΔH_m is the enthalpy change at T_m, ΔS_m is the entropy change at T_m, and ΔC_p is the heat capacity change, which has been assumed here to be temperature independent. In the fit, $\theta_F(T)$ was treated as a linear line (i.e., $\theta_F(T) = a + bT$, where a and b are constants), whereas $\theta_U(T)$ was determined from the CD signals of the unfolded state of K5I and K5I/K39V in 6 M guanidinium hydrochloride (GuHCl) solution.[17] The best fits for both mutants yielded the following thermodynamic parameters for unfolding: (K5I) $\Delta H_m = 48$ kcal mol^{-1}, $\Delta S_m = 130$ cal mol^{-1} K^{-1}, $\Delta C_p = 192$ cal mol^{-1} K^{-1}, $T_m = 96$ °C; (K5I/K39V) $\Delta H_m = 51$ kcal mol^{-1}, $\Delta S_m = 137$ cal mol^{-1} K^{-1}, $\Delta C_p = 804$ cal mol^{-1} K^{-1}, $T_m = 99$ °C. It is worth pointing out that though the folding thermodynamics can be evaluated by this approach, the resulting parameters may be subject to large uncertainties due to the fact that even to the highest temperature the CD signals only catch part of the thermal unfolding transition.

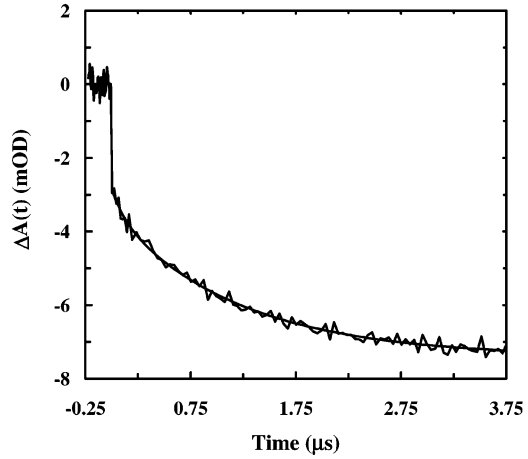

Figure 3. Representative relaxation trace of K5I/K39V corresponding to a T-jump of 84.3 to 94.5 °C, probed at 1630 cm^{-1}. The solid line is the fit to the following function, $\Delta A(t) = A[1 - B \exp(-t/\tau)]$, with $A = -7.4$, $B = 0.53$, and $\tau = 1.2$ μs.

The FTIR spectra of K5I and K5I/K39V in the amide I′ (amide I band of proteins in D_2O) region show typical features of helical proteins (Supporting Information). Globally analyzing[18] the spectra collected between 6 and 82 °C indicates that no significant unfolding takes place below 80 °C, consistent with the CD results. As shown (Figure 3), the T-jump relaxation kinetics of K5I/K39V following a temperature jump of ∼10 °C, monitored at 1630 cm^{-1}, where helical conformations absorb,[16] contain two phases. The fast phase, which is not resolved in time due to the ∼15 ns response time of the detection system, is the result of temperature-induced spectral shift and/or imperfect background subtraction. Within experimental uncertainties, the slow phase can be modeled by a single-exponential function, indicating that K5I/K39V folds in a two-state manner. Qualitatively, the T-jump relaxation kinetics of K5I show behaviors similar to those of K5I/K39V. In a T-jump experiment, the observed relaxation measures the time required for the system to relax from the initial equilibrium (corresponding to the initial temperature) to the final equilibrium (corresponding to the final temperature). For a two-state scenario, it is easy to show that $k_R = k_F + k_U$, where k_R is the measured relaxation rate constant, whereas k_F and k_U are the folding and unfolding rate constants, respectively. Therefore, k_F and k_U can be determined in two-state folding if K_{eq} is known because $K_{eq} = k_F/k_U$. Using these relationships, we calculated the temperature-dependent folding and unfolding rate constants for both K5I and K5I/K39V based on the measured relaxation rate constants as well as the thermodynamic parameters determined from the CD experiment. As shown (Figure 4), both proteins exhibit ultrafast folding behaviors. For example, the maximum folding rate constant for K5I is $(2.5 \pm 0.5\ \mu s)^{-1}$, whereas K5I/K39V folds even faster. The latter has a maximum folding rate constant of $(1 \pm 0.2\ \mu s)^{-1}$, which is equal to the estimated

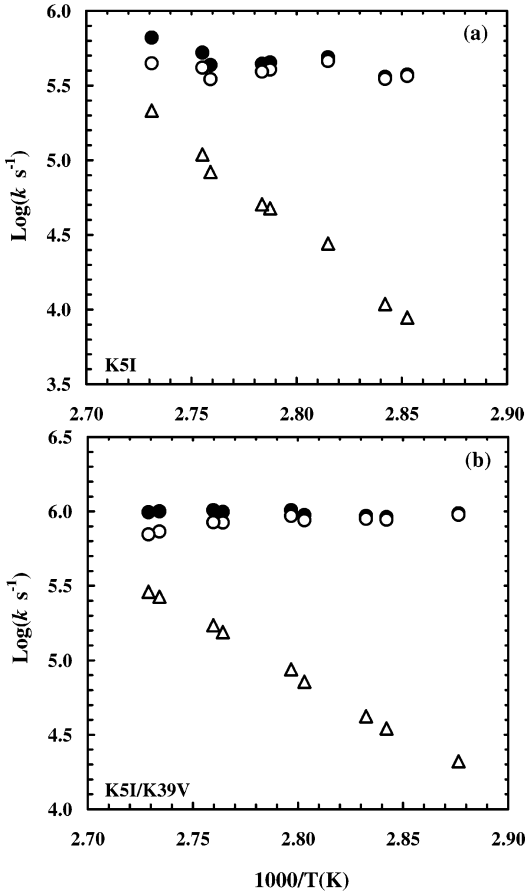

Figure 4. (a) Arrhenius plot of the observed (●), folding (○), and unfolding (△) rate constants of K5I. (b) Arrhenius plot of the observed (●), folding (○), and unfolding (△) rate constants of K5I/K39V.

folding speed limit suggested by Hagen et al.[9] Although the uncertainties associated with the equilibrium thermal unfolding CD analysis would lead to uncertainties in the determined equilibrium constants, the conclusions reached above are not affected because most of the kinetic measurements were made at temperatures lower than T_m. Under this condition the folding rate dominates the relaxation rate.

Because helix formation is a submicrosecond event,[16] helix bundle proteins tend to fold on very fast time scales due to the fact that they have relatively small contact orders.[19] For a given fold or topology, however, the folding rate may be fine-tuned by changing its hydrophobic content. Using a self-solvation corrected hydrophobicity scale (ΔG) of individual amino acids, which is obtained on the basis of the transfer of solutes from water to alkanes,[20] we calculated the mean hydrophobicity ($\Delta\bar{G}$) of the three proteins studied here and found that for the wild type $1prb_{5-73}$, $\Delta\bar{G} = 0.071$ kcal mol^{-1}; for K5I, $\Delta\bar{G} = -0.044$ kcal mol^{-1}; and for K5I/K39V, $\Delta\bar{G} = -0.142$ kcal mol^{-1}. Indeed, the maximum folding rates of these proteins correlate with $\Delta\bar{G}$ (Figure 5) with a relation that a faster folding rate coincides with a smaller $\Delta\bar{G}$ (or higher mean hydrophobicity). This dependence underscores the importance of the average hydrophobicity of a protein in controlling its folding rate. A similar conclusion was also drawn recently by Calloni et al.[21] Additionally, these results are also consistent with those obtained by Klimov et al., who have shown by using simple cubic lattice models that the folding times for two-state folders correlate with their stabilities.[22]

The folding of $1prb_{7-53}$ brings side chains into tertiary contacts that are separated by up to 30 residues in sequence. Two residues separated by this distance in sequence would be

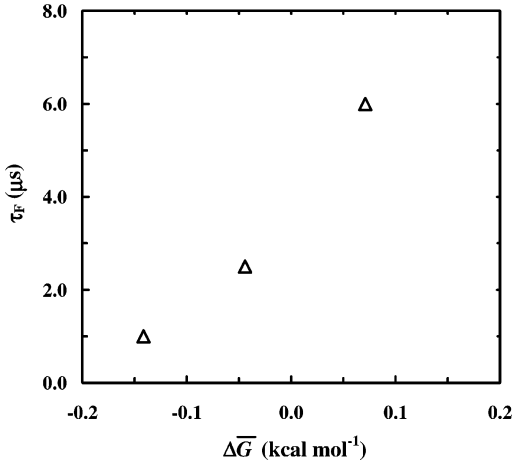

Figure 5. Minimum folding time vs mean hydrophobicity ($\Delta\bar{G}$) for $1prb_{7-53}$, K5I, and K5I/K39V.

expected to collide on the 0.3−1 μs time scale,[9] yet one of its mutants, i.e., K5I/K39V, is able to achieve the fully folded state on the same time scale. This suggests that the folding energy landscape of K5I/K39V is probably very smooth and the free energy barrier for folding is small. Nevertheless, the identification here of a protein capable of folding with a first-order rate constant of $\sim 1\ (\mu s)^{-1}$ may also suggest that the prefactor of the transition state folding models is quite likely $>1\ (\mu s)^{-1}$, at least for small, single domain proteins.[10,11] Consistent with this view, Thirumalai and co-workers[23] also suggested that the $\sim 1\ (\mu s)^{-1}$ could serve as a practical estimate for the prefactor. Similarly, Eaton and co-workers[24] suggested that the speed limit for a generic N-residue single-domain protein is approximately $N/(100\ \mu s)$. This relationship predicts that the maximum folding rate for $1prb_{7-53}$ is $\sim(0.5\ \mu s)^{-1}$. Taken together, these results suggest that the folding of K5I/K39V indeed approaches the folding speed limit and encounters only a small free energy barrier ($\sim$0.5 kcal mol^{-1}). Finally, although topology,[19] size,[23−25] and stability[26] have been shown to be important determinants of protein folding rate, results from this and several other studies suggest that other factors, such as the mean hydrophobicity, may also affect greatly the folding free energy barrier.

In summary, we show that a double mutant of $1prb_{7-53}$, K5I/K39V, exhibits the fastest folding rate known to date. Its maximum folding rate constant is $\sim 1\ (\mu s)^{-1}$, which is equal to the estimated theoretical speed limit for protein folding. In addition, we found that for a given tertiary fold its mean hydrophobicity plays an important role in fine-tuning its folding rate. In light of these findings, we believe that $1prb_{7-53}$ and its mutants are ideal model systems for further experimental and computational studies because of their small size, simple fold, and fast folding rate.

Acknowledgment. We thank the NSF (CHE-0094077 and DMR00-79909) for funding. We also thank Drs. J. G. Saven, S. Takada, and A. Tamura for helpful discussion.

Supporting Information Available: Temperature-dependent FTIR spectra of K5I and K5I/K39V. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Mayor, U.; Guydosh, N. R.; Johnson, C. M.; Grossmann, J. G.; Sato, S.; Jas, G. S.; Freund, S. M.; Alonso, D. O.; Daggett, V.; Fersht, A. R. *Nature* **2003**, *421*, 863.

(2) Yang, W. Y.; Gruebele, M. *Nature* **2003**, *423*, 193.

(3) Gillespie, B.; Vu, D. M.; Shah, P. S.; Marshall, S. A.; Dyer, R. B.; Mayo, S. L.; Plaxco, K. W. *J. Mol. Biol.* **2003**, *330*, 813.

(4) Qiu, L.; Pabit, S. A.; Roitberg, A. E.; Hagen, S. J. *J. Am. Chem. Soc.* **2002**, *124*, 12952.

(5) Kubelka, J.; Eaton, W. A.; Hofrichter, J. *J. Mol. Biol.* **2003**, *329*, 625.

(6) Zhu, Y.; Alonso, D. O.; Maki, K.; Huang, C. Y.; Lahr, S. J.; Daggett, V.; Roder, H.; DeGrado, W. F.; Gai, F. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 15486.

(7) Garcia, A. E.; Onuchic, J. N. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13898.

(8) Camacho, C. J.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 1277.

(9) Hagen, S. J.; Hofrichter, J.; Szabo, A.; Eaton, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 11615.

(10) Portman, J. J.; Takada, S.; Wolynes, P. G. *J. Chem. Phys*. **2001**, *114*, 5082.

(11) Chang, I. J.; Lee, J. C.; Winkler, J. R.; Gray, H. B. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 3838.

(12) Bieri, O.; Wirz, J.; Hellrung, B.; Schutlowski, M.; Drewello, M.; Kiefhaber, T. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9597.

(13) Johansson, M. U.; de. Château, M.; Björck, S.; Drakenberg, T.; Wikström, M. *FEBS Lett.* **1995**, *374*, 257.

(14) Takada, S. *Proteins* **2001**, *42*, 85.

(15) Waldburger, C. D.; Jonsson, T.; Sauer, R. T. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 2629.

(16) Huang, C. Y.; Getahun, Z.; Zhu, Y.; Klemke, J. W.; DeGrado, W. F.; Gai, F. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2788.

(17) Pace, C. N. *Methods Enzymol.* **1986**, *131*, 266.

(18) Huang, C. Y.; Wang, T.; Gai, F. *Chem. Phys. Lett.* **2003**, *371*, 731.

(19) Plaxco, K. W.; Simons, K. T.; Baker, D. *J. Mol. Biol.* **1998**, *277*, 985.

(20) Roseman, M. A. *J. Mol. Biol*. **1988**, *200*, 513.

(21) Calloni, G.; Taddei, N.; Plaxco, K. W.; Ramponi, G.; Stefani, M.; Chiti, F. *J. Mol. Biol*. **2003**, *330*, 577.

(22) Klimov, D. K.; Thirumalai, D. *J. Chem. Phys.* **1998**, *109*, 4119.

(23) Li, M. S.; Klimov, D. K.; Thirumalai, D. *Polymer* **2004**, *45*, 573.

(24) Kubelka, J.; Hofrichter, J.; Eaton, W. A. *Curr. Opin. Struct. Biol.*, in press.

(25) Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 6170.

(26) Dinner, A. R.; Karplus, M. *Nature Struct. Biol.* **2001**, *8*, 21.