Solvent-tuning the collapse and helix formation time scales of λ_{6-85}^*

CHARLES DUMONT,¹ YOSHITAKA MATSUMURA,² SEUNG JOONG KIM,¹ JINSONG LI,²
ELENA KONDRASHKINA,^{3,5} HIROSHI KIHARA,² AND MARTIN GRUEBELE^{1,4}

¹Department of Physics, University of Illinois at Urbana–Champaign, Urbana, Illinois 61801, USA

²Department of Physics, Kansai Medical University, 18-89 Uyama-Higashi, Hirakata 573-1136, Japan

³BioCAT at Advanced Photon Source, BCPS Department, Illinois Institute of Technology, Illinois 60439, USA

⁴Department of Chemistry, and Center for Biophysics and Computational Biology, University of Illinois at Urbana–Champaign, Urbana, Illinois 61801, USA

(RECEIVED March 30, 2006; FINAL REVISION June 20, 2006; ACCEPTED August 3, 2006)

Abstract

The λ_{6-85}^* pseudo-wild type of lambda repressor fragment is a fast two-state folder ($k_f \approx 35 \mu\text{sec}^{-1}$ at 58°C). Previously, highly stable λ_{6-85}^* mutants with $k_f > 30 \mu\text{sec}^{-1}$ have been engineered to fold nearly or fully downhill. Stabilization of the native state by solvent tuning might also tune λ_{6-85}^* away from two-state folding. We test this prediction by examining the folding thermodynamics and kinetics of λ_{6-85}^* in a stabilizing solvent, 45% by weight aqueous ethylene glycol at –28°C. Detection of kinetics by circular dichroism at 222 nm (sensitive to helix content) and small angle X-ray scattering (measuring the radius of gyration) shows that refolding from guanidine hydrochloride denatured conditions exhibits very different time scales for collapse and secondary structure formation: the two processes become decoupled. Collapse remains a low-barrier activated process, while the fastest of several secondary structure formation time scales approaches the downhill folding limit. Two-state folding of λ_{6-85}^* is not a robust process.

Keywords: collapse; heterogeneous kinetics; cryosolvent; stopped flow

In the conventional view of protein folding, the native state is reached after the unfolded protein has crossed one or more obligatory and substantial ($\gg kT$) barriers (Baldwin 1995). Small proteins and protein domains often encounter only one barrier (two-state folding), whereas for larger proteins or multi-domain proteins, population may pool temporarily in nonnative basins. In 1995, downhill folding was proposed as an alternative mechanism (Bryngelson et al. 1995): proteins could switch from two- or multi-state

folding to downhill folding when the native state becomes strongly stabilized (for example, when the temperature approaches the optimal folding temperature; Fig. 1). Kinetic signatures of this transition, such as switching from slow single exponential kinetics to fast nonexponential dynamics (Sabelko et al. 1999; Yang and Gruebele 2003), and probe-dependent kinetics (Yang and Gruebele 2004b; Ma and Gruebele 2005) have been observed experimentally.

The possibility of downhill folding opens up an important question: how robust is two-state (or multi-state) folding really? If a wild-type protein subject to a few mutations or changes in solvent conditions, tuning its free energy landscape by a few kT , can be made to fold downhill, and if small proteins can fold in a few microseconds (Kubelka et al. 2004), folding barriers are not a necessary prerequisite for folding. It has even been suggested that downhill folding could extend beyond the apparent melting temperature (“one-state folding”), leaving

⁵Present address: Rigaku Innovative Technologies, Inc., 1900 Taylor Rd., Auburn Hills, MI 48326, USA.

Reprint requests to: Martin Gruebele, University of Illinois at Urbana–Champaign, 600 South Mathews Ave., Box 5-6, Urbana, IL 61801, USA; e-mail: gruebele@scs.uiuc.edu; fax: (217) 244-3186; or Hiroshi Kihara, Kansai Medical University, 18-89 Uyama-Higashi, Hirakata 573-1136, Japan; e-mail: kihara@makino.kmu.ac.jp; fax: 001-81-72-850-0733.

Article and publication are at <http://www.proteinscience.org/cgi/doi/10.1110/ps.062257406>.

Figure 1. (Top) Tuning two-state or multi-state scenarios toward downhill barrier or downhill scenarios by stabilizing the native state. On the left side, structure is formed uphill over the main barrier $\Delta G_2^\ddagger \gg kT$. If $\Delta G_1^\ddagger$ is also large enough, an intermediate can accumulate between the two barriers. (Intermediates could also occur after the main barrier.) On the right side, native bias has reduced $\Delta G_1^\ddagger$ such that substantial structure forms downhill, while a barrier crossing may remain over $\Delta G_2^\ddagger$. If the bias is strong enough such that $\Delta G_2^\ddagger < 3 kT$, then the entire folding reaction approaches downhill folding. (Bottom) Location of Trp 22 (blue), Y/Q33 (green), and A/G37,49 (red) in λ_{6-85}^* and λ_{6-85}^* triple mutant. The illustration is obtained from the PDB lambda repressor structure after amino acid substitution and molecular dynamics relaxation of the side chains using VMD software (Humphrey et al. 1996). Previous work has shown that helix 5 is less structured than shown here (Larios et al. 2006).

a thermodynamic signature (Muñoz 2002). Experimental evidence for this case has been reported and debated (Garcia-Mira et al. 2002; Ferguson et al. 2004; Naganathan et al. 2005).

It has been known for some time that small proteins are not generally optimized for the fastest possible folding (Kim et al. 1998). In one proposal, downhill folding is not common among natural proteins because evolution for function and against aggregation usually disfavors full kinetic optimization (Gruebele 2005). The WW domain has been studied recently as one example of a structure-folding-function tradeoff (Jäger et al. 2006). However, cases where downhill folding would be favored by evolution for certain types of functions have also been discussed (Garcia-Mira et al. 2002; Muñoz 2006). Even if most small wild-type proteins fold via an activated two-state process, the existence of downhill folding points out that rich energy landscape is just barely hidden by one or a few non-obligatory bottlenecks. The folding rates of many small proteins or protein domains are not far from the downhill “speed limit,” and some have been engineered and tuned near or to that limit (Qiu et al. 2002; Kubelka et al. 2003, 2004; Yang and Gruebele 2003; Yang et al. 2003; Zhu et al. 2003; Arora et al. 2004; Faraone-Mennella et al. 2005; Nguyen et al. 2005). To the extent that native topology controls folding rates, there is good reason to believe that additional slower folding proteins can be engineered or placed in solvent conditions where their folding is no longer described by activated rate theory.

In the past, fast phases have been interpreted as non-specific contraction of the protein preceding the critical barrier crossing (Krantz and Sosnick 2000), and thermodynamic titration baselines are usually subtracted from data without detailed justification to enable two-state fitting. Strictly speaking, true two-state folders must have single-exponential kinetics and no baselines. If the deviations from two-state behavior are small, it is reasonable to approximate folding as an activated process between two slightly moving targets (the unfolded and native basins), and this may well apply to many proteins. However, when baselines or new kinetic phases become comparable in magnitude to the apparent two-state signals, this view is simply not tenable (Muñoz 2002; Qin et al. 2002; Larios et al. 2004). Intermediates must be invoked if the new kinetics are slow enough to correspond to barriers $\gg kT$ or titrations show steps corresponding to distinct states. Downhill folding must be invoked if the kinetics are near the speed limit, nonexponential and probe-dependent, or if no evidence of separate steps is discernible in thermodynamic measurements with large baselines.

In this paper, we solvent-tune the five helix bundle λ_{6-85}^* (Fig. 1) away from two-state folding. In aqueous solvent, λ_{6-85}^* is a two-state folder with a minimal folding time of 35 μsec at 58°C (Yang and Gruebele 2004c). The pseudo-wild type λ_{6-85}^* (itself a fluorescent Y22W mutant of the λ_{6-85}^* wild type) (Ghaemmaghani et al. 1998) was previously engineered toward downhill folding by site-directed mutagenesis (Yang and Gruebele 2004a). Here, we ask how far the pseudo-wild type can be tuned by a stabilizing solvent alone. We monitor kinetics by using two spectroscopic probes correlated with two very different reaction coordinates (Arai et al. 1998). Compactness is probed by small-angle X-ray scattering (Arai et al. 1998; Segel et al. 1999), yielding the radius of gyration R_g . Secondary structure is probed by circular dichroism (CD) monitoring helical content at 222 nm (Kuwaitima et al. 1996). We study folding in a cryogenic solvent (45% ethylene glycol in water) at low temperature (−28°C). The solvent not only stabilizes secondary structure, but through its high viscosity substantially slows down the dynamics of the system by affecting the prefactor, not just the activation energy.

We observe up to four kinetic phases for λ_{6-85}^* ; protein collapse and various stages of secondary structure formation no longer coincide under our solvent conditions. Our lower limit on the rate coefficient of one of the CD-detected phases is sufficiently fast to correspond to downhill secondary structure formation. We make a mutant with lower secondary structure propensity than λ_{6-85}^* , and show that its fast CD phases decrease in magnitude and slow down. Thus, the fast kinetic phases are correlated with sequence and native secondary structure

formation propensity. At the same time, no contraction of the chain occurs. λ_{6-85}^* does not fold downhill under our solvent conditions, but its helix formation and collapse processes have become completely decoupled, and the speed of much secondary structure formation approaches the downhill limit. Two-state folding is not a robust process for λ_{6-85}^* .

Results

Steady-state SAXS and CD

The radius of gyration (R_g) obtained by SAXS for native λ_{6-85}^* in 0.7 M guanidine hydrochloride buffer is 14 ± 1 Å. R_g is 23 ± 2 Å in the unfolded state (Fig. 2). The native radius of gyration of λ_{6-85}^* estimated from the 6–85 fragment of the lambda repressor X-ray crystal structure using the program CRY SOL is 14 Å. CRY SOL already includes an estimate of the solvation shell effect on R_g (11 Å bare, 3 Å solvation shell) (Svergun et al. 1995). The slightly larger final R_g compared to the bare radius could also result from helix 5 of λ_{6-85}^* not being highly structured compared to the X-ray crystal structure of the whole protein (Larios et al. 2006).

The random coil radius for λ_{6-85}^* is expected to be 26 Å, based on the power law scaling determined by Kohn et al. (2004). The observed value of 23 Å is slightly more compact, indicating some residual interactions in 5 M guanidine hydrochloride at -28°C . Some uncertainty in R_g is also introduced by the nonlinearity of the Guinier plot at small values of Q^2 , resulting from transient aggregation, as discussed in Materials and Methods.

Figure 2. Guinier plots of folded (top, 0.7 M guanidine hydrochloride in 45% ethylene glycol at $-28 \pm 1^\circ\text{C}$) and unfolded (bottom, 5 M guanidine hydrochloride) λ_{6-85}^* . The double-exponential fits to Equation 4 are shown in red; the single exponential fits are shown in blue. Both types of fits yielded radii of gyration in agreement within 1 Å.

Figure 3. (Top) Circular dichroism spectra of λ_{6-85}^* at $-28 \pm 1^\circ\text{C}$ concentrations (from thin to thick line): 0, 0.95, 1.36, 1.86, 2.47, 2.91, 3.39, 3.90, 4.41, 4.84, 5.30 M. The insert shows spectra reconstructed from two SVD functions. (Bottom) Guanidine hydrochloride melt of λ_{6-85}^* detected by circular dichroism at $-28 \pm 1^\circ\text{C}$ in 45% ethylene glycol 50 mM phosphate buffer.

Similar SAXS results were obtained for the Q33Y/A37G/A49G mutant of λ_{6-85}^* .

CD spectra of λ_{6-85}^* as a function of guanidine hydrochloride concentration are shown in Figure 3. In the absence of denaturant, the protein shows a typical α -helical signature with peaks at 208 and 222 nm. The native steady-state value is for $[\theta]_{222}$ is $-14,000 \pm 300$ deg $\text{m}^{-1} \text{M}^{-1}$. In 5.3 M denaturant, the CD spectrum is slightly positive, indicative of a random coil secondary structure. Similar results were obtained for the mutant. Singular value decomposition of the λ_{6-85}^* spectra in the 208–250 nm region reveals a simple sigmoidal transition with the two largest singular values (1, 0.035) (Materials and Methods; Henry and Hofrichter 1992). The inset in Figure 3 shows how the two largest components reconstruct the CD spectra at different denaturant concentrations.

Thermodynamic titration

The unfolding titration of λ_{6-85}^* detected by circular dichroism at 222 nm is shown in Figure 3. The midpoint lies at 3.02 M guanidine hydrochloride concentration. Table 1 summarizes the thermodynamic parameters obtained for the two proteins studied here, as well as

Table 1. Thermodynamic parameters of λ_{6-85} Y22W (λ_{6-85}^*) and λ_{6-85} Y22W/Q33Y/A37G/A49G

Protein	−28.5°C, 45% ethylene glycol		22°C, 0% ethylene glycol	
	C_m (molar)	M (kJ/mol)	C_m (molar)	M (kJ/mol)
λ_{6-85}^*	3.02	7.7	2.04	9.0
Mutant	3.04	7.9	1.61	7.8

the parameters from Yang and Gruebele (2004c) at 22°C in aqueous phosphate buffer. The fit in Figure 3 is to the two-state linear free energy relationship (Tanford 1968)

$$\Delta G = m([GuHCl] - C_m). \quad (1)$$

The midpoint of the unfolding transition in the crysolvent lies nearly 1 M higher than under conditions previously used to investigate folding of λ_{6-85}^* mutants (Yang and Gruebele 2004c). The mutant is stabilized slightly more. As discussed below, the parameters in Equation 1 should not be taken to confirm two-state behavior far below the denaturation midpoint.

Refolding kinetics

Figure 4 summarizes the SAXS and circular dichroism kinetics obtained for the pseudo-wild type. The radius of gyration smoothly decays from 24 ± 1 Å to 15 ± 1 Å. The observed SAXS kinetics can be fitted within the signal-to-noise ratio to a single exponential decay with an observed rate coefficient of $k_{\text{obs}} = 3.0 \pm 1 \text{ sec}^{-1}$. The mutant fits similarly with $1.8 \pm 1 \text{ sec}^{-1}$. Kinetic fits to the λ_{6-85}^* and mutant data are summarized in Table 2.

The CD-monitored refolding kinetics of λ_{6-85}^* differ from the SAXS data in three important aspects. First, a burst phase of $\tau_{1/e} < 2$ msec overshoots to 150% of the native value of $[\theta]_{222} = -13,500 \pm 1000 \text{ deg m}^{-1} \text{ M}^{-1}$. The 2-msec limit was set based on the signal-to-noise ratio, size of the burst phase, and instrument response function of 5.3 msec (see Materials and Methods). The burst phase rate is thus over 150 times faster than the SAXS rate. Second, the fastest resolvable recovery phase fits to a single exponential with $k_{\text{obs}} = 35 \text{ sec}^{-1}$. This rate is still 10 times faster than the collapse rate determined by SAXS, and thus also not concomitant with the collapse kinetics. Finally, the resolvable CD phase of the pseudo-wild type is somewhat better fitted by a double or stretched exponential than by a single exponential (Fig. 4, bottom panel). All of these results were reproduced over several measurements. The CD kinetic parameters are summarized in Table 2 and Figure 4. The arrows in the figure indicate the values measured for native and

unfolded conditions under steady state, closely matching the kinetic endpoints.

The SAXS-detected kinetics of λ_{6-85}^* and its mutant are very similar. In contrast, the CD-detected kinetics of the Q33Y/A37G/A49G mutant of λ_{6-85}^* with lower helix propensity differ significantly from the pseudo-wild type (see inset in Fig. 4). The burst phase now takes up only 78% of the amplitude. The rate of the recovery still differs from the SAXS measurement, but now by less than a factor of two. Last, the recovery phase fit is not improved by a double or stretched exponential fitting function.

Discussion

If site-directed mutagenesis can disrupt two-state folding of λ_{6-85}^* and even induce downhill folding (Yang and Gruebele 2004a), how robust is two-state folding against changes in solvent conditions only? Our answer, based on the results outlined above, is that a $<0.2 \text{ kT}$ change of

Figure 4. Refolding kinetics of λ_{6-85}^* upon dilution from 5 M to 0.7 M guanidine hydrochloride by stopped flow. (Top) Radius of gyration detected by SAXS; the inset shows the mutant data. (Middle) $[\theta]_{222}$ detected by circular dichroism; the inset shows the traces for the mutant. (Bottom) Detailed view of the λ_{6-85}^* exponential and stretched exponential fits.

Table 2. Kinetic parameters of λ_{6-85} Y22W (λ_{6-85}^*) and λ_{6-85} Y22W/Q33Y/A37G/A49G at -28.5°C in 45% ethylene glycol: SAXS rate, burst phase amplitude, single exponential CD fit, stretched exponential CD fit, and double exponential CD fit

Protein	k_{obs} , sec^{-1} (SAXS)	A_{burst} , % (CD)	k_{obs} , sec^{-1} (CD)	k_{stretch} , sec^{-1} (CD)	β (CD)	$k_{\text{bi},1}$, sec^{-1}	$k_{\text{bi},2}$, sec^{-1}
λ_{6-85}^*	3.0 ± 1	150%	35 ± 4	41 ± 6	0.50 ± 0.04	66 ± 16	4.9 ± 3.0
Mutant	1.8 ± 1	78%	4.9 ± 0.5	—	—	—	—

The large unresolved burst amplitude has $k \geq 500 \text{ sec}^{-1}$. Uncertainties are two standard deviations.

thermal energy and switching to a more secondary structure-stabilizing solvent are sufficient to disrupt two-state folding, decoupling the collapse and secondary-structure formation time scales of the pseudo-wild type λ_{6-85}^* . The main activation barrier at 25°C in aqueous buffer is not highly dominant over other free energy landscape roughness, which contributes to the heterogeneous kinetics described above. So, it comes as less of a surprise that signatures of downhill folding have been detected in mutants of λ_{6-85}^* (Yang and Honig 1993; Yang and Gruebele 2003, 2004a; Ma and Gruebele 2005).

The ethylene glycol/water solvent used here slows down the prefactor for folding such that collapse and some of the secondary structure formation can be resolved on the millisecond-to-second time scale, and additionally suppresses the cold denaturation ordinarily observed for lambda repressor fragment (Yang and Gruebele 2005). Our goal was to take λ_{6-85}^* and tune its stability to a value comparable to the most stable lambda repressor fragment mutants, such as Y22W/Q33Y/G46A/G48A (Yang and Gruebele 2004c). The latter has a guanidine midpoint concentration of $C_m = 3.27 \text{ M}$ compared to 2.04 M for λ_{6-85}^* at 22°C in aqueous buffer. As shown in Table 1, 45% ethylene glycol at -28°C shifted the stability of λ_{6-85}^* up to a 3.02 M midpoint. The Y22W/Q33Y/A37G/A49G mutant is similarly shifted to $C_m = 3.04 \text{ M}$. These values nearly approach those of the most stable mutants.

Our most important result is that the SAXS-detected refolding rate, which probes the radius of gyration and is thus sensitive to protein collapse, and the CD-detected refolding rates, which probe secondary structure formation, differ by factors of 10–150 for λ_{6-85}^* . Collapse occurs on a 300 msec time scale, as measured by the SAXS radius of gyration (Table 1). A CD burst phase of 150% of the native $[\theta]_{222}$ value whose tail cannot be resolved with our instrument response function requires $k_{\text{burst}} \geq 2 \text{ msec}^{-1}$, >150 times larger than the collapse rate. The subsequent resolved relaxation of the CD overshoot back to the native baseline is still over 10 times faster than the collapse rate of λ_{6-85}^* . Thus, neither the initial secondary structure overshoot, nor its subsequent relaxation back to the native value, coincide with collapse to a compact state. There are at least three kinetic phases. Unstructured contraction in the unfolded well preceding two-state folding could explain one small fast phase in

addition to the rate-limiting step, but not three very different fast phases, one of which exceeds the native signal level by 50%. Under our solvent conditions, λ_{6-85}^* , native-like helical content precedes collapse substantially.

Our second important finding is that the fast CD phases are sequence-specific. If substantial helical structure forms during the fast CD phases, then decreasing the helical propensity of the protein should decrease the magnitude and/or the rate associated with the fast CD phases. To test this hypothesis, we made the Y22W/Q33Y/A37G/A49G mutant. The Q33Y and A37G/A49G mutations have opposite effects on protein stability, producing a mutant just about as stable as the pseudo-wild type (Yang and Gruebele 2004c; Larios et al. 2006) (Table 2). The effect of the mutations on secondary-structure stability is very different. AGADIR analysis (Lacroix et al. 1998) indicates that helix 2 propensity is not seriously perturbed by the Q33Y substitution, while helices 2 and 3 propensities are significantly reduced by the A37G/A49G pair of mutations. The combination of the Q33Y and A37G/A49G mutations thus allows us to reduce helix propensity without changing the overall stability of the protein significantly. The kinetics results are clear: The SAXS rate remains similar to λ_{6-85}^* , so the glycine substitutions have no large effect on the collapse rate. The CD-detected kinetics change substantially. The mutant still has a burst phase, but its amplitude (78%) is only half that of λ_{6-85}^* . In addition, the resolvable recovery phase now is seven times slower than in λ_{6-85}^* .

A nonspecific contraction does not occur in λ_{6-85}^* before it folds to the native state. Rather, secondary structure forms rapidly on at least two time scales, followed by slower collapse. Both rapid collapse and collapse concomitant with the rate-limiting step of folding have been reported (Plaxco et al. 1999; Sadqi et al. 2003). The CD overshoot most likely corresponds to excess helix formation, followed by reduction in helix content and formation of loops needed to acquire native topology: λ_{6-85}^* has a normal helix bundle CD spectrum at all denaturant concentrations studied, with a 208/222 nm double minimum even above the denaturation midpoint (Fig. 3). The CD kinetics at 222 nm thus monitor helix content. Extended structures of proteins with high helical content have been observed previously under certain denaturing conditions, such as the acid denatured A state

of ubiquitin (Brutscher et al. 1997). The ensemble occupied by λ_{6-85}^* within 2 msec of folding initiation at -28°C in 45% ethylene glycol is probably similarly disposed.

Our third finding is not as strong, but also supports the view that our solvent conditions reveal heterogeneity of the λ_{6-85}^* free energy landscape not observed at room temperature. The bottom panel in Figure 4 shows the fits of the CD overshoot relaxing back to the native value. A stretched (or double, data not shown) exponential fit accounts for the data somewhat better than a single exponential fit, when both are constrained to match the observed amplitude and baseline. The figure shows the raw data in 3-msec increments, so there is ample sampling on the scale of the fitted rate coefficients shown in Table 2. The χ^2 of the stretched fit is 1.5 times better because it accounts for the region from 0.02–0.15 sec more accurately. Thus, secondary-structure formation of λ_{6-85}^* may exhibit three or more phases, and the overall folding process four or more phases, but certainly at least the three (<2 msec, ≈ 30 msec, ≈ 300 msec) that are clearly separated.

There is no strong evidence for deviations from two-state behavior in the thermodynamic data (Fig. 3). Just like the room temperature thermal titration data (Yang and Gruebele 2004c), denaturant titration under cryogenic conditions is well-approximated by a two-state transition. As discussed elsewhere (Yang and Gruebele 2004a,c), this does not imply two-state folding of λ_{6-85}^* in the absence of denaturant. It simply means that the transition region along the reaction coordinate is sufficiently destabilized by denaturant (or high temperature) that the protein acquires a dominant folding barrier and makes a transition toward two-state folding. Similarly, it has been shown that downhill folding variants of λ_{6-85}^* can be turned into two-state folders by stressing the protein with high temperature (Ma and Gruebele 2005). For other small proteins, such as BBL, there is evidence that downhill relaxation can survive heat stress (Garcia-Mira et al. 2002; Muñoz 2002; Muñoz and Sanchez-Ruiz 2004; Naganathan et al. 2005).

Finally, we compare the measured rate coefficients in cryogenic solvent and at room temperature. The cryogenic solvent conditions affect the folding barrier(s), as well as the solvent viscosity. Increased viscosity slows down the diffusional dynamics of the protein during its conformational search and reduces the prefactor for folding in activated rate models in the Kramers overdamped regime (Kramers 1940).

We previously showed that folding rate coefficients of known two-state folding mutants of λ_{6-85}^* can be fitted to a simple quadratic free energy model over a wide temperature range, including cold denaturation temperatures down to -4.7°C (Yang and Gruebele 2004c, 2005). We

thus use as our zeroth-order model for two-state folding the Kramers expression (Kramers 1940)

$$k_f(T) = \nu^\dagger(T_m) \left(\frac{\eta(T_m, \chi)}{\eta(T, \chi)} \right)^a e^{-[\Delta G_0^\dagger + \Delta G_1^\dagger(T-T_m) + \Delta G_2^\dagger(T-T_m)^2]/RT} \quad (2)$$

where $\nu^\dagger$ (61°C) ≈ 1.5 μsec (Yang and Gruebele 2003), and $a \approx 1$. The viscosity is discussed in more detail in the Materials and Methods section. Figure 5 plots the refolding rates of λ_{6-85}^* from room temperature fluorescence measurements, as well as Equation 2 over its expected range of validity. The fastest CD phase reported here is not likely to be reached by such an extrapolation, even if higher order corrections were added to the free energy. The SAXS rate reported here is more likely to be reached by such an extrapolation. Using a viscosity-dependent prefactor of $\nu^\dagger = \nu^\dagger(T_m)\eta(T_m, \chi=0)/\eta(T, \chi=0.19) \approx 72$ μsec , we obtain an activation energy of 16.8 kJ/mol for the collapse of λ_{6-85}^* at -28°C in 45% ethylene glycol. The upper limit set on the barrier for excess secondary structure formation (the CD burst phase) is even more stringent, $\Delta G^\dagger < 6.4$ kJ/mol. This lies near the 3 kT limit where transition theory begins to break down and downhill folding takes over (Yang and Gruebele 2003). Thus, the cryogenic conditions provide substantial stabilization of the transition state region of the free energy surface, not just of the native well.

We therefore conclude that cryogenic ethylene glycol solvent speeds up both collapse and secondary structure formation relative to models that adequately fit previous room temperature and cold denaturation data obtained in aqueous solvents. Helix formation and collapse of λ_{6-85}^* are decoupled into at least three, and possibly more,

Figure 5. Refolding rates observed for λ_{6-85}^* by SAXS and CD (black circles) in this work, as well as fluorescence-detected rates measured in aqueous buffer from 58°C – 85°C . The fit of the room temperature data to Equation 2 is shown over the expected range of validity. The extrapolated folding rate of the fast Q33Y,G46A,G48A mutant of λ_{6-85}^* at 37°C is shown as an open square.

distinguishable time scales. Excess helical structure is formed very rapidly in λ_{6-85}^* , followed by a reduction in helix content toward the native value, followed finally by collapse of the protein and arrangement of the native tertiary structure. The cartoon in Figure 1 shows this sequence of events mapped onto a single global reaction coordinate, but two coordinates (e.g., fractional helix content and radius of gyration) may be more suitable for a complete description of the process. A mutant with reduced secondary structure propensity but comparable stability does not accumulate excess secondary structure and relaxes to the native secondary structure more slowly, so the fast processes are sequence-specific.

Although the solvent conditions at -28°C in 45% ethylene glycol may seem exotic, we must remember that they do not induce a large perturbation of the free energy landscape. λ_{6-85}^* is stabilized over “physiological” conditions by only about 1 M guanidine hydrochloride, corresponding to an average native bias of $\Delta\Delta G \approx 7.7$ kJ/mol using the *m*-value in Table 1. For such a small change to disrupt two-state folding indicates that two-state folding is not a very robust process in λ_{6-85}^* to begin with. Even without genetic engineering, small protein-free energy barriers can be tuned sufficiently so free energy landscape roughness can contribute to the observed folding kinetics.

Materials and methods

Protein

The wild type of λ repressor is a DNA-binding phage regulatory protein, which controls the lambda switch in bacterial cells. The engineered protein λ_{6-85} , containing 80 residues and five α -helices (MW ~ 9160 Da) of the full protein, was the first reported submillisecond two-state folder (Huang and Oas 1995). A Tyr22Trp mutant of the wild type (abbreviated here as λ_{6-85}^*), made to enable fluorescence detection, does not significantly affect either the thermodynamic stability or folding kinetics of λ_{6-85}^* (Ghaemmaghami et al. 1998). We made Y22W and Y22W/Q33Y/A37G/A49G mutants by site-directed mutagenesis (Stratagene Quickchange kit), starting with a wild-type plasmid donated by Terry Oas. Genes were inserted into the PET-15b vector with a histidine tag coding region and expressed in RosettaTM(DE3) pLysS cells (Novagen Inc.). After lysing cells on a French press and collecting supernatant by centrifugation, the overexpressed protein was purified through Ni-ATA and Sephacryl S-200 HR columns (Pharmacia). The histidine tag was removed by thrombin cleavage. The purity of λ_{6-85}^* and its mutant was confirmed by matrix-assisted laser desorption ionization mass spectrometry and sodium dodecyl sulfate polyacrylamide gel electrophoresis.

Cryosolvent conditions

The lyophilized proteins were resuspended in a 45%/55% ethylene glycol/deionized water mixture (Qin et al. 2001). The

mixed solvent was buffered to pH 7.0 with 50 mM phosphate buffer. As initial condition for stopped flow studies, 5 M guanidine hydrochloride (relative to total solvent volume) was added. The concentration of GuHCl was calibrated by refractive index measurements (Nozaki 1970; Pace 1986). Other guanidine hydrochloride buffer concentrations were used for unfolding titrations to determine thermodynamic stability. Low-temperature measurements were carried out at $-28 \pm 1^\circ\text{C}$.

Stopped-flow setup

Nearly identical stopped-flow mixers were used in Japan (CD) and in the United States (SAXS). The cryogenic stopped-flow mixers are custom instruments built to operate at high viscosity and low temperature by UNISOKU. The instruments allow the fast mixing of two solutions in the ratio 1:6 without significant leakage at temperatures as low as -30°C .

For all protein measurements, one part of denatured protein solution was mixed with six parts of either the denaturant buffer (5 M GuHCl, for calibration) or the refolding buffer (0 M GuHCl, yielding 0.7 M final guanidine hydrochloride concentration). The cell containing the mixed solution has a 1-mm path length and 50- μm sapphire windows, suitable for both CD and SAXS steady-state and kinetic measurements. For cryogenic X-ray measurements, the cell was insulated with a cork/foam envelope and cooled by an ethylene glycol/water mix with a Neslab ULT-80DD model bath (rated 250 W at -70°C).

The dead time for the stopped flow was calibrated to be 5.3 msec in 45% ethylene glycol at -28°C , using the fast reduction of 2,6-dichlorophenol sodium by ascorbic acid at pH 4. The rate of this bimolecular reaction increases linearly with ascorbic acid concentration. As the rate increases, the fraction of resolvable signal change decreases exponentially, allowing us to determine the concentration of ascorbic acid where $k_{\text{obs}} = k_{\text{dead}}$ (we take it as the 1/e point of the signal decrease). Inserting this concentration (24.75 mM) into the linear dependence of k_{obs} on ascorbic acid concentration yields the dead time, as shown in Figure 6.

Small-angle X-ray scattering

SAXS steady-state and refolding data of λ_{6-85}^* were collected at the BioCAT 18 undulator beamline of the Advanced Photon Source at Argonne National Laboratory (Fischetti et al. 2004). The X-ray beam (nominal wavelength 1.0 Å) was collimated to

Figure 6. Observed rate of the dead time calibration reaction as a linear function of ascorbic acid concentration. The observable signal change decreased by 1/e when the concentration was raised to 24.75 mM; inserting this value into the fitted relationship between rate and concentration yields the “dead-rate,” or its inverse, the dead time of 5.3 msec.

a spot size of 300×130 microns at the sample. The scattering signal was detected with a MAR 165 CCD X-ray detector with 2048×2048 pixels of $79\text{-}\mu\text{m}$ pixel size. The detector was placed 1.90 m away from the sample cell. The useful range of $Q = 4\pi \sin\theta/\lambda$ for the detector was $0.006\text{--}0.23 \text{ \AA}^{-1}$. Final protein concentration in SAXS experiments was $\approx 450 \text{ }\mu\text{M}$.

X-ray data were collected for exposure times between 100 msec (at $t = 0$) and 600 msec ($t \geq 1 \text{ sec}$). The scattering patterns were azimuthally averaged, with logarithmic weighting in Q . Scattering regions with contributions from crystalline structure of the sample cell windows were masked out. A 0.7-M guanidine hydrochloride buffer reference processed equally is subtracted from the sample intensity to yield a final Guinier plot.

λ_{6-85}^* mutants are known to undergo transient aggregation in aqueous buffer at similar concentrations ($100\text{--}400 \text{ }\mu\text{M}$) to those used here, causing up to a 10% shift in observed relaxation times (Yang and Gruebele 2003). We did observe curvature of the Guinier plot below $Q^2 = 0.001$, indicative of transient aggregation (Fig. 2). Several types of data fitting were therefore tested to make sure that the extracted radii of gyration R_g are robust. We used a generalized Guinier formula (Guinier and Fournet 1955)

$$I(Q) = A \exp(-R_g^2 Q^2/3) + B \exp(-R_{agg}^2 Q^2/3). \quad (3)$$

Here, R_{egg} is an effective radius of gyration for the ensemble of aggregate states, which contribute at small values of Q .

We performed fits over a range of Q^2 between 0.0005 (below which baseline subtraction was not accurate) and $Q^2 = 0.004$ (above which the Guinier constraint $R_g \cdot Q_{\max} \leq 1.3$ is no longer satisfied), using one or two terms in Equation 3. We found that Equation 3 with two terms fitted over the full range of Q^2 , and Equation 3 with only one term fitted from $Q^2 = 0.002$ to 0.003 (where the aggregate term in Equation 3 makes only a small contribution), yielded the same results within $1 \text{ \AA}$ for all data. As a result, we report radii of gyration with an uncertainty of $\pm 1 \text{ \AA}$, as opposed to the achievable experimental precision of $\pm 0.1 \text{ \AA}$. The maximum variation in the SAXS rate of λ_{6-85}^* was $\pm 1 \text{ sec}^{-1}$, depending on the range of Q and number of exponentials used. These are the uncertainties reported in Table 2. To provide a consistent analysis, the same range of Q^2 (0.002 to 0.003) was used for all kinetic data points. The transient aggregation phenomenon will be the subject of future SAXS studies at small Q , where transient aggregation and folding compete at comparable signal amplitudes, as in a recent study of U1A transient aggregation (Yang and Gruebele 2006).

Circular dichroism

Unfolded λ_{6-85}^* and mutant were diluted from 5 M guanidine buffer sevenfold (1:6) so as to initiate refolding. The refolding process was monitored by circular dichroism (CD) at 222 nm ($[\theta]_{222}$) at each temperature. Measurements were repeated and accumulated to obtain a good signal/noise ratio. The averaged data were normalized to yield molar ellipticity with $300 \text{ deg m}^{-1} \text{ M}^{-1}$ accuracy for steady-state measurements, and $\pm 1000 \text{ deg m}^{-1} \text{ M}^{-1}$ accuracy for kinetic measurements. For each kinetic measurement, in addition to collecting refolding data, we always performed two more experiments: mixing unfolded protein in unfolding buffer with the same unfolding buffer. This gave the initial CD level. After the refolding experiments, the solution was left to equilibrate for 20 min, and then remeasured. This gave the final level, or CD signal of the native state in

0.7 M guanidine. Final protein concentration in the CD experiments ranged from 32 to $37 \text{ }\mu\text{M}$.

Steady-state CD data were analyzed by singular value decomposition (SVD) in the 208–250 nm region (Henry and Hofrichter 1992). At smaller wavelengths, the signal-to-noise ratio of the CD data was not sufficient for reliable SVD due to low signals from denaturant absorption. As can be seen in Figure 3, the first two SVD basis functions are sufficient to reconstruct the most important features of the CD spectra as a function of denaturant concentration, at least over the wavelength region probed. Similarly, the signal at 222 nm follows a sigmoidal curve with only small baselines. However, other reported variants of λ_{6-85}^* , particularly ones folding downhill or near-downhill, do have significant baselines (Yang and Gruebele 2004a; 2004c).

Viscosity formula

The viscosity η depends on the mole fraction χ of ethylene glycol and temperature as

$$\eta(T, \chi) = (0.004757 + 0.047\chi) \{1 + (221 + 573\chi)e^{-0.048T/^\circ\text{C}} + (154 - 69\chi)e^{-0.01T/^\circ\text{C}}\}. \quad (4)$$

This empirical expression was obtained by fitting 22 viscosity data points for water/ethylene glycol mixtures at various concentrations and temperatures (Bolz and Tuve 1973; Weast 1997). Note that in Equation 2, we adopt a power law dependence of the rate coefficient on viscosity. Internal protein friction, leading to a $(\eta_{\text{Protein}} + \eta_{\text{Solvent}})^{-1}$ have also been proposed, but it is not currently known what, if any, temperature correction applies to the internal friction in such models, and it does not fit available data well (Plaxco and Baker 1998; Bhattacharyya and Sosnick 1999). Our previous work on cold denaturation of lambda repressor showed that $a \approx 1$ can account for the observed data down to 3°C using Equation 2; other experiments and modeling also indicate that $a \approx 1$ for protein folding (Klimov and Thirumalai 1997; Plaxco and Baker 1998), so the $a \approx 1$ is justified by the available theoretical and experimental data.

Acknowledgments

H.K. and Y.M. were supported by the Japan–U.S. Cooperative Science program. M.G. and C.D. were supported by National Science Foundation grant NSF MCB-0316925. S.J.K. was supported by the Human Frontiers in Science Program. Use of the Advanced Photon Source under proposal GU-2830 was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under Contract No. W-31-109-22ENG-38. The BioCAT beamline is a National Institutes of Health supported Research Center, Grant RR-08630.

References

- Arai, M., Ikura, T., Semisotnov, G.V., Kihara, H., Amemiya, V., and Kuwajima, K. 1998. Kinetic refolding of β -lactoglobulin—Studies by synchrotron X-ray scattering, and circular dichroism, absorption and fluorescence spectroscopy. *J. Mol. Biol.* **275**: 149–162.
- Arora, P., Oas, T.G., and Myers, J.K. 2004. Fast and faster: A designed variant of the B-domain of protein A folds in 3 μsec . *Protein Sci.* **13**: 847–853.

- Baldwin, R.L. 1995. The nature of protein folding pathways: The classical versus the new view. *J. Biomol. NMR* **5**: 103–109.
- Bhattacharyya, R.P. and Sosnick, T.R. 1999. Viscosity dependence of the folding kinetics of a dimeric and monomeric coiled coil. *Biochemistry* **38**: 2601–2609.
- Bolz, R.E. and Tuve, G.L. 1973. *CRC handbook of tables for applied engineering science*, 2nd ed. CRC Press, Boca Raton, FL.
- Brutscher, B., Bruschweiler, R., and Ernst, R.R. 1997. Backbone dynamics and structural characterization of the partially folded state of ubiquitin by ^1H , ^{13}C , and ^{15}N nuclear magnetic resonance spectroscopy. *Biochemistry* **36**: 13043–13053.
- Bryngelson, J.D., Onuchic, J.N., Socci, N.D., and Wolynes, P.G. 1995. Funnels, pathways, and the energy landscape of protein folding: A synthesis. *Proteins* **21**: 167–195.
- Faraoane-Mennella, J., Gray, H.B., and Winkler, J.R. 2005. Early events in the folding of four-helix-bundle heme proteins. *Proc. Natl. Acad. Sci.* **102**: 6315–6319.
- Ferguson, N., Schartau, P.J., Sharpe, T.D., Sato, S., and Fersht, A.R. 2004. One-state downhill versus conventional protein folding. *J. Mol. Biol.* **344**: 295–301.
- Fischetti, R., Stepanov, S., Rosenbaum, G., Barrea, R., Black, E., Gore, D., Heurich, R., Kondrashkina, E., Kropf, A.J., Wang, S., et al. 2004. The BioCAT undulator beamline 18ID: A facility for biological non-crystalline diffraction and X-ray absorption spectroscopy at the advanced photon source. *J. Synchrotron Radiat.* **11**: 399–405.
- Garcia-Mira, M., Sadqi, M., Fischer, N., Sanchez-Ruiz, J.M., and Muñoz, V. 2002. Experimental identification of downhill protein folding. *Science* **298**: 2191–2195.
- Ghaemmaghami, S., Word, J.M., Burton, R.E., Richardson, J.S., and Oas, T.G. 1998. Folding kinetics of a fluorescent variant of monomeric λ repressor. *Biochemistry* **37**: 9179–9185.
- Gruebele, M. 2005. Downhill protein folding: Evolution meets physics. *C.R. Biol.* **328**: 701–712.
- Guinier, A. and Fournet, G. 1955. *Small-angle scattering of X-rays*. John Wiley & Sons, Inc., New York.
- Henry, E.R. and Hofrichter, J. 1992. Singular value decomposition: Application to analysis of experimental data. *Methods Enzymol.* **210**: 129–192.
- Huang, G.S. and Oas, T.G. 1995. Submillisecond folding of monomeric λ repressor. *Proc. Natl. Acad. Sci.* **92**: 6878–6882.
- Humphrey, W.F., Dalke, A., and Schulten, K. 1996. VMD visual molecular dynamics. *J. Mol. Graph.* **14**: 33–38.
- Jäger, M., Zhang, J., Bieschke, J., Nguyen, H., Dendle, G., Bowman, M., Noel, J., Gruebele, M., and Kelly, J. 2006. The structure-function-folding relationship in a WW domain. *Proc. Natl. Acad. Sci.* **108**: 10648–10653.
- Kim, D.E., Gu, H., and Baker, D. 1998. The sequences of small proteins are not extensively optimized for rapid folding by natural selection. *Proc. Natl. Acad. Sci.* **95**: 4982–4986.
- Klimov, D.K. and Thirumalai, D. 1997. Viscosity dependence of the folding rates of proteins. *Phys. Rev. Lett.* **79**: 317–320.
- Kohn, J.E., Millet, I.S., Jacob, J., Zagrovic, B., Dillon, T.M., Cingel, N., Dothager, R.S., Seifert, S., Thiagarajan, P., Sosnick, T.R., et al. 2004. Random-coil behavior and the dimensions of chemically unfolded proteins. *Proc. Natl. Acad. Sci.* **101**: 12491–12496.
- Kramers, H.A. 1940. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* **7**: 284–304.
- Krantz, B.A. and Sosnick, T.R. 2000. Distinguishing between two-state and three-state models for ubiquitin folding. *Biochemistry* **39**: 11696–11701.
- Kubelka, J., Eaton, W.A., and Hofrichter, J. 2003. Experimental tests of villin subdomain folding simulations. *J. Mol. Biol.* **329**: 625–630.
- Kubelka, J., Hofrichter, J., and Eaton, W.A. 2004. The protein folding “speed limit.” *Curr Opin. Struct. Biol.* **14**: 76–88.
- Kuwajima, K., Yamaya, H., and Sugai, S. 1996. The burst-phase intermediate in the refolding of β -lactoglobulin studied by stopped-flow circular dichroism and absorption spectroscopy. *J. Mol. Biol.* **264**: 806–822.
- Lacroix, E., Viguera, A.R., and Serrano, L. 1998. Elucidating the folding problem of α -helices: Local motifs, long-range electrostatics, ionic-strength dependence and prediction of NMR parameters. *J. Mol. Biol.* **284**: 173–191.
- Larios, E., Li, J.S., Schulten, K., Kihara, H., and Gruebele, M. 2004. Multiple probes reveal a native-like intermediate during low-temperature refolding of ubiquitin. *J. Mol. Biol.* **340**: 115–125.
- Larios, E., Pitera, J.W., Swope, W.C., and Gruebele, M. 2006. Correlation of early orientational ordering of engineered λ 6–85 structure with kinetics and thermodynamics. *Chem. Phys.* **323**: 45–53.
- Ma, H. and Gruebele, M. 2005. Kinetics are probe-dependent during downhill folding of an engineered λ 6–85 protein. *Proc. Natl. Acad. Sci.* **102**: 2283–2287.
- Muñoz, V. 2002. Thermodynamics and kinetics of downhill protein folding investigated with a simple statistical mechanical model. *Int. J. Quantum Chem.* **90**: 1522–1528.
- Muñoz, V., ed. 2006. *Protein folding and misfolding*. Royal Society of Chemistry, Cambridge, UK.
- Muñoz, V. and Sanchez-Ruiz, J.M. 2004. Exploring protein-folding ensembles: A variable-barrier model for the analysis of equilibrium unfolding experiments. *Proc. Natl. Acad. Sci.* **101**: 17646–17651.
- Naganathan, A.N., Perez-Jimenez, R., Sanchez-Ruiz, J.M., and Muñoz, V. 2005. Robustness of downhill folding: Guidelines for the analysis of equilibrium folding experiments on small proteins. *Biochemistry* **44**: 7435–7449.
- Nguyen, H., Jaeger, M., Kelly, J., and Gruebele, M. 2005. Engineering a β -sheet protein towards the folding speed limit. *J. Phys. Chem. B* **109**: 15182–15186.
- Nozaki, Y. 1970. The preparation of guanidine hydrochloride. *Methods Enzymol.* **26**: 43–50.
- Pace, C.N. 1986. Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol.* **131**: 266–280.
- Plaxco, K.W. and Baker, D. 1998. Limited internal friction in the rate-limiting step of a two-state protein folding reaction. *Proc. Natl. Acad. Sci.* **95**: 13591–13596.
- Plaxco, K.W., Millett, S., Segel, D.J., Doniach, S., and Baker, D. 1999. Chain collapse can occur concomitantly with the rate-limiting step in protein folding. *Nat. Struct. Biol.* **6**: 554–556.
- Qin, Z., Hu, D., Shimada, L., Nakagawa, T., Arai, M., Zhou, J., and Kihara, H. 2001. Refolding of β -lactoglobulin studied by stopped-flow circular dichroism at subzero temperatures. *FEBS Lett.* **507**: 299–302.
- Qin, Z., Ervin, J., Larios, E., Gruebele, M., and Kihara, H. 2002. Formation of a compact structured ensemble without fluorescence signature early during ubiquitin folding. *J. Phys. Chem. B* **106**: 13040–13046.
- Qiu, L.L., Pabit, S.A., Roitberg, A.E., and Hagen, S.J. 2002. Smaller and faster: The 20-residue Trp-cage protein folds in 4 μs . *J. Am. Chem. Soc.* **124**: 12952–12953.
- Sabelko, J., Ervin, J., and Gruebele, M. 1999. Observation of strange kinetics in protein folding. *Proc. Natl. Acad. Sci.* **96**: 6031–6036.
- Sadqi, M., Lapidus, L.J., and Muñoz, V. 2003. How fast is protein hydrophobic collapse. *Proc. Natl. Acad. Sci.* **100**: 12117–12122.
- Segel, D.J., Bachmann, A., Hofrichter, J., Hodgson, K., Doniach, S., and Kiefhaber, T. 1999. Characterization of transient intermediates in lysozyme folding with time-resolved small-angle X-ray scattering. *J. Mol. Biol.* **288**: 489–499.
- Svergun, D., Barberato, C., and Koch, M.H.J. 1995. CRYSOLE: A program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.* **28**: 768–773.
- Tanford, C. 1968. Protein denaturation. *Adv. Protein Chem.* **23**: 121–282.
- Weast, R.C. 1997. *CRC handbook of chemistry and physics*. CRC Press Inc., Cleveland, OH.
- Yang, W.Y. and Gruebele, M. 2003. Folding at the speed limit. *Nature* **423**: 193–197.
- Yang, W.Y. and Gruebele, M. 2004a. Folding λ -repressor at its speed limit. *Biophys. J.* **87**: 596–608.
- Yang, W.Y. and Gruebele, M. 2004b. Detection-dependent folding kinetics as a probe of folding landscape microstructure. *J. Am. Chem. Soc.* **126**: 7758–7759.
- Yang, W.Y. and Gruebele, M. 2004c. Rate-temperature relationships in λ repressor fragment 6–85 folding. *Biochemistry* **43**: 13018–13025.
- Yang, W.Y. and Gruebele, M. 2005. Kinetic equivalence of the heat and cold structural transitions of λ 6–85. *Philos. Trans. R. Soc. Lond. B Biol. Soc.* **43**: 13018–13025.
- Yang, W.Y. and Gruebele, M. 2006. Binary and ternary aggregation with tethered protein constructs. *Biophys. J.* **90**: 2930–2937.
- Yang, A. and Honig, B. 1993. On the pH dependence of protein stability. *J. Mol. Biol.* **231**: 459–474.
- Yang, W., Larios, E., and Gruebele, M. 2003. On the extended β -conformational propensity of polypeptides at high temperature. *J. Am. Chem. Soc.* **125**: 16220–16227.
- Zhu, Y., Alonso, D.O.V., Maki, K., Huang, C.-Y., Lahr, S.J., Daggett, V., Roder, H., DeGrado, W.F., and Gai, F. 2003. Ultrafast folding of α -3-D, a de novo designed three-helix bundle protein. *Proc. Natl. Acad. Sci.* **100**: 15486–15491.