doi:10.1016/j.jmb.2006.03.034 J. Mol. Biol. (2006) 359, 546–553

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

COMMUNICATION

Sub-microsecond Protein Folding

Jan Kubelka[1,3], Thang K. Chiu[2], David R. Davies[2], William A. Eaton[1]* and James Hofrichter[1]*

[1]*Laboratory of Chemical Physics National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892-0520, USA*

[2]*Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda MD 20892-0520, USA*

[3]*Department of Chemistry University of Wyoming Laramie, WY 82071, USA*

**Corresponding authors*

We have investigated the structure, equilibria, and folding kinetics of an engineered 35-residue subdomain of the chicken villin headpiece, an ultrafast-folding protein. Substitution of two buried lysine residues by norleucine residues stabilizes the protein by 1 kcal/mol and increases the folding rate sixfold, as measured by nanosecond laser T-jump. The folding rate at 300 K is $(0.7\ \mu s)^{-1}$ with little or no temperature dependence, making this protein the first sub-microsecond folder, with a rate only twofold slower than the theoretically predicted speed limit. Using the 70 ns process to obtain the effective diffusion coefficient, the free energy barrier height is estimated from Kramers theory to be less than ~1 kcal/mol. X-ray crystallographic determination at 1 Å resolution shows no significant change in structure compared to the single-norleucine-substituted molecule and suggests that the increased stability is electrostatic in origin. The ultrafast folding rate, very accurate X-ray structure, and small size make this engineered villin subdomain an ideal system for simulation by atomistic molecular dynamics with explicit solvent.

Published by Elsevier Ltd.

Keywords: protein crystallography; protein folding kinetics

The historical gap between the time-scale accessible to atomistic molecular dynamics trajectories and measurements of protein folding kinetics is rapidly narrowing.[1–6] Advances on the computational side have come from improved force fields, increased CPU power and the use of distributed computing. On the experimental side the introduction of optical triggering methods with nanosecond laser pulses[7–10] and NMR methods[11] have dramatically improved the time-resolution in investigating protein folding kinetics. A question that immediately arises is: how fast can a protein possibly fold?[12] There is as yet no rigorous theoretical estimate, analogous to the diffusion-limited rate of Smoluchowski for bimolecular reactions. However, both empirical and theoretical arguments have combined to produce an estimate of $\sim(N/100\ \mu s)^{-1}$ for a protein folding speed limit, where N is the number of residues in the polypeptide chain.[13] The notion of a speed limit has sparked the search for ultrafast folding proteins and the engineering of fast-folding proteins to make them fold even faster. There are now a number of proteins that fold in less than 100 μs.[1,2,11,14–26] In addition to narrowing the computation-experiment gap, another rationale for increasing folding speed is to eliminate the free energy barrier between folded and unfolded states to produce what has been called a "downhill" or "one-state" folder.[27–33] For downhill folders it may be possible to devise an experiment in which a rapid change in solution conditions converts the system from two-state to one-state, and thereby enables direct interrogation of intermediate structures along the folding pathways by spectroscopy.[34]

Here we report the results of protein engineering experiments aimed at increasing the folding rate of the 35 residue subdomain from the chicken villin headpiece (HP-35). HP-35 is the smallest naturally occurring polypeptide that autonomously folds into a globular structure. It consists of three helices surrounding a compact hydrophobic core. Figure 1 shows the structure of HP-35 recently determined by X-ray crystallography to 1 Å resolution.[35] Laser temperature jump experiments on wild-type HP-35 revealed distinctly bi-exponential kinetics, with a nanosecond fast phase and a slower, microsecond phase. We interpreted the slower phase as the

Abbreviations used: GdmCl, guanidinium chloride; HP-35, the 35 residue subdomain from the chicken villin headpiece.

E-mail addresses of the corresponding authors: eaton@helix.nih.gov; jameshof@niddk.nih.gov

0022-2836/$ - see front matter Published by Elsevier Ltd.

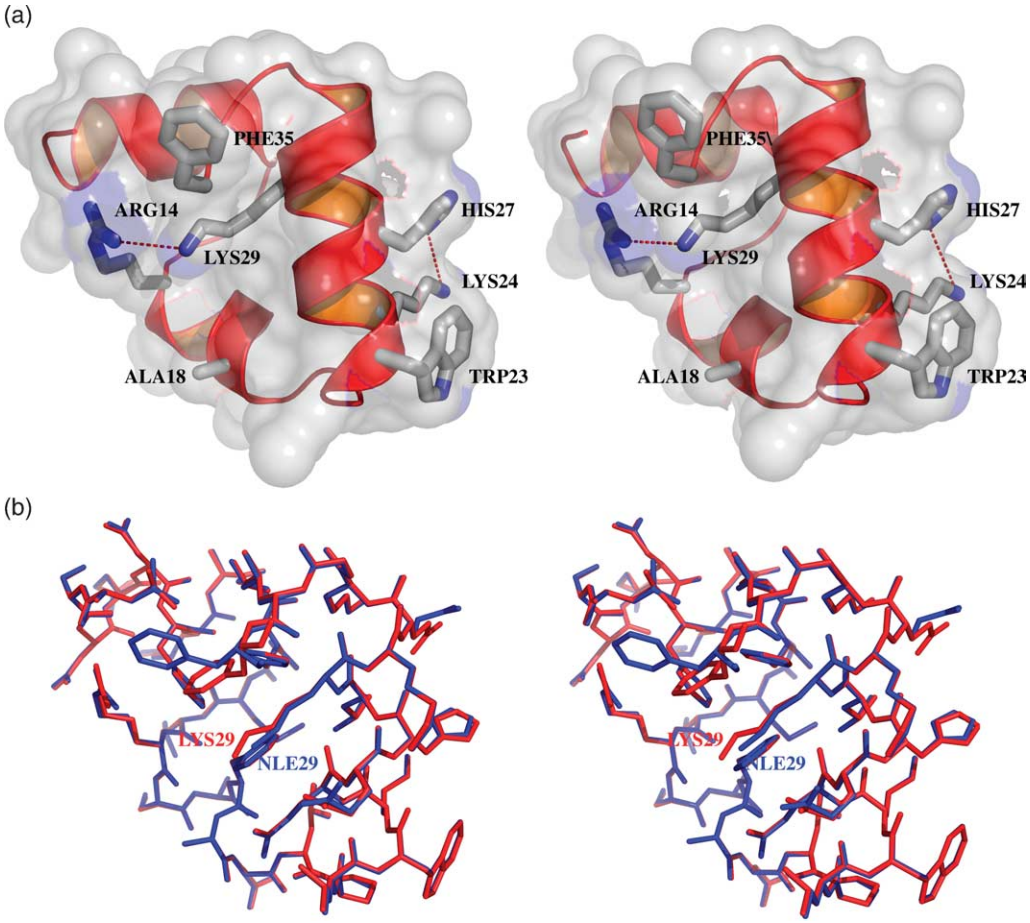

Figure 1. (a) X-ray structure of villin headpiece subdomain (PDB 1YRF) showing key residues:[35] Trp23, the fluorescence probe; His27, replaced Asn27 to quench Trp23 fluorescence upon helix formation (referred to as the wild-type in all of our studies); Phe35, no effect on stability or folding rate when replaced by alanine; Ala18, destabilized by replacement with valine or serine, but no effect on folding rate; Lys24, stabilizes protein and increases folding rate when replaced by norleucine (its ε-amino nitrogen is 6.1 Å from the nearest protonated nitrogen of the imidazole ring of His27); Lys29, stabilizes protein and increases folding rate when replaced by norleucine (its ε-amino nitrogen is 5.3 Å from the nearest charged amino atom of Arg14. (b) Comparison of the X-ray structures of Lys24Nle mutant (red) (PDB 1WY3) with Lys24Nle/Lys29Nle double mutant (blue) (PDB 2F4K). Removal of two charged amino groups in the Lys24Nle/Lys29Nle peptide resulted in a substantial decrease in solubility. The solubility was increased from ~15 mg/ml to at least 100 mg/ml by adding 7.5% (v/v) triflouroethanol (TFE). Crystals were grown by vapor diffusion at 4 °C from hanging drops containing 1 μl of 90 mg/ml peptide and 1 μl of precipitant (100 mM Bicine (pH 9), 800 mM $(NH_4)_2SO_4$ and 10% TFE) equilibrated over the 1 ml of the same solution in the reservoir. After one week, the reservoir solution was replaced with 1 ml of 100 mM Bicine (pH 9), 1.6 M $(NH_4)_2SO_4$ and 7.5% (w/w) TFE. Crystals isomorphous with the Lys24Nle mutant appeared after one week, and were cryoprotected in buffer lacking TFE: 100 mM Bicine (pH 9), 2.2 M $(NH_4)_2SO_4$, 25% (w/v) sucrose. X-ray diffraction data to 1.05 Å resolution were collected at 95 °C by Mail-In-Crystallography at the Advanced Photon Source (Argonne National Laboratory, Argonne, IL), and merged with lower resolution data previously collected on the same crystal at a home source. Statistics are: $R_{sym} = 8.6\%$, $I/\sigma = 22.5$, completeness = 96.3%, redundancy = 7.7. The structure was solved by molecular replacement with the Lys24Nle mutant as the starting model, and refined as described[35] to a final R_{work}/R_{free} of 13.95%/16.64% and RMS (bonds/angles) of 0.015 Å/0.033 Å. The RMSD among C^{α} atoms of the single and double Nle mutant structures is 0.08 Å, and no TFE molecule was observed.

overall unfolding/refolding kinetics and obtained a folding rate of $(4.3 \pm 0.6\ \mu s)^{-1}$ at 300 K.[2] Our interpretation of the microsecond phase as unfolding/refolding was confirmed by measurements of the kinetics of quenching of the tryptophan triplet state by contact formation with cysteine,[36] which occurs only in the unfolded chain and depends on the rate of exchange between the folded and unfolded conformation.[37,38]

The first attempt to increase folding speed came from theoretical studies. Using atomistic molecular dynamics simulations with implicit solvent Zagrovic *et al.* determined a folding time of 5 (+11, −3) μs, in remarkably good agreement with the measured time.[39] From a comparison of the successful and unsuccessful trajectories they suggested that dissociation of the terminal phenylalanine (Phe35) from a misfolded hydrophobic

cluster with the three core phenylalanine residues is rate-limiting. However, replacement of Phe35 with alanine changed neither the stability, consistent with almost complete exposure of Phe35 to solvent in the native structure (Figure 1(a)),[35] nor the folding rate.[2]

Another suggestion for increasing folding speed came from a combination of coarse-grained and atomistic molecular dynamics by Fernandez *et al.*[40] These authors argued that alanine 18 near the N terminus of the C-terminal helix (Figure 1(a)) influences C-terminal helix nucleation by desolvating the helical hydrogen bonds. Substitution of Ala18 by valine, which is more hydrophobic and therefore more effective in the desolvation, was predicted to destabilize the folded state, yet to dramatically speed up folding. By contrast, substitution of Ala18 by serine was predicted to result in a non-folding protein, because the C-terminal helix would not form. We have synthesized both mutants and tested these predictions experimentally. The results in Figure 2 show that both these substitutions decrease protein stability by about 0.6 kcal/mol, but have no effect on the folding rate (kinetic data for Ala18Val not shown).

We therefore turned to the X-ray structure to look for clues. HP-35 contains two buried lysine residues at positions 24 and 29. We have previously shown that removing the charge of Lys24 by substituting norleucine slightly stabilizes the protein, and all of the 0.5 kcal/mol increase in stability is reflected in the folding time, which decreases from 4.3($\pm$0.6) μs to 1.7($\pm$0.3) μs (norleucine, denoted Nle, is lysine without the ε-amino group).[35] The high resolution X-ray structure showed that the substitution increases burial of the aliphatic side-chain of residue 24, reducing its accessible surface area, which would increase stability.[35] Moreover, Zhu *et al.* had also shown that replacement of a buried lysine in a 47 residue three-helix bundle protein (prb_{7-53}, the GA module of an albumin binding domain) both stabilized and increased the folding rate.[23] Anticipating the possibility of increased stabilization and folding rate,[35] we also replaced Lys29 with norleucine and examined the structural, thermodynamic, and kinetic consequences.

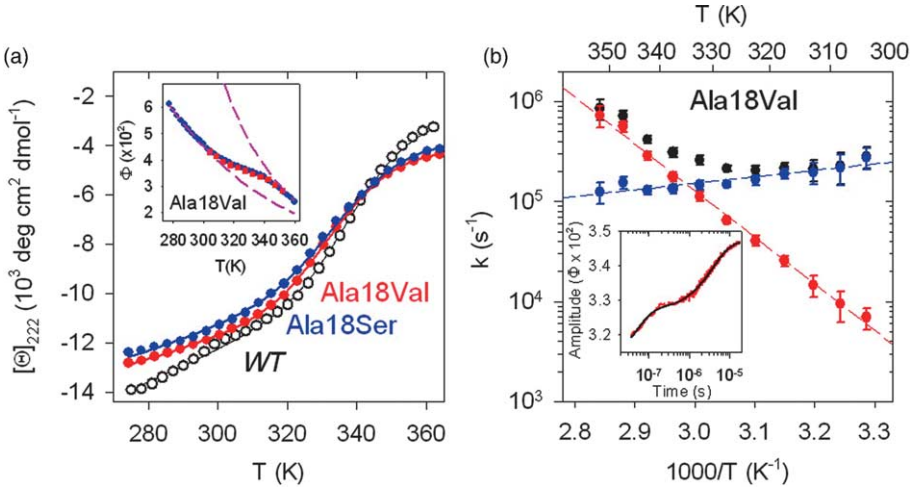

Figure 2. Folding thermodynamics and kinetics data for substitutions of Ala18. The peptides were synthesized by FMOC solid-phase synthesis and purified to >95% by HPLC, as evidenced by mass spectrometry. Equilibrium and kinetic data were collected at pH 4.8 (20 mM sodium acetate buffer) to ensure protonation of His27. (a) Thermal unfolding monitored by CD at 222 nm for Ala18Val (red circles) and Ala18Ser (blue circles) mutants of HP-35 compared to wild-type (black open circles).[2] The CD spectra were measured on a JASCO J-720 spectropolarimeter in a 100 μm pathlength cuvette and at protein concentration of ~200 μM. Inset: equilibrium thermal unfolding of HP-35 Ala18Val monitored by tryptophan fluorescence quantum yield (Φ). The fluorescence was measured using a SPEX Fluorolog spectrofluorimeter in a standard 1 cm cuvette and at sample concentration of ~20 μM. The red squares are the final values of Φ observed in the kinetics after complete relaxation. The continuous curves are global thermodynamic fits to the CD and fluorescence data that assume a two-state model (F $\rightleftharpoons$ U) with $K_{eq} = [U]/[F] = \exp[-\Delta H/R(1/T - 1/T_m)]$. We further assume linearly temperature-dependent CD baselines for U and F, and temperature dependence of the Φ given as: $\Phi_{F(U)} = 1/[1 + A_{F(U)} + B_{F(U)}\exp(-E_{F(U)}/RT)]$,[2] where $A_{F(U)}$, $B_{F(U)}$ and $E_{F(U)}$ are adjustable parameters. The optimized fluorescence baselines are shown for HP-35 Ala18Val in the inset. The resulting thermodynamic parameters are: for HP-35 Ala18Val $\Delta H = 24.3$ kcal/mol, $T_m = 335$ K; for HP-35 Ala18Ser $\Delta H = 21.6$ kcal/mol, $T_m = 337$ K. (b) Folding kinetics of HP-35 Ala18Val. Inset: fluorescence relaxation trace (red) following a 10 ns laser temperature jump from 323 K to 333 K. Continuous black curve is a bi-exponential fit to the data. From the relaxation rates of the slower kinetic phase (black circles) folding (blue circles) and unfolding (red circles) rates were obtained by the two-state analysis. The error bars were obtained by standard sensitivity analysis of χ^2.[47] Broken lines are fits to the folding/unfolding data using the Arrhenius relationship $k = k_0 \exp[-\Delta H^{\ddagger}/R(1/T - 1/T_0)]$, with $T_0 = 300$ K, which yield the following values: folding rate, $k_0 = (2.3 \pm 0.2) \times 10^5\ s^{-1}$, $\Delta H^{\ddagger} = -2.7(\pm 0.5)$ kcal/mol, unfolding rate, $k_0 = (3.8 \pm 0.4) \times 10^3\ s^{-1}$, $\Delta H^{\ddagger} = 21.0(\pm 0.6)$ kcal/mol. The same two-state analysis of the kinetic data for HP-35 Ala18Ser (not shown) yields: folding rate, $k_0 = (2.6 \pm 0.2) \times 10^5\ s^{-1}$, $\Delta H^{\ddagger} = -(2.7 \pm 0.5)$ kcal/mol, unfolding rate, $k_0 = (5.3 \pm 0.5) \times 10^3\ s^{-1}$, $\Delta H^{\ddagger} = 18.6(\pm 0.6)$ kcal/mol.

Figure 1(b) compares the structure of the doubly norleucine-substituted molecule (Lys24Nle/Lys29Nle), which we have solved by X-ray crystallography to 1 Å resolution, with the previously determined structure of the singly substituted molecule. Because both molecules crystallize in the same space group, a detailed comparison can be made, which shows essentially no difference (alpha carbon RMSD of 0.08 Å), and, in particular, no increased burial of the aliphatic side-chain of residue 29 for Lys24Nle/Lys29Nle. Nevertheless, the stability increases by an additional 0.5 kcal/mol compared to Lys24Nle (Figure 3). Two electrostatic effects may contribute

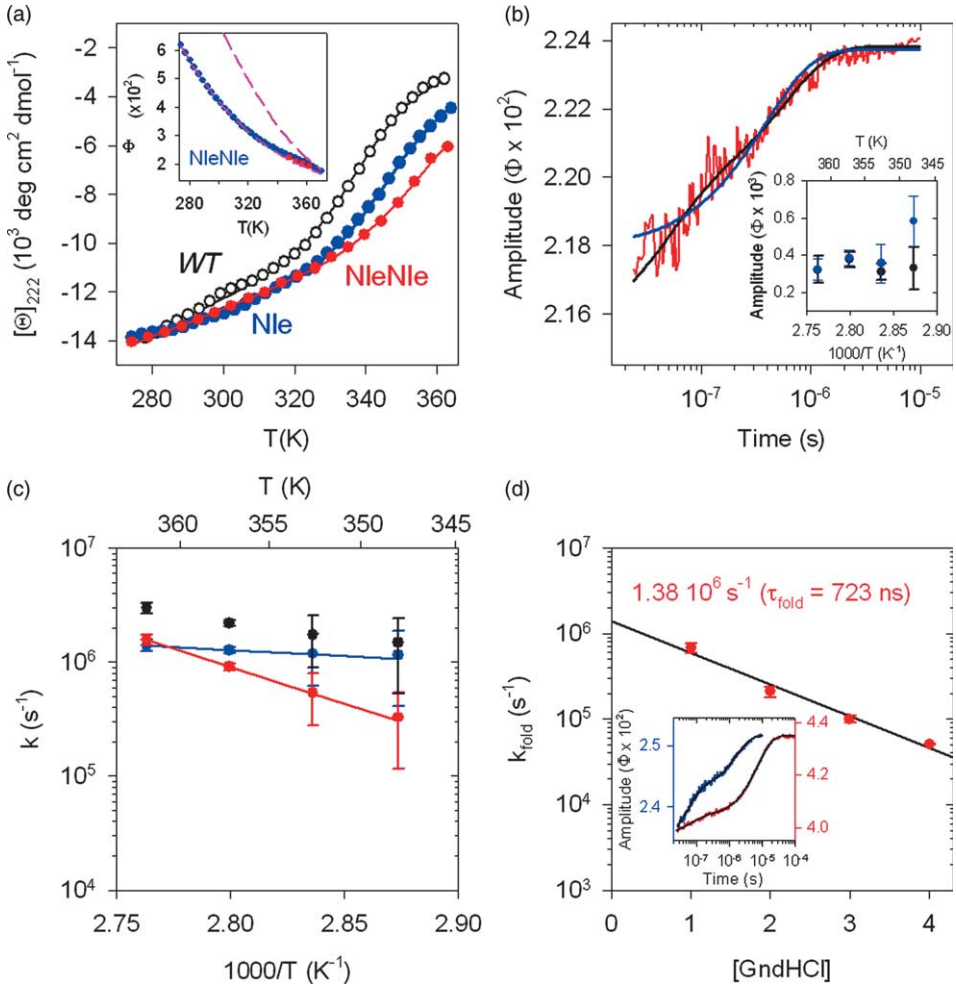

Figure 3. Folding thermodynamics and kinetics data for the HP-35 Lys24Nle/Lys29Nle. The data were collected under the same conditions as those shown in Figure 2. (a) CD thermal unfolding of HP-35 Lys24Nle/Lys29Nle (NleNle, red circles) compared to HP-35 Lys24Nle (Nle, blue circles) and the wild-type (WT, open circles). Inset: Temperature-dependent fluorescence quantum yield for Lys24Nle/Lys29Nle. As in Figure 2 the pink broken lines are the temperature-dependent fluorescence baselines and the red squares are the quantum yields from kinetic measurements after the complete relaxation. The two-state thermodynamic analysis (detailed in the legend to Figure 2) yields: $\Delta H =$ 25 kcal/mol, $T_m = 361$ K. (b) Relaxation of fluorescence quantum yield in response to the 10 ns laser temperature jump from 343 K to 348 K. The continuous blue curve is an exponential and the black curve a bi-exponential fit to the data. Inset: temperature dependence of kinetic amplitudes obtained from the single exponential fit (blue circles) and kinetic amplitudes of the slower relaxation from the bi-exponential fit (black circles). (c) Two-state analysis of the slower phase of the bi-exponential fits to the kinetic progress curves: observed relaxation rates (black circles) and folding (blue circles) and unfolding (red circles). Arrhenius fits $k = k_0 \exp[-\Delta H^{\ddagger}/R(1/T - 1/T_0)]$, with $T_0 = 361$ K, which yield the following values: folding rate, $k_0 = (1.4 \pm 0.1) \times 10^6\ s^{-1}$, $\Delta H^{\ddagger} = 4(\pm 14)$ kcal/mol, unfolding rate, $k_0 = (1.41 \pm 0.4) \times 10^6\ s^{-1}$, $\Delta H^{\ddagger} =$ 29($\pm$14) kcal/mol. (d) Folding rates of Lys24Nle/Lys29Nle as a function of denaturant GdmCl concentration at 300 K. Laser temperature jump experiments were carried out at four GdmCl concentrations, 1.0, 2.0, 3.0 and 4.0 M at pH 4.8 (20 mM acetate buffer). The exact concentration was in each case confirmed by refractometry. Relaxation kinetics at ten to twelve temperatures through the unfolding transition was measured for each denaturant concentration. The inset shows the kinetic traces at 1.0 M GdmCl (blue, temperature jump from 323 K to 333 K) and at 4.0 M GdmCl (red, temperature jump from 303 to 313 K). The folding rates at 300 K at each denaturant concentration were calculated from two-state analysis and Arrhenius fit to the temperature-dependent relaxation rates, in a fashion identical to the kinetic analysis presented in (c) and in Figure 2. Linear extrapolation to zero denaturant concentration yields for the folding rate at 300 K, $k_0 = (1.38 \pm 0.05) \times 10^6\ s^{-1}$.

to stabilization. One is removal of charged groups that are likely to be less solvated in the folded compared to unfolded structures. A second is decreased electrostatic repulsion in the folded structure. Examination of the distance matrix for the 11 charged residues of the wild type shows that the shortest distances are between Lys24 and protonated His27 (6.1 Å) and between Lys29 and Arg14 (5.3 Å). It would be interesting to investigate this issue further with a realistic electrostatic calculation.

The folding kinetics of Lys24Nle/Lys29Nle are shown in Figure 3. The additional stabilization shifts the unfolding temperature midpoint to near 90 °C. This high temperature creates problems for kinetic measurements. A fluorescence change in response to the rapid heating by a nanosecond laser pulse requires both a shift in the population from folded to unfolded and a difference in the quantum yields between the folded and unfolded states. As can be seen from Figure 3(a) (inset), the relative efficiency of tryptophan fluorescence quenching by protonated histidine (in the folded state) to that of quenching of free tryptophan in solution (the unfolded state), is much smaller at higher temperature. This is due to stronger temperature dependence of the free tryptophan quenching. As a consequence, there is a much smaller fluorescence change upon unfolding and therefore a much smaller kinetic signal amplitude than at lower temperature. At lower temperature (below about 70 °C), on the other hand, the protein stays essentially folded and no signal can be observed because there is almost no population shift between the folded and unfolded states. Lastly, the artifacts in the laser temperature jump experiments, such as cavitation, become much more severe when the temperature nears the boiling point.[41]

Nevertheless, we were able to measure the temperature jump relaxation unfolding/refolding kinetics in the temperature range 75–90 °C. We observe sub-microsecond relaxation kinetics, with the relaxation time at the unfolding temperature midpoint of 370 ns, which is more than five times faster than the corresponding relaxation time of wild-type HP-35, and about a factor of 2.5 faster than that of the molecule with the single norleucine substitution (Lys24Nle). The agreement between the equilibrium fluorescence and the final value of the quantum yield after the relaxation is complete (Figure 1(a) inset) confirms that there is no slower kinetic phase.

An important question is whether the relaxation kinetics is still bi-exponential as observed for the wild-type and all the mutants that we have investigated thus far. We have fitted the data with both single and double-exponential functions (Figure 3(b)). Unfortunately, it is not possible to clearly distinguish between the two scenarios due to the signal-to-noise challenge discussed above, further complicated by the comparable time-scales of the fast and the slower kinetic phases. However, single exponential fits to the kinetic data yield rather unusual temperature dependence of the kinetic amplitudes (Figure 3(b) inset), which appear to suddenly increase towards the lower temperature, but disappear completely at the next lower temperature point. On the other hand, the bi-exponential fit yields temperature dependence for the amplitude of the slower phase which is consistent with that expected for a two-state process, with a maximum near the transition midpoint. The temperature dependence of the kinetic amplitudes is therefore consistent with bi-exponential relaxation kinetics, further underlined by the fact that we obtain the same relaxation rate for the fast phase (~70 ns) as we have for each single mutant of HP-35 that we have measured thus far.

An alternative interpretation is that the free energy barrier to folding has disappeared and that the observed kinetics correspond to a single non-exponential process arising from reconfiguration in the folded well (the "one state"). The argument against this interpretation is that an analysis of the clearly bi-exponential kinetics of the wild type indicates that the fast phase corresponds to a conformational relaxation on the folded side of the main free energy barrier, and this relaxation rate is not significantly affected by the 12 mutations that we have studied so far (J.K., unpublished results). Consequently, we would expect a relaxation time for the one-state protein close to 70 ns, in contrast to the observed 350 ns relaxation time.

Our conclusion, then, is that there are still two processes and that folding can be modeled as a two-state process. The two-state analysis of the slower relaxation phase (Figure 3(c)) yields the folding rate at the transition midpoint ($T_m = 360$ K) of $(730(\pm 50)\ \text{ns})^{-1}$. The doubly norleucine-substituted HP-35 is therefore the first protein to fold in less than 1 μs (but only slightly faster than the 47 residue prb_{7-53} with two lysine residues replaced by hydrophobic residues, which folds in 1.0 μs[23]). The folding rate of Lys24Nle/Lys29Nle is only weakly temperature-dependent, which has been a general observation for virtually all HP-35 mutants. However, it is impossible to obtain the folding rate for Lys24Nle/Lys29Nle at 300 K with any reasonable level of confidence, since the Arrhenius fit to the folding rate for the available data points when extrapolated to 300 K is subject to huge error (this is obvious from the fitted value of the activation energy = 4(±14) kcal/mol).

To obtain the folding rate at 300 K and, at the same time, have an independent confirmation of the record breaking folding speed of HP-35, we measured the folding kinetics as a function of added denaturant, guanidinium chloride (GdmCl) (Figure 3(d)). The presence of GdmCl lowers the temperature denaturation midpoint, which greatly improves the signal-to-noise by increasing the change in fluorescence quantum yield. Linear extrapolation to 0 M GdmCl yields the folding rate of $(730(\pm 30)\ \text{ns})^{-1}$, the same value that we have obtained without denaturant at the transition midpoint temperature (360 K). The excellent

agreement between the values for both folding rates, extrapolated from denaturant at 300 K, and without denaturant but at higher temperature suggest that HP-35 Lys24Nle/Lys29Nle indeed folds in ~720 ns at 300 K.

Even with a sub-microsecond folding rate it appears that folding of HP-35 Lys24Nle/Lys29Nle still requires crossing a free energy barrier, albeit a very small one. A very rough estimate for the barrier height of 0.4 kcal/mol can be obtained by assuming that the speed limit of $[(N/100\ \mu s)^{-1} = (350\ ns)^{-1}]$ corresponds to the pre-exponential factor, τ_o (i.e. $\tau_f = \tau_o \exp(\Delta G_f^*/k_B T)$). An alternative estimate can be made by using Kramers theory and assuming that the 70 ns process corresponds to a conformational relaxation in the folded state. For a harmonic potential, $U(x) = 1/2\omega^2 x^2$, this relaxation time is given by:

$$\tau = \frac{k_B T}{\omega^2 D} = 70\ \text{ns} \quad (1)$$

where k_B is Boltman's constant, T is the absolute temperature, ω^2 is the curvature of the folded well, and D is the diffusion coefficient for motion in the well. The Kramers folding time for diffusion over a (high) one-dimensional barrier is given by:[42,43]

$$\tau_f = \frac{2\pi k_B T}{\omega N \omega^* N D^*} \exp\left(\frac{\Delta G_f^*}{k_B T}\right) = 720\ \text{ns} \quad (2)$$

where ωN^2 is the curvature of the unfolded well, ω^{*2} the curvature of the inverted barrier, D^* is the diffusion coefficient at the barrier top, and ΔG_f^* is the free energy barrier to folding. Assuming the same curvature for the folded and unfolded wells and the inverted barrier, and equal diffusion coefficients†, the free energy barrier height is simply:

$$\Delta G_f^* \approx k_B T \ln\left(\frac{t_f}{2\pi\tau}\right) = 0.3\ \text{kcal/mol} \quad (3)$$

Although there are many approximations in both estimates, they strongly suggest that the barrier to folding is very small, probably ~1 kcal/mol or less, and that additional stabilization could eliminate it to produce a "downhill" or one-state folder. At this point it will be interesting to use co-solvents that do not stabilize helices and to employ protein design algorithms to further stabilize the native state and possibly make the villin subdomain fold significantly faster. There is the interesting possibility that once the barrier disappears completely, additional stabilization, e.g. by co-solvents, could even slow folding by decreasing the effective diffusion coefficient.[34]

An important remaining question concerns the structure of the unfolded state. Is folding ultrafast because the helices are already preformed and simply dock together to form the folded structure or is the unfolded state more random-coil-like?[4,15,24,44] NMR results point to the latter. Havlin & Tycko have recently examined the alpha carbon and carbonyl carbon chemical shifts of eight carbon-13 labeled residues that sample all three helices and find no evidence for helix formation in either the thermally or chemically denatured state (R. H. Havlin & R. Tycko, unpublished results). There is also some evidence from tryptophan quenching experiments, at least in GdmCl, that the unfolded state is behaving more like a random coil. At 2 M GdmCl the rate of quenching of the tryptophan triplet by cysteine in the unfolded wild-type protein is $(3\ \mu s)^{-1}$.[38] This rate is very similar to the $(1\ \mu s)^{-1}$ rate recently measured under the same conditions for a completely disordered peptide $(W(AGQ)_nC)$ containing the same number of residues and a much higher proportion of glycine.[45] It is also similar to the rate of $(5\ \mu s)^{-1}$ measured for the quenching rate in the unfolded state of the cold shock protein of *Thermotoga maritima*, a five-stranded beta barrel protein (after normalizing for the different number of intervening residues assuming an $N^{-3/2}$ dependence). Nevertheless, there may be some native-like structure in the unfolded state of HP-35 in the absence of denaturant at 300 K[46] that could contribute to its ultrafast rate below the unfolding temperature.

Finally, we should point out that the sub-microsecond folding time, very accurate X-ray structure, and small size, together with the likelihood that the denatured protein is unstructured at the unfolding temperature, make this engineered villin subdomain an ideal system for simulation by atomistic molecular dynamics with explicit solvent.

† If the 70 ns relaxation occurs on the folded side of the folding/unfolding barrier, as mentioned above, any difference between D and D^* would be such that $D^* > D$, making the estimate of 0.3 kcal/mol a lower limit.

Acknowledgements

We thank Dr Wai-Ming Yau for synthesizing and purifying the peptides used in this work and Dr Attila Szabo for helpful discussion. We also thank Dr Zhongmin Jin at beamline 22-ID of the Advance Photon Source (US Department of Energy, Office of Basic Energy Sciences, contract W-31-109-Eng-38) for collecting X-ray diffraction data.

References

1. Snow, C. D., Nguyen, N., Pande, V. S. & Gruebele, M. (2002). Absolute comparison of simulated and experimental protein-folding dynamics. *Nature*, **420**, 102–106.
2. Kubelka, J., Eaton, W. A. & Hofrichter, J. (2003). Experimental tests of villin subdomain folding simulations. *J. Mol. Biol.* **329**, 625–630.

3. Snow, C. D., Qiu, L. L., Du, D. G., Gai, F., Hagen, S. J. & Pande, V. S. (2004). Trp zipper folding kinetics by molecular dynamics and temperature-jump spectroscopy. *Proc. Natl Acad. Sci. USA*, **101**, 4077–4082.
4. Mayor, U., Guydosh, N. R., Johnson, C. M., Grossmann, J. G., Sato, S., Jas, G. S. *et al*. (2003). The complete folding pathway of a protein from nanoseconds to microseconds. *Nature*, **421**, 863–867.
5. Ferguson, N., Day, R., Johnson, C. M., Allen, M. D., Daggett, V. & Fersht, A. R. (2005). Simulation and experiment at high temperatures: ultrafast folding of a thermophilic protein by nucleation–condensation. *J. Mol. Biol.* **347**, 855–870.
6. Snow, C. D., Sorin, E. J., Rhee, Y. M. & Pande, V. S. (2005). How well can simulation predict protein folding kinetics and thermodynamics? *Annu. Rev. Biophys. Biomol. Struct.* **34**, 43–69.
7. Jones, C. M., Henry, E. R., Hu, Y., Chan, C. K., Luck, S. D., Bhuyan, A. *et al*. (1993). Fast events in protein-folding initiated by nanosecond laser photolysis. *Proc. Natl Acad. Sci. USA*, **90**, 11860–11864.
8. Callender, R. H., Dyer, R. B., Gilmanshin, R. & Woodruff, W. H. (1998). Fast events in protein folding: the time evolution of primary processes. *Annu. Rev. Phys. Chem.* **49**, 173–202.
9. Gruebele, M. (1999). The fast protein folding problem. *Annu. Rev. Phys. Chem.* **50**, 485–516.
10. Eaton, W. A., Munoz, V., Hagen, S. J., Jas, G. S., Lapidus, L. J., Henry, E. R. & Hofrichter, J. (2000). Fast kinetics and mechanisms in protein folding. *Annu. Rev. Biophys. Biomol. Struct.* **29**, 327–359.
11. Myers, J. K. & Oas, T. G. (2002). Mechanisms of fast protein folding. *Annu. Rev. Biochem.* **71**, 783–815.
12. Hagen, S. J., Hofrichter, J., Szabo, A. & Eaton, W. A. (1996). Diffusion-limited contact formation in unfolded cytochrome *c*: estimating the maximum rate of protein folding. *Proc. Natl Acad. Sci. USA*, **93**, 11615–11617.
13. Kubelka, J., Hofrichter, J. & Eaton, W. A. (2004). The protein folding "speed limit". *Curr. Opin. Struct. Biol.* **14**, 76–88.
14. Spector, S. & Raleigh, D. P. (1999). Submillisecond folding of the peripheral subunit-binding domain. *J. Mol. Biol.* **293**, 763–768.
15. Mayor, U., Johnson, C. M., Daggett, V. & Fersht, A. R. (2000). Protein folding and unfolding in microseconds to nanoseconds by experiment and simulation. *Proc. Natl Acad. Sci. USA*, **97**, 13518–13522.
16. Ferguson, N., Johnson, C. M., Macias, M., Oschkinat, H. & Fersht, A. (2001). Ultrafast folding of WW domains without structured aromatic clusters in the denatured state. *Proc. Natl Acad. Sci. USA*, **98**, 13002–13007.
17. Qiu, L. L., Pabit, S. A., Roitberg, A. E. & Hagen, S. J. (2002). Smaller and faster: the 20-residue Trp-cage protein folds in 4 μs. *J. Am. Chem. Soc.* **124**, 12952–12953.
18. Yang, W. Y. & Gruebele, M. (2003). Folding at the speed limit. *Nature*, **423**, 193–197.
19. Wang, M. H., Tang, Y. F., Sato, S. S., Vugmeyster, L., McKnight, C. J. & Raleigh, D. P. (2003). Dynamic NMR line-shape analysis demonstrates that the villin headpiece subdomain folds on the microsecond time scale. *J. Am. Chem. Soc.* **125**, 6032–6033.
20. Nguyen, H., Jager, M., Moretto, A., Gruebele, M. & Kelly, J. W. (2003). Tuning the free-energy landscape of a WW domain by temperature, mutation, and truncation. *Proc. Natl Acad. Sci. USA*, **100**, 3948–3953.
21. Zhu, Y., Alonso, D. O. V., Maki, K., Huang, C. Y., Lahr, S. J., Daggett, V. *et al*. (2003). Ultrafast folding of α_3D: a *de novo* designed three-helix bundle protein. *Proc. Natl Acad. Sci. USA*, **100**, 15486–15491.
22. Qiu, L. L. & Hagen, S. J. (2004). Internal friction in the ultrafast folding of the tryptophan cage. *Chem. Phys.* **307**, 243–249.
23. Zhu, Y. J., Fu, X. R., Wang, T., Tamura, A., Takada, S., Savan, J. G. & Gai, F. (2004). Guiding the search for a protein's maximum rate of folding. *Chem. Phys.* **307**, 99–109.
24. Ferguson, N., Sharpe, T. D., Schartau, P. J., Sato, S., Allen, M. D., Johnson, C. M. *et al*. (2005). Ultra-fast barrier-limited folding in the peripheral subunit-binding domain family. *J. Mol. Biol.* **353**, 427–446.
25. Nguyen, H., Jager, M., Kelly, J. W. & Gruebele, M. (2005). Engineering beta-sheet protein toward the folding speed limit. *J. Phys. Chem. B*, **109**, 15182–15186.
26. Hagen, S. J., Qiu, L. L. & Pabit, S. A. (2005). Diffusional limits to the speed of protein folding: fact or friction? *J. Phys. Condens. Matter*, **17**, S1503–S1514.
27. Bryngelson, J. D., Onuchic, J. N., Socci, N. D. & Wolynes, P. G. (1995). Funnels, pathways, and the energy landscape of protein-folding—a synthesis. *Proteins: Struct. Funct. Genet.* **21**, 167–195.
28. Sabelko, J., Ervin, J. & Gruebele, M. (1999). Observation of strange kinetics in protein folding. *Proc. Natl Acad. Sci. USA*, **96**, 6031–6036.
29. Garcia-Mira, M. M., Sadqi, M., Fischer, N., Sanchez-Ruiz, J. M. & Munoz, V. (2002). Experimental identification of downhill protein folding. *Science*, **298**, 2191–2195.
30. Oliva, F. Y. & Munoz, V. (2004). A simple thermodynamic test to discriminate between two-state and downhill folding. *J. Am. Chem. Soc.* **126**, 8596–8597.
31. Yang, W. Y. & Gruebele, M. (2004). Folding lambda-repressor at its speed limit. *Biophys. J.* **87**, 596–608.
32. Naganathan, A. N., Perez-Jimenez, R., Sanchez-Ruiz, J. M. & Munoz, V. (2005). Robustness of downhill folding: guidelines for the analysis of equilibrium folding experiments on small proteins. *Biochemistry*, **44**, 7435–7449.
33. Ma, H. R. & Gruebele, M. (2005). Kinetics are probe-dependent during downhill folding of an engineered lambda(6–85) protein. *Proc. Natl Acad. Sci. USA*, **102**, 2283–2287.
34. Eaton, W. A. (1999). Searching for "downhill scenarios" in protein folding. *Proc. Natl Acad. Sci. USA*, **96**, 5897–5899.
35. Chiu, T. K., Kubelka, J., Herbst-Irmer, R., Eaton, W. A., Hofrichter, J. & Davies, D. R. (2005). High-resolution X-ray crystal structures of the villin headpiece subdomain, an ultrafast folding protein. *Proc. Natl Acad. Sci. USA*, **102**, 7517–7522.
36. Lapidus, L. J., Eaton, W. A. & Hofrichter, J. (2000). Measuring the rate of intramolecular contact formation in polypeptides. *Proc. Natl Acad. Sci. USA*, **97**, 7220–7225.
37. Buscaglia, M., Schuler, B., Lapidus, L. J., Eaton, W. A. & Hofrichter, J. (2003). Kinetics of intramolecular contact formation in a denatured protein. *J. Mol. Biol.* **332**, 9–12.
38. Buscaglia, M., Kubelka, J., Eaton, W. A. & Hofrichter, J. (2005). Determination of ultrafast protein folding rates from loop formation dynamics. *J. Mol. Biol.* **347**, 657–664.
39. Zagrovic, B., Snow, C. D., Shirts, M. R. & Pande, V. S. (2002). Simulation of folding of a small alpha-helical protein in atomistic detail using worldwide-distributed computing. *J. Mol. Biol.* **323**, 927–937.

40. Fernandez, A., Shen, M. Y., Colubri, A., Sosnick, T. R., Berry, R. S. & Freed, K. F. (2003). Large-scale context in protein folding: villin headpiece. *Biochemistry*, **42**, 664–671.
41. Wray, W. O., Aida, T. & Dyer, R. B. (2002). Photoacoustic cavitation and heat transfer effects in the laser-induced temperature jump in water. *Appl. Phys. B-Lasers Opt.* **74**, 57–66.
42. Kramers, H. A. (1940). Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, **VII**, 284–304.
43. Berne, B. J., Borkovec, M. & Straub, J. E. (1988). Classical and modern methods in reaction-rate theory. *J. Phys. Chem. B*, **92**, 3711–3725.
44. Myers, J. K. & Oas, T. G. (2001). Preorganized secondary structure as an important determinant of fast protein folding. *Nature Struct. Biol.* **8**, 552–558.
45. Buscaglia, M., Lapidus, L. J., Eaton, W. A. & Hofrichter, J. (2006). Effects of denaturants on the dynamics of loop formation in polypeptides. *Biophys. J.* In the press.
46. Tang, Y. F., Rigotti, D. J., Fairman, R. & Raleigh, D. P. (2004). Peptide models provide evidence for significant structure in the denatured state of a rapidly folding protein: the villin headpiece subdomain. *Biochemistry*, **43**, 3264–3272.
47. Bevington, P. R. & Robinson, D. K. (1992). *Data Reduction and Error Analysis for the Physical Sciences* (2nd edit.), WCB/McGraw-Hill, Boston.

Edited by C. R. Matthews

(Received 1 December 2005; received in revised form 7 March 2006; accepted 16 March 2006)
Available online 31 March 2006