doi:10.1016/S0022-2836(02)01029-X available online at http://www.idealibrary.com on IDEAL® J. Mol. Biol. (2002) 324, 359–371

JMB

Fast and Slow Intermediate Accumulation and the Initial Barrier Mechanism in Protein Folding

Bryan A. Krantz[1], Leland Mayne[2], Jon Rumbley[2], S. Walter Englander[2]* and Tobin R. Sosnick[1,3]*

[1]*Department of Biochemistry and Molecular Biology University of Chicago Chicago, IL 60637, USA*

[2]*Department of Biochemistry and Biophysics, The Johnson Research Foundation University of Pennsylvania 242 Anatomy-Chemistry Bldg Philadelphia, PA 19104-6059 USA*

[3]*Institute for Biophysical Dynamics, University of Chicago, Chicago, IL 60637 USA*

**Corresponding authors*

Do stable intermediates form very early in the protein folding process? New results and a quantity of literature that bear on this issue are examined here. Results available provide little support for early intermediate accumulation before an initial search-dependent nucleation barrier.

Keywords: protein folding; barriers in protein folding; intermediates in protein folding; burst phase; initial barrier hypothesis

Introduction

Earlier work suggested that protein folding rates might be limited by an initial conformational search to organize some sizable part of the native state topology.[1] This structure provides a transition state, a large topological nucleus, that can for the first time guide and support forward folding steps in an energetically downhill manner. One implication is that intermediates that are stable relative to the unfolded state cannot form and accumulate before the initial nucleation step. They can form only later, and can then accumulate only when blocked by later barriers, which we believe are often due to non-obligatory misfold reorganization events.[2]

An initial barrier mechanism, rate-limited by the organization of a large-scale topological nucleus, would have multiple consequences. The extensive organization of the transition state would account for the large surface burial commonly found ($m_f^{\ddagger}$). A whole molecule search to find the rate-limiting transition state would explain the correlation between folding rate and various measures of how rapidly residues that are in contact in the native topology can find each other.[3–5] Most interesting for the present discussion, a rate-limiting initial barrier mechanism would explain the prevalence of two-state folding because stable intermediates can occur only afterwards.

Following a period of surprise that any protein could fold in a kinetically two-state manner, without the population of intermediates, about 40 small proteins are now known to do so (S. Jackson, personal communication), consistent with our expectation.[6] However, exceptions have been reported in which intermediates appear to accumulate in an early phase, before the initial search-dependent barrier. The present work examines the major exceptions.

Abbreviations used: Cyt *c*, cytochrome *c*; RNase A or H, ribonuclease A or H; Ub, ubiquitin; HX, hydrogen exchange; Fl, fluorescence; GdmCl, guanidinium chloride.

E-mail addresses of the corresponding authors: walter@hx2.med.upenn.edu; trsosnic@midway.uchicago.edu

Results

Burst phase signals

In folding experiments some proteins display "burst phase" signals on a fast time-scale that is

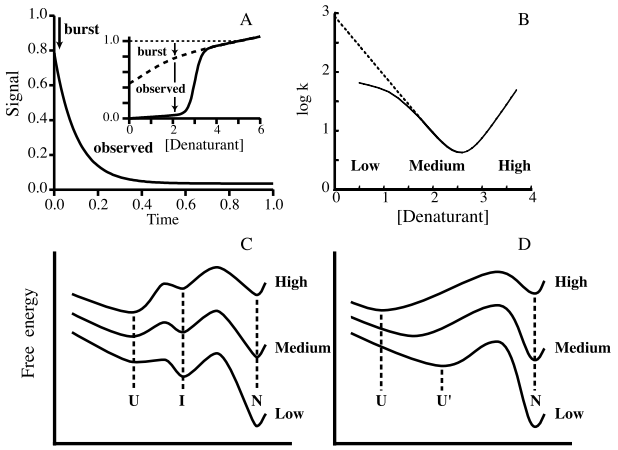

Figure 1. Burst phase and chevron rollover. (A) Illustrative kinetic trace for folding after denaturant dilution, showing an unmeasured burst phase and a slower observable phase. The denaturant-dependence of both phase amplitudes is illustrated in the inset. (B) Illustrative kinetic chevron showing the usual linear denaturant-dependence of folding and unfolding rates, but with a rollover in the rate of folding to N at a low concentration of denaturant (rate and *m* value less than expected). (C) and (D) Folding free energy profiles that might explain initial burst phase and rollover behavior. In (C), a discrete intermediate is pictured to form in a fast barrier-crossing event, producing burst phase signals (small barrier) and folding rates slower than expected (chevron rollover) at a low concentration of denaturant. In (D), the unfolded molecules when placed into a lower concentration of denaturant relax downhill to a more compact but still unfolded ensemble, producing burst phase signals and chevron rollover.

inaccessible to the usual stopped-flow instrumentation (Figure 1(A)). The same proteins tend to exhibit a rollover in folding rate at a low concentration of denaturant (Figure 1(B)). These observations have been rationalized in terms of a structured intermediate, formed in a fast barrier-crossing event and stabilized by the decreasing denaturant, as diagrammed in Figure 1(C).[7,8] However other alternatives can generate the same signals. When the unfolded polypeptide is mixed from a high concentration of denaturant (good solvent) into a lower concentration of denaturant (poor solvent), the U state ensemble can be expected to experience some solvent-dependent chain contraction/collapse, as in Figure 1(D).[1,2,9–14] Other possibilities include aggregation,[15] some trivial interaction that affects the signal (e.g. the fluorescence (Fl) of the exceedingly hydrophobic tryptophan ring), and experimental or interpretational error.

As one test of the initial barrier view, the analysis described here considers a variety of observations that have been interpreted in terms of the early formation and accumulation of a structured intermediate, before the initial search succeeds.

Protein G, B1 domain (GB1; 56 residues)

An early Fl study of GB1 as a function of temperature found two-state folding,[16] but subsequent work seems to disagree. Structural information was sought by Kuszewski *et al*.[17] in a hydrogen exchange (HX) competition labelling experiment, with folding initiated by a pH jump (zero denaturant). Measured HX protection was interpreted in terms of a fast-folding intermediate. However, when the calculated HX protection factors in the fast phase are corrected using the more recent intrinsic (unprotected) HX rates reported by Bai *et al*.,[18] almost all are below 3.5 (24 out of 26 residues measured). Even then, the low protection factors exhibit a strong correlation ($r = 0.83$) with the intrinsic amide HX rates, presenting the unlikely result that the core hydrophobic residues in the putative intermediate are the least protected and the polar surface residues the most protected. These observations might be explained by a small systematic error with true protection factors close to 1, indicating no fast intermediate formation. In agreement, Baker and co-workers observed that the folding reactions of GB1 and many mutants are well described by a single exponential without missing amplitude and with adherence to the chevron criteria (see below).[19]

A more recent conflict has arisen. Park *et al*. measured the folding of GB1 in 0.4 M Na_2SO_4 by fast continuous-flow (0.2–1.3 ms; Figure 2(A)) and slower stopped-flow (>3 ms; Figure 2(B)) methods, using a single tryptophan Fl probe.[20] When these non-overlapping kinetic traces were merged, they appeared to diverge from a single exponential (residuals <5%).[20] A better fit was obtained with two exponentials with rates differing by a small factor. This result was taken to indicate a sequential three-state model with two different barrier-crossing events and, on that basis, a distinct populated intermediate.[20]

To explore these ambiguities, we measured GB1 refolding in the presence and in the absence of 0.4 M Na_2SO_4, at 10 °C and 20 °C. We modified our instrument to reduce the dead-time to 1 ms and increase the signal to noise ratio. All the kinetic data could then be obtained within one instrument, spanning the gap where the biexponential break found by Park *et al*. was most pronounced (e.g. see Figure 2(C) and (D)). All of the data show accurate mono-exponential behavior with no missing amplitude, and they produce a linear chevron plot with no rollover (Figure 2(E) and (F)).

Most definitively, the results satisfy the stringent chevron criteria for two-state folding without intermediate accumulation (Table 1). In this test, one compares the free energy difference between U and N states, measured independently in equilibrium melting and in kinetic folding experiments ($\Delta G^\circ = -RT \ln k_u/k_f$). A similar comparison can be made for surface burial parameters ($m^0 = m_f^\ddagger - m_u^\ddagger$). Agreement between the equilibrium and kinetic parameters ensures that no

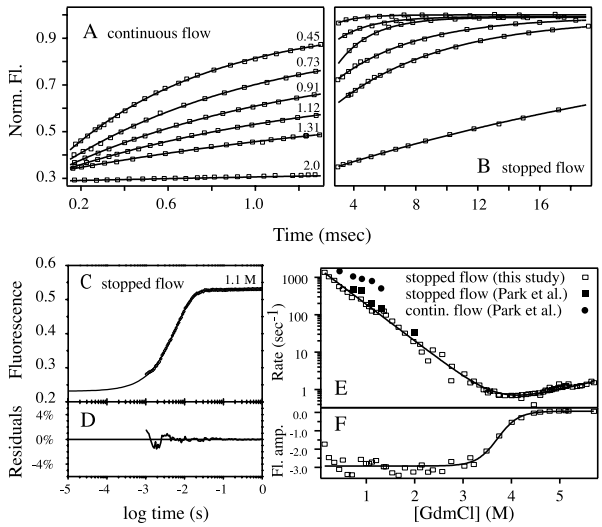

Figure 2. GB1 folding. (A) and (B) Kinetic traces from Park *et al*.[20] (Only a fraction of the ~1000 continuous flow data points are shown.) Each data trace fits a mono-exponential decay with residuals less than 1% (curves shown; GdmCl concentrations listed). New data obtained with a 1 ms dead-time instrument with and without 0.4 M Na_2SO_4 at 10 °C and 20 °C show mono-exponential folding (representative trace in (C)), no rollover ((E); 0.4 M Na_2SO_4, 20 °C), no missing amplitude ((F)), and adherence to the two-state chevron criteria (Table 2). The fitted rates from the data of Park *et al*. (filled symbols in (E)) are in good agreement with our stopped-flow data but the continuous-flow data diverge slightly, which leads to apparent biexponential behavior when the non-overlapping data sets are merged.

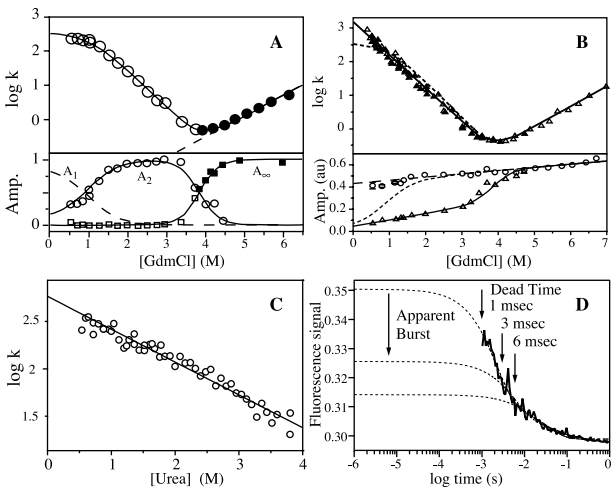

Figure 3. Ubiquitin folding and the dead-time artifact. (A) Folding chevron (top panel, Fl, 25 °C) and fitted normalized amplitudes (lower panel) for the apparent burst phase (A1), N (A2), and U (A3).[8,21] (B) New chevron and amplitude data at the same conditions showing no rollover or missing amplitude. Broken lines compare the rates and amplitudes from (A). (C) Folding rates down to a low concentration of urea (0.4 M Na_2SO_4, pH 5.0, 10 °C) show no rollover because the 1 ms dead-time and other precautions avoided factoring in the slower phases. In agreement, these and other results at a variety of conditions obey the chevron criteria for two-state folding (Table 2). (D) A dead-time artifact. When folding time constant approaches the dead-time, computer fitting of Ub data tends to include some of the slower phase. This underestimates the folding rate (chevron rollover), and extrapolation produces an apparent missing burst phase amplitude.

intermediate with significant stabilization free energy or surface burial is populated, i.e. that folding is operationally two-state (although non-populating intermediates may exist).

A review of the data presented by Park *et al*. finds that each individual kinetic trace can be fit with a mono-exponential decay with residuals of 1% or less (Figure 2(A) and (B)). When fit in this way, the stopped-flow rates match our data closely under the same conditions, but the continuous-flow results are slightly divergent (Figure 2(E)), leading to the non-monoexponential appearance of the merged data sets.

In summary, experiment finds HX protection factors less than 3.5,[17] accurately mono-exponential folding, no missing amplitude, no chevron rollover, and strict compliance with the chevron criteria. In contradiction, two sets of kinetic data taken in different instruments on non-overlapping time scales do not merge perfectly.[20] The weight of evidence appears to favor two-state folding without intermediate accumulation.

Table 1. Chevron parameters for GB1 folding (0.4 M Na_2SO_4)

	$\Delta G°$ (kcal mol^{-1})		m^0 (kcal mol^{-1} (M GdmCl)$^{-1}$)	
	Equilibrium	Kinetic	Equilibrium	Kinetic
10 °C	6.68 ± 0.89	6.84 ± 0.18	1.77 ± 0.24	1.81 ± 0.03
20 °C	6.80 ± 0.77	6.56 ± 0.22	1.83 ± 0.13	1.76 ± 0.04

Other parameters are: at 10 °C, $RT \ln k_f^{H_2O} = 3.98(\pm 0.10)$ kcal mol^{-1}, $m_f = 1.36(\pm 0.03)$ kcal mol^{-1} M^{-1}; at 20 °C, $RT \ln k_f^{H_2O} = 4.36(\pm 0.09)$ kcal mol^{-1}, $m_f = 1.32(\pm 0.02)$ kcal mol^{-1} M^{-1}.

Ubiquitin (Ub; 76 residues)

Khorasanizadeh *et al*. engineered a single tryptophan residue (F45W) into the core of ubiquitin to provide a Fl probe and used it to measure folding as a function of guanidinium chloride (GdmCl) concentration.[8,21] Folding at 25 °C but not at 8 °C exhibited a burst phase loss in Fl amplitude and a rollover in the folding chevron, already evident at 1 M GdmCl (Figure 3(A)). These results were interpreted in terms of the formation of a stable intermediate in the burst phase. In contradiction, HX labeling studies in the first 2 ms of folding at 20 °C (protection factor ~ 1), and stopped flow, far-UV CD, found no evidence for an early intermediate.[22]

We studied the fast-folding behavior of F45W Ub under these and other conditions using both GdmCl[23] and urea. We used a stopped-flow

Table 2. Chevron parameters for ubiquitin folding

	$\Delta G°$ (kcal mol^{-1})		m^0 (kcal mol^{-1} (M GdmCl)$^{-1}$)	
Condition	Equilibrium	Kinetic	Equilibrium	Kinetic
1	9.50 ± 0.37	10.2 ± 0.5	2.13 ± 0.09	2.23 ± 0.09
2	9.39 ± 0.37	9.70 ± 0.34	2.23 ± 0.09	2.28 ± 0.06
3	8.19 ± 0.10	8.40 ± 0.24	2.18 ± 0.02	2.17 ± 0.04
4	8.71 ± 0.22	9.28 ± 0.27	2.24 ± 0.05	2.24 ± 0.05
5	8.42 ± 0.25	8.20 ± 0.20	2.23 ± 0.06	2.18 ± 0.04

Conditions: (1) 25 °C, 0.23 M Na_2SO_4, pH 5; (2) 25 °C, 1 M NaCl, 20 mM sodium acetate, pH 4.5; (3) 25 °C, 20 mM sodium acetate, pH 5.0; (4) 15 °C, 1 M NaCl, 20 mM sodium acetate, pH 4.5; and (5) 5 °C, 1 M NaCl, 20 mM sodium acetate, pH 4.5.

apparatus with 1 ms dead-time and enhanced signal to noise ratio, and we used continuous-flow and double-jump protocols to minimize slow-folding phases due to proline mis-isomerization and other effects.[23] Our data are in excellent agreement with the GdmCl results presented by Khorasanizadeh *et al*. except for the chevron rollover (Figure 3(B), upper) and burst phase amplitude loss (Figure 3(B), lower). With our shorter instrument dead-time and minimal slow phases, these indicators did not appear, even in the presence of stabilizing Na_2SO_4, or down to 0.5 M urea in 0.4 M Na_2SO_4 (Figure 3(C)). The rollover previously observed at 300 s^{-1}, using instrumentation with a 2–3 ms dead-time, is not present. Further, our Ub data taken under a variety of conditions all obey the stringent chevron criteria for two-state folding, as shown in Table 2.

These observations suggest an interesting effect that might account for the discrepant results. Ub folding is heterogeneous with slower folding phases that account for ~23% of the total Fl amplitude.[8,21] When a part of the fast phase is lost within the instrument dead-time, straightforward computer fitting will inadvertently mix part of the slower phase into the rate determination. This effect, illustrated in Figure 3(D), will mimic a decreased folding rate (chevron rollover) and exaggerate the burst phase amplitude, which is obtained by extrapolating the artificially slow folding phase back to zero time.

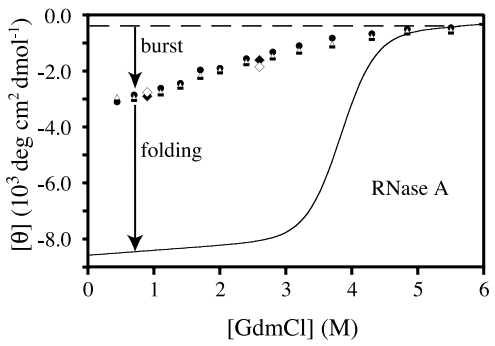

Figure 4. Ribonuclease A. Equilibrium and kinetic results for RNase A and its non-folding disulfide-broken variant (pH 7, 5 °C).[10] The continuous line is the equilibrium melting curve for the intact protein. Open symbols define the analogous melting curve for the non-folding disulfide broken analog, obtained both at equilibrium and after the kinetic burst phase upon denaturant dilution. Filled symbols show the identical burst phase for the intact protein. Open diamonds are from Houry *et al*.,[24,25] which led to the hypothesis of two different intermediates.

Ribonuclease A (RNase A; 124 residues)

Houry & Scheraga used double-jump experiments to avoid mis-isomerized proline residues and found burst phase CD behavior, suggesting stable intermediate formation very early in RNase A folding (Figure 4).[24,25] They measured the CD burst at two concentrations of GdmCl, assumed that these signals reflect two distinct folding intermediates in rapid equilibrium, and sought structural information by HX pulse labeling. The HX results, when interpreted with the assumption that exchange occurs only through the minor unprotected intermediate, suggested protection factors up to 100 in the more protected intermediate.

To help interpret the RNase A burst phase, we did control experiments with a non-folding disulfide broken analog. In a high concentration of denaturant, the non-folding polypeptide and the intact protein have little CD_{222} (Figure 4). When denaturant is decreased continuously, CD_{222} for the polypeptide increases in a continuous non-cooperative way, tracking changes in the equilibrium unfolded ensemble. When the polypeptide is diluted rapidly from a high concentration of denaturant, the CD jumps from its equilibrium high-denaturant (good solvent) value to its equilibrium low-denaturant (poor solvent) value in a submillisecond phase. Intact RNase A duplicates the burst phase jump quantitatively in the same non-cooperative way over the whole range of denaturant concentration (and then folds more slowly to N).

These results suggest that the burst phase signals reflect the solvent-dependence of the U state ensemble. When the published HX protection results[24] are computed straightforwardly, without any assumption about the character of possible intermediates, protection factors of 2.7 and 4 are found for the two most protected amide protons (Met29 and the disulfide-bridged Cys84) and 0.8–1.8 for the remaining 19 residues measured.[10]

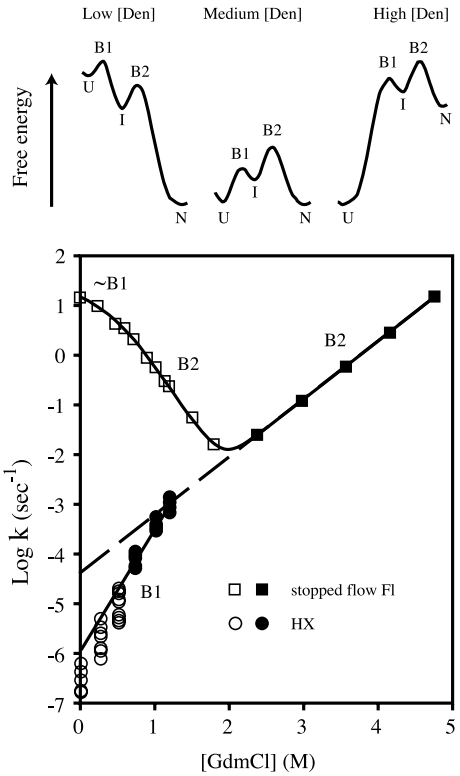

Figure 5. Barnase folding and unfolding. Unfolding rates curve down at low concentrations of denaturant (open and filled circles, calculated by Takei *et al.*[27] for multiple amide protons from HX data of Clarke & Fersht[68] in EX1 conditions). The curve-down accounts quantitatively for the missing ΔG_u° initially assigned to a burst phase intermediate. Fersht uses the pathway free-energy profiles shown to explain the folding roll-over, unfolding roll down, and other data in terms of a denaturant-dependent crossing over of barriers (B1 and B2), consistent with a late on-pathway intermediate that forms after the initial barrier, rather than an initial burst phase intermediate.[32]

These observations, i.e. the non-cooperative melting of the burst phase species, the identical burst phase signals for the protein and the non-folding polypeptide, and the absence of HX protection, do not support distinct, stable burst phase intermediates. All of the data consistently point to a simple readjustment of the U state ensemble to the new solvent condition.

Barnase (110 residues)

Indications that barnase might form a very fast intermediate include a burst phase in Fl, curvature in the folding chevron (Figure 5), HX pulse-labeling protection, and a failure to satisfy the ΔG and m chevron criteria for two-state folding.[5,23–26]

The evidence for a burst phase intermediate in barnase folding has been re-examined by Bai and co-workers.[26–28] They find no HX protection before native state formation (HX pulse labeling at pH 8.6). Earlier results to the contrary[29,30] apparently depended on aggregation.[28,32] HX competition experiments (pH 6.3–7.2) find protection factors of 2 or less, indicating that no significantly protected intermediate accumulates much before final folding.[26,27] This is true even in the presence of stabilizing concentrations of Na_2SO_4 up to 0.4 M.[28] In agreement, melting analysis of the burst phase amplitude as a function of denaturant concentration shows that any burst phase species has stability very close to U.[27] The previously missing 3 kcal mol^{-1} (1 cal = 4.184 J) of wild-type barnase stability (chevron criteria failure) is accounted for[27] by the fact that unfolding rates curve sharply downward at a low concentration of denaturant, well below the value that one finds by simple linear extrapolation from the chevron unfolding arm (Figure 5).

In response, Fersht[32] proposes that a real intermediate, perhaps unstable under the HX labeling conditions, may have been fortuitously stabilized by aggregation. He focuses on evidence for a late on-pathway kinetically significant intermediate and refers only secondarily to the apparent burst phase intermediate that most engaged Bai *et al.* The late intermediate escapes direct detection because it has no Fl signal, forms at a rate close to the final folding rate (50 ms *versus* 70 ms), and does not accumulate. Fersht proposes that the hidden intermediate is revealed by the sharp curve-down in the unfolding rate, which he explains in terms of two discrete barriers that cross over with changing denaturant (Figure 5 top).[32] An alternative, formally identical but physically distinct, is a denaturant-dependent change in the effective transition state position within a single broad barrier region.[28,31,33] However, Fersht argues that such a transition would not lead to a sharp break but a shallow curve.

Both of these alternatives account quantitatively for the missing stabilization free energy and the chevron rollover in terms of the denaturant dependence of the barriers. This leaves no missing stabilization free energy or surface burial for a putative burst phase intermediate, consistent with other indications that any early intermediate, if one exists, is no more stable than U. Fersht notes that a protection factor of about 2 measured in the HX competition study is quantitatively consistent with the slow formation rate of the late intermediate. This leaves no protection at all for any earlier intermediate. The late intermediate may have escaped detection because it is unstable at the high pH of the HX labeling pulse (in the absence of aggregation), but the HX competition experiments that find no protected fast intermediate were done at neutral pH. Thus the post-burst phase barnase population has free energy and surface exposure equal to U and no additional HX protection.

The remaining evidence for fast intermediate accumulation rests on the Fl burst phase. However, the denaturant dependence of the burst phase amplitude (melting analysis) indicates zero ΔG_u° relative to U for any burst phase intermediate.[27]

Further, the burst phase signals are fully consistent with an expected solvent-dependent readjustment of the U state ensemble (Figure 1(D)), as seen for RNase A (above) and cytochrome *c* (Cyt *c*) (below).

In summary, there is no substantive evidence for a burst phase barnase intermediate. Upon denaturant dilution, unfolded barnase produces burst phase signals but they record some fast solvent-dependent readjustment of the denatured ensemble.

Cytochrome *c* (Cyt *c*; 104 residues + heme)

A variety of observations (rollover, burst phase signals, HX protection, barrier crossing) have been interpreted in terms of the submillisecond formation of a well-structured Cyt *c* intermediate. We consider each data category here.

Chevron rollover

Cyt *c* shows a distinct chevron rollover in folding rate, suggesting an earlier condensed intermediate (Figure 6(A)). It appears, however, that the rollover reflects some aggregation process rather than intermediate formation. The rollover disappears when Cyt *c* folds at a very low concentration of protein (unpublished results), when the heme iron is protected by bound imidazole[34] (Figure 6(B)) or azide,[35] or in a mutant with the two peripheral histidine residues replaced by asparagine (Figure 6(C)). These observations indicate that the rollover depends on aggregation that is mediated by histidine to heme misligation.

The absence of chevron rollover when aggregation does not occur (Figure 6(B) and (C)) is inconsistent with the formation of a burst phase intermediate because the burial of any appreciable amount of surface before the rate-limiting step must reduce the slope of the chevron folding arm. Further, the red chevron in Figure 6(C) exhibits a strong H/^{2}H kinetic isotope effect, indicating that the majority of the helical H-bonds form in the rate-limiting step leading to N, long after the burst phase.[35,36]

Burst phase

Even in the absence of rollover, burst phase Fl loss continues to be seen (Figure 6(C), bottom). Energy transfer to the heme is extremely sensitive to small chain contraction, especially, since the Trp59 to heme distance is close to the Förster distance in the initial unfolded condition (both ~32 Å, see Figure 6(E), right axis).

To help interpret the burst phase signals, two large fragments, which lack one (1–80) or two (1–65) of the three major Cyt *c* helices, were used as non-folding models to calibrate the CD_{222} and Fl of the U state under native conditions (Figure 6(D) and (E)).[9] In concentrated denaturant, the non-folding analogs match the CD and Fl of the denatured protein. When denaturant is decreased continuously, the U ensemble changes its equilibrium distribution in a continuous non-cooperative way. In rapid dilution experiments, the CD and Fl of the truncated fragments change from their equilibrium high-denaturant (good solvent) values to their equilibrium low-denaturant (poor solvent) values in a submillisecond burst phase. When the same experiment is repeated with intact Cyt *c*,[9] it duplicates quantitatively the burst phase CD and Fl amplitudes found for the non-folding analogs in the same non-cooperative way over the whole range of denaturant concentration (and then folds more slowly to N).

Bhuyan & Udgaonkar[37] found the identical result for intact Cyt *c* held unfolded at pH 1.5, using urea as denaturant, and at higher temperature using GdmCl as denaturant. S. J. Hagen measured the burst phase kinetics of the fragments and the intact protein and found very similar rates (personal communication). Winkler and co-workers show that the post-burst phase Cyt *c* population is heterogeneneous[38–40] with approximately equal compact and extended fractions in rapid equilibrium, indicating that the collapsed conformation is not much more stable than the extended one.[39,40]

All of these results are inconsistent with the accumulation of a burst phase intermediate that is demonstrably different from U. The burst phase changes appear to reflect the solvent-dependence of the U state ensemble and track the time-scale for U state equilibration.

CD spectrum

In elegant continuous-flow experiments, Akiyama *et al*. were able to extend these observations by recording the major part of the post-burst phase CD spectrum with a 0.4 ms dead-time (Figure 6(F)).[41] When folding was initiated by denaturant dilution (from 4.4 M to 0.7 M GdmCl), the CD spectrum changed in the dead-time to one that is identical with the acid-unfolded U state,[41] the thermally unfolded U state,[9] and the non-folding Cyt *c* fragments.[9] When folding was initiated by a pH jump from the acid-unfolded U state (pH 2, zero denaturant, no salt, jumped to pH 4.5), the measured CD extrapolates back to the same acid-unfolded U state spectrum, i.e. there is no burst phase CD change.

The post-burst phase CD spectrum (Figure 6(F)) deconvolves to 8% helix (about eight helical residues, or less if aromatic residues contribute; native Cyt *c* has 40% helix).[41] The CD_{222} (per residue basis) matches the truncated fragments (Figure 6(D)) even though the fragments lack much (1–80) or most (1–65) of the segments that are helical in N, indicating that even the small CD found may be unrelated to native-like structure formation.

In summary, the CD spectrum recorded after the burst phase documents the equilibrium U state.

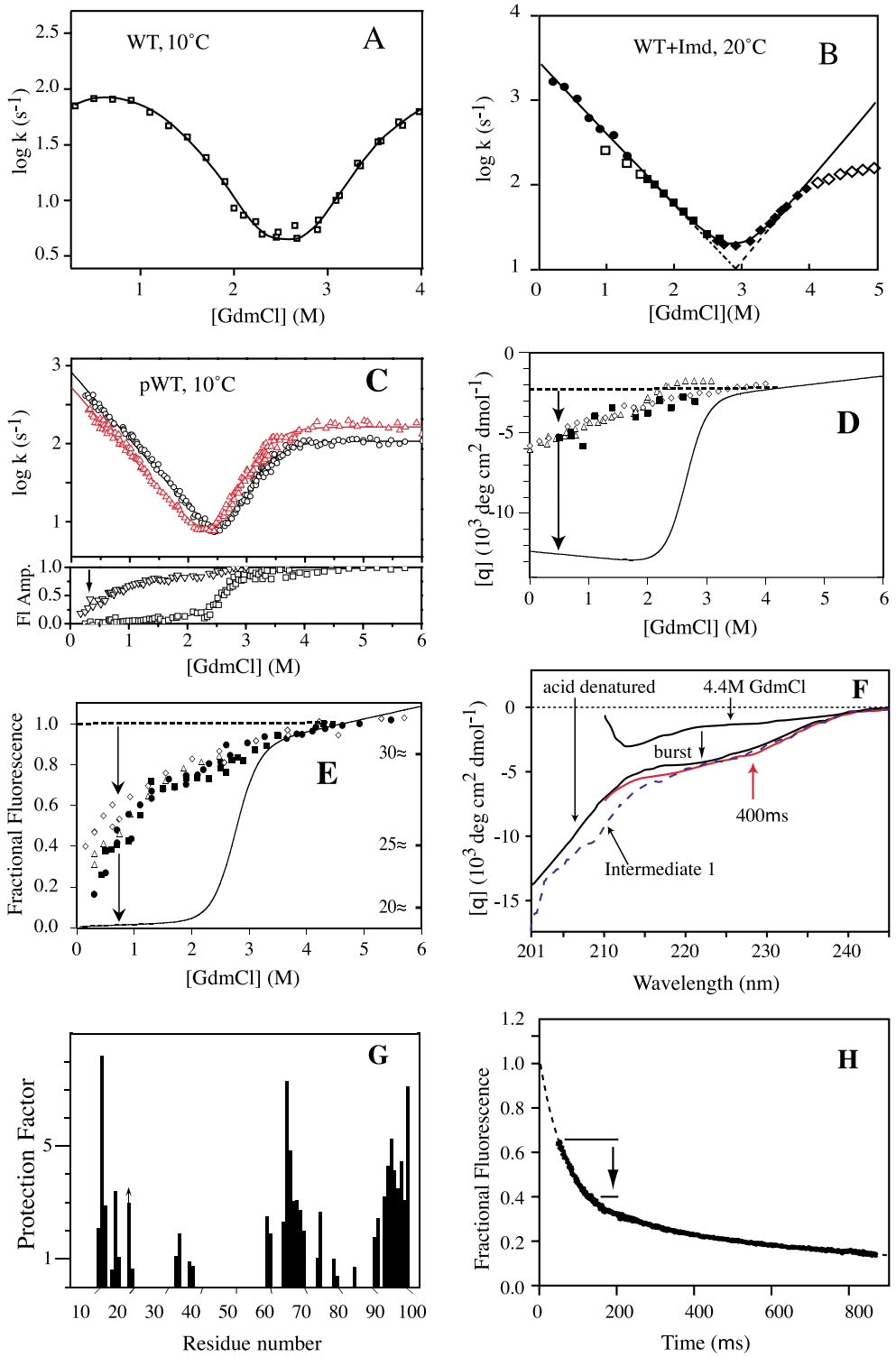

Figure 6. Cyt *c* folding. (A)–(C) Chevron plots for wild-type, imidazole-bound,[34] and pWT Cyt *c* (new data). Elimination of His to heme misligation eliminates aggregation and the resulting folding rollover. The red chevron in (C) shows the H/^{2}H isotope effect (Cyt *c* in H_2O but deuterated at exchangeable amide protons). (D) and (E) Equilibrium and kinetic results for intact Cyt *c* (closed symbols) and two non-folding analogs (open symbols; fragments 1–65 and 1–80; pH 4.9, 10 °C).[9] The continuous line is the equilibrium melting curve for the intact protein. The data show identical burst phase amplitudes for the non-folding analogs and the intact protein. (F) CD spectra. Upon denaturant dilution, Cyt *c* relaxes (400 μs, pH 4.9, 22 °C) to a condition with a CD spectrum (red[41]) equal to the acid-denatured state (shown), the thermally denatured state, and the non-folding fragments of (D) and (E) (computed as Intermediate I by Akiyama *et al.*, for which short wavelength data were poorly defined). (G) HX protection factors measured[42] for Cyt *c* by pH competition during 2 ms of folding at elevated pH (0.4 M Na_2SO_4, 0.3 M GdmCl, 10 °C). (H) Fast Fl quenching data for unfolded Cyt *c* (pH 2, low salt) jumped to native conditions (pH 4.5, 22 °C).[48] Data fitting includes the mono-exponential burst phase and slower phases. Horizontal lines mark the measured range of the ~50 μs burst phase.

HX protection

More detailed structural information for the burst phase intermediate was pursued in an HX competition experiment. Sauder & Roder found widespread HX protection in Cyt *c* within 2 ms of folding, with protection factors less than 5 for 32 of the 36 residues measured (in 0.4 M Na_2SO_4 at high pH; Figure 6(G)), and interpreted this in terms of the fast formation of the near-native helical content.[42] This interpretation is inconsistent with the absence of burst phase surface burial (Figure 6(B) and (C)), CD and Fl that are equal to the U state (Figure 6(D)–(F)), indications that half of the post-burst molecules have an extended conformation,[38–41] and amide H/^{2}H kinetic isotope results (Figure 6(C)), which indicate that the majority of the helical H-bonds form in the rate-limiting step leading to N, long after the burst phase.[35]

One does not know what level of HX protection should be considered significant in respect to specific structure formation. It is known that HX is impressively sensitive to local hydrophobic blocking.[43,44] For example, a single neighboring Val or Ile side-chain slows exchange fivefold.[18] Environmental slowing seems especially likely in the concentrated Na_2SO_4 used, which drives polypeptide chain condensation. Bieri & Kiefhaber find generalized low-level HX protection ($P \sim 5$) in the initially collapsed but unstructured form of hen lysozyme.[45] A contribution to the measured HX protection in Cyt *c* may come from aggregation (e.g. Figure 6(A) *versus* Figure 6(B) and (C)).

General blocking effects provide an alternative explanation for the small burst phase HX protection found in concentrated Na_2SO_4 for Cyt *c* and for CD2.D1.[46,47] It may seem indicative that the rapidly protected hydrogen atoms occur only in segments that are protected in the native protein; however, these are the only hydrogen atoms that can be measured.

Barrier crossing

An indirect argument favoring intermediate formation is that the Cyt *c* burst phase appears to represent a barrier-crossing event, and therefore distinct structure formation (Figure 1(C)), insofar as the rate exhibits mono-exponential kinetics, has significant activation enthalpy (7.5 kcal mol^{-1}), and is insensitive to conditions.

Shastry & Roder initiated folding from the acid-unfolded state of Cyt *c* and observed a fast multi-exponential decay of Fl due to Förster transfer as Trp59 approaches the heme (Figure 5(H)).[48] The earliest measured part of the burst phase Fl decay appears to be mono-exponential. However, data recording tracked less than half of the putative mono-exponential burst phase (Figure 5(H)), and this merges into slower phases. Thus, the fractional decay measured is likely to appear exponential in any case. The appearance of exponential relaxation was taken to demonstrate a barrier-crossing event. However, it has been pointed out that mono-exponentiality is not diagnostic for barrier crossing.[13] In agreement, simulations of a barrier-less Cyt *c* chain collapse (Figure 1(D)) show that the kinetics measured by Fl transfer can be indistinguishable from a single exponential.[49]

Shastry & Roder found that the burst phase rate (but not the amplitude) is insensitive to initial conditions (but not final conditions), including pH and the presence of heme ligand (0.2 M imidazole), as might be expected for a barrier crossing U → I reaction (Figure 1(C)).[48] However, these same characteristics are expected for a U → U′ model (Figure 1(D)). Similarly, the activation enthalpy measured for the Fl burst phase (7.5 kcal mol^{-1}) is close to expectations for simple chain contraction,[49] which should involve 4.4 kcal mol^{-1} for solvent viscosity plus other contributions, including main-chain bond rotations.[50]

Summary

The post-burst phase Cyt *c* ensemble exhibits little buried surface (no rollover under conditions that avoid aggregation). The Fl and the CD_{222} burst phase (amplitude and kinetics) match the U state and non-folding Cyt *c* models. The entire CD spectrum after the burst phase duplicates the U state and non-folding Cyt *c* models. Marginal HX protection is found even in concentrated osmolyte. The native helical H-bonds form much later (H/^{2}H isotope effect). The significance of a putative mono-exponential decay and other indicators is questionable on several grounds. All of this evidence is uniformly against the formation of a structured burst phase intermediate.

CD2.D1 (98 residues)

Clarke, Parker and co-workers studied the folding of the β-protein CD2.D1 following denaturant dilution.[44,45,49,50,52] They found chevron rollover and distributed low-level HX protection in the burst phase in the concentrated sulfate used, with HX protection factors less than 6 for 37 of the 41 NHs measured. These observations were taken to imply a burst phase intermediate with widespread H-bonding. In conflict, native-state HX measurements found no stably H-bonded intermediate.[53] Well-determined mutational ϕ_I values for six core mutants were 0.0 ± 0.1,[51,52] suggesting very little side-chain organization in the burst phase. The significance of low-level HX protection was considered above.

In light of the insecure relationship of chevron rollover and low-level HX protection to fast intermediate formation, and additional conflicting results, these observations do not strongly favor a stable burst phase intermediate.

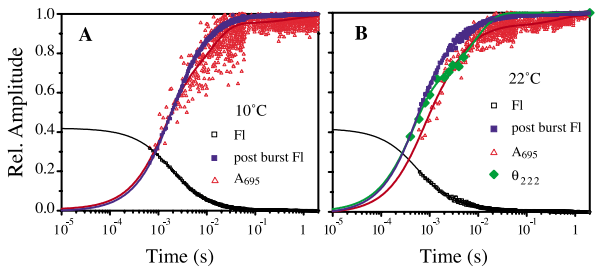

Figure 7. Heterogeneity and two-state folding in Cyt *c*. Folding kinetics are measured by Fl (initial collapse) and 695 nm absorbance (final N state acquisition) after a pH jump (pH 2 denatured jumped to pH 4.5 for folding). For direct comparison, the kinetically observable amplitudes of A695 (red) and Fl (blue) are each normalized to unity. (The Fl observed (descending) is shown normalized to total Fl.) After subtracting the ultrafast Fl solvent-response of the U ensemble, the probes for chain condensation (Fl) and final folding (A695) exhibit closely similar kinetics for the major phase. Thus, the large majority of the population folds in a two-state manner without intermediate accumulation (pH 4.5). Green points in (B) are from Akiyama *et al*.[41] The slowest kinetic phase, referred to by Akiyama *et al*. as intermediate II to N,[41,57] in fact reflects the folding of some small population fraction trapped by the optional pH-dependent barrier rather than an obligatory late barrier encountered by the entire population.

Other proteins

The literature reports on a number of other apparent burst phase intermediates with even less substantiation, most usually dependent on a Fl burst phase. In other cases, authentic intermediates, observed by methods with long multi-millisecond dead-times, have been referred to as burst phase intermediates. These are not discussed individually here.

Burst-phase summary

Major examples of burst phase structure formation were re-examined. In every case, an interpretation in terms of the accumulation of an intermediate is found to be unconvincing.

Late intermediates and folding heterogeneity

The results just reviewed are against the accumulation of intermediates before the initial search-nucleation step. However, well-structured on-pathway folding intermediates are known to accumulate in many cases. The thesis developed here suggests that intermediates can accumulate only later, blocked by barriers after the initial intrinsically rate-limiting step. These later barriers often seem to represent some non-intrinsic error-repair process rather than some difficulty intrinsic to the folding process, as for the initial search-nucleation barrier.[1,2,54]

The principles are these. For an intermediate to accumulate, it must occupy a well that is lower than all prior wells and must encounter a barrier that is higher (trough to peak) than all prior barriers. A later barrier (small-scale search) is very likely to be smaller than the initial intrinsic barrier (large-scale search). This is why folding tends to appear two-state, even though on-pathway intermediates that are stable relative to U do exist. A late barrier can be enlarged by corrupting the prior intermediate with interactions that must be removed (error repair) in the subsequent transition state. This can slow folding and cause accumulation of the corrupted intermediate. (Note that the stability of the intermediate is not the determining factor. Both slowing and accumulation are obligately linked consequences of the increased trough to peak height, which encodes the additional time required to reorganize the misfold.)

Cyt *c* provides an instructive case study. At low pH (e.g. pH 5), most of the Cyt *c* population folds to N in a two-state manner, limited by an initial barrier.[1] At neutral pH, Cyt *c* encounters an additional, later, misfold-reorganization barrier and reaches N more slowly, in a three-state manner.[2,55] This occurs because histidine residues in their high-pH form misligate avidly to the heme, inserting a reorganization barrier that slows folding and causes an otherwise normally occurring, but in this case somewhat corrupted, intermediate to accumulate.[56] In the middle pH range where histidine residues are partially titrated, Cyt *c* folding is heterogeneous, with a fast two-state fraction and a slower three-state histidine-misligated fraction.[2] The probability for barrier insertion and intermediate accumulation can be set between zero and 1 by titrating the histidine residues.[2,55] It is noteworthy that heterogeneous folding can easily be mistaken for parallel pathways or for multiple sequentially populated intermediates (see Figure 7).[41,57]

Heterogeneous folding with faster and slower fractions dependent on ambient conditions is seen for many proteins. The accumulation of a partially misfolded intermediate in Im7 with three of the four helices formed is promoted by reduced pH.[58,59] Lysozyme folds heterogeneously with faster and slower phases; misfolding and intermediate accumulation are promoted by acidic pH[12] and by high salt.[60] Many other proteins, like RNase A,[61–63] encounter optional barriers and populate late intermediates fractionally due to non-native proline isomers that pre-exist in the initial U-state ensemble. The probability for proline-dependent barrier insertion can be adjusted by varying the time that the protein is allowed to spend in the unfolded form.

Heterogeneous folding might reasonably be taken to show that a late barrier, encountered by only a fraction of the population, is optional and probabilistic rather than due to some intrinsically difficult step in structure formation. Once some advanced native-like intermediate is reached, it is hard to see why continuing structure formation,

able to build onto the existing structure, should be more difficult and slower. In 1994, we suggested that late barriers, intermediate accumulation, and heterogeneous folding might all be due to probabilistic misfolding and the need for error repair.[2] This view is now supported by results for Cyt *c*, Im7, lysozyme, RNase A and other proteins, but its generality remains an open issue.

We draw from the present analysis the observation that intermediates do not accumulate before the initial intrinsically rate-limiting barrier. This result is observed nearly universally, at least for small proteins, and therefore provides support for the initial barrier hypothesis. A corollary suggestion is that later intermediate accumulation, when it occurs, depends on a non-intrinsic error repair barrier inserted into the the folding pathway by some misfolding event.

Discussion

Major current issues in protein folding concern the rate-limiting barriers, their placement (early or late), their character (intrinsic or optional), and their height (how fast does folding go), and the character and role of intermediates (discrete or continuous, obligatory or optional, constructive or obstructive). The initial barrier hypothesis touches on all of these issues.

The present work was motivated by the belief that protein folding begins with an uphill search through U-state space.[1] The initial search culminates when it chances to find some conformation that is sufficiently large and native-like to support (nucleate) the downhill formation of stable structure. This view leads to the expectation that stable intermediates should not accumulate before the initial barrier. Many proteins are known to behave in this way,[6] but others have seemed to contradict this view. Here, we consider the major exceptions. Under closer examination, these exceptions disappear. These well-studied cases do not demonstrate that stable intermediates accumulate before the initial rate-limiting search-nucleation step.

The absence of stable intermediates on the unfolded side of the intrinsic initial barrier appears to be due to thermodynamic constraints. Collapse to a distinct species is inhibited by the loss of conformational entropy and the small number of possible stable intermediates. Enthalpically, the burial and desolvation of polar groups, both main-chain and side-chain, can occur only in a context that satisfies their H-bonding requirements,[36] with each failure costing several kcal mol^{-1}.[64] These exacting requirements tend to inhibit non-native chain condensation and make the conformational search for a functional transition state uphill from U in free energy.

The fact that the intrinsic rate-limiting barrier may occur very early in folding has been obscured by several factors. The rate-limiting barrier appears to be late in two-state folding when folding progress is measured by relative surface burial ($m_f^{\ddagger}/m^0$). Rather, major surface burial in the transition state ($m_f^{\ddagger}$) is a consequence of the extensive initial nucleation. Another confusing factor has been the accumulation of later intermediates in three-state folding, caused by late rate-limiting barriers. These may well be adventitious misfolding barriers not intrinsic to the uncorrupted folding process. Finally, it is noteworthy that the burst phase signals considered here relate to U state behavior. The character of the equilibrium denatured state is important for protein folding,[65] but it is important to distinguish pre-existing structure and subsequent structure formation.

The initial barrier view can be taken to question the interpretation of ultrafast folding events more generally. The fate of this expectation remains to be seen.

Materials and Methods

Proteins

Horse heart Cyt *c* and RNase A of the highest grade were purchased from Sigma Chemical Company. The CNBr fragments of Cyt *c* have been described.[9] Standard methods were used to reduce RNase A disulfides.[10] The preparation of the H26N,H33N variant of Cyt *c* (pWT) is described elsewhere.[66] Deuterated versions of Cyt *c* were made by incubating the protein in ^2H_2O in concentrated GdmCl. On dilution into H_2O at pH 4.5 and 10 °C, amide groups remain deuterated for tens of seconds. Ub was prepared as described.[23] The His_6 B1 domain of protein G was prepared from standard *Escherichia coli* expression lysates using an expression vector kindly provided by D. Baker.

Kinetic measurements

Rapid mixing experiments used a Biologic SFM4 or 400 stopped-flow apparatus as described.[23] To improve the dead-time performance, illumination through a fibre optic cable was restricted to the bottom 20% of the 0.8 mm cross-section FC-08 cuvette. High light intensity was obtained using a PTI A101 elliptical light source with 100 W Hg arc lamp. Background scattering was decreased with light baffles, and the PMT signal was filtered with an 11-pole elliptical filter with a 50 μs cut-off. The dead-time was calibrated using a dye-quenching reaction as well as multiple continuous-flow measurements at different flow speeds under the folding conditions.[23] Fluorescence spectroscopy used excitation wavelengths of 280–290 nm and emission wavelengths of 310–400 nm.

Data analysis

Kinetic data were analyzed using the chevron analysis[67] with the free energy of equilibrium folding and the activation free energy for kinetic folding and unfolding dependent on denaturant according to the standard equations. Parameters were fit using a

non-linear, least-squares algorithm implemented in the Microcal Origin software package.

Acknowledgements

We thank Y. Bai, D. Baker, R. L. Baldwin, W. Eaton, N. R. Kallenbach, I. Morishima, T. Pan, S. Radford, H. Roder, D. Rousseau, and members of our groups for numerous helpful discussions. This work was supported by grants from the NIH, the Mathers Foundation, and the Packard Foundation Interdisciplinary Science Program.

References

1. Sosnick, T. R., Mayne, L. & Englander, S. W. (1996). Molecular collapse: the rate-limiting step in two-state cytochrome *c* folding. *Proteins: Struct. Funct. Genet.* **24**, 413–426.
2. Sosnick, T. R., Mayne, L., Hiller, R. & Englander, S. W. (1994). The barriers in protein folding. *Nature Struct. Biol.* **1**, 149–156.
3. Plaxco, K. W., Simons, K. T. & Baker, D. (1998). Contact order, transition state placement and the refolding rates of single domain proteins. *J. Mol. Biol.* **277**, 985–994.
4. Gromiha, M. M. & Selvaraj, S. (2001). Comparison between long-range interactions and contact order in determining the folding rate of two-state proteins: application of long-range order to folding rate prediction. *J. Mol. Biol.* **310**, 27–32.
5. Mirny, L. & Shakhnovich, E. (2001). Protein folding theory: from lattice to all-atom models. *Annu. Rev. Biophys. Biomol. Struct.* **30**, 361–396.
6. Jackson, S. E. (1998). How do small single-domain proteins fold? *Fold. Des.* **3**, R81–R91.
7. Matouschek, A., Kellis, J. T., Jr., Serrano, L., Bycroft, M. & Fersht, A. R. (1990). Transient folding intermediates characterized by protein engineering. *Nature*, **346**, 440–445.
8. Khorasanizadeh, S., Peters, I. D., Butt, T. R. & Roder, H. (1993). Folding and stability of a tryptophan-containing mutant of ubiquitin. *Biochemistry*, **32**, 7054–7063.
9. Sosnick, T. R., Shtilerman, M. D., Mayne, L. & Englander, S. W. (1997). Ultrafast signals in protein folding and the polypeptide contracted state. *Proc. Natl Acad. Sci. USA*, **94**, 8545–8550.
10. Qi, P. X., Sosnick, T. R. & Englander, S. W. (1998). The burst phase in ribonuclease A folding and solvent dependence of the unfolded state. *Nature Struct. Biol.* **5**, 882–884.
11. Englander, S. W., Sosnick, T. R., Mayne, L. C., Shtilerman, M., Qi, P. X. & Bai, Y. (1998). Fast and slow folding in cytochrome *c*. *Accts Chem. Res.* **31**, 737–744.
12. Chen, L., Wildegger, G., Kiefhaber, T., Hodgson, K. O. & Doniach, S. (1998). Kinetics of lysozyme refolding: structural characterization of a non-specifically collapsed state using time-resolved X-ray scattering. *J. Mol. Biol.* **276**, 225–237.
13. Parker, M. J. & Marqusee, S. (1999). The cooperativity of burst phase reactions explored. *J. Mol. Biol.* **293**, 1195–1210.
14. Bhuyan, A. K. & Udgaonkar, J. B. (1999). Multistate kinetics of folding and unfolding of barstar. In *Old and New Views of Protein Folding* (Kuwajima, K. & Arai, M., eds), vol. 1194, pp. 261–270, Elsevier, Amsterdam.
15. Silow, M. & Oliveberg, M. (1997). Transient aggregates in protein folding are easily mistaken for folding intermediates. *Proc. Natl Acad. Sci. USA*, **94**, 6084–6086.
16. Alexander, P., Orban, J. & Bryan, P. (1992). Kinetic analysis of folding and unfolding of the 56 amino acid IgG-binding domain of streptococcal protein G. *Biochemistry*, **31**, 7243–7248.
17. Kuszewski, J., Clore, G. M. & Gronenborn, A. M. (1994). Fast folding of a prototypic polypeptide: the immunoglobulin binding domain of streptococcal protein G. *Protein Sci.* **3**, 1945–1952.
18. Bai, Y., Milne, J. S., Mayne, L. & Englander, S. W. (1993). Primary structure effects on peptide group hydrogen exchange. *Proteins: Struct. Funct. Genet.* **17**, 75–86.
19. McCallister, E. L., Alm, E. & Baker, D. (2000). Critical role of beta-hairpin formation in protein G folding. *Nature Struct. Biol.* **7**, 669–673.
20. Park, S. H., Shastry, M. C. & Roder, H. (1999). Folding dynamics of the B1 domain of protein G explored by ultrarapid mixing. *Nature Struct. Biol.* **6**, 943–947.
21. Khorasanizadeh, S., Peters, I. D. & Roder, H. (1996). Evidence for a 3-state model of protein folding from kinetic analysis of ubiquitin variants with altered core residues. *Nature Struct. Biol.* **3**, 193–205.
22. Gladwin, S. T. & Evans, P. A. (1996). Structure of very early protein folding intermediates: new insights through a variant of hydrogen exchange labelling. *Fold. Des.* **1**, 407–417.
23. Krantz, B. A. & Sosnick, T. R. (2000). Distinguishing between two-state and three-state models for ubiquitin folding. *Biochemistry*, **39**, 11696–11701.
24. Houry, W. A. & Scheraga, H. A. (1996). Structure of a hydrophobically collapsed intermediate on the conformational folding pathway of ribonuclease A probed by hydrogen–deuterium exchange. *Biochemistry*, **35**, 11734–11746.
25. Houry, W. A., Rothwarf, D. M. & Scheraga, H. A. (1996). Circular dichroism evidence for the presence of burst-phase intermediates on the conformational folding pathway of ribonuclease A. *Biochemistry*, **35**, 10125–10133.
26. Chu, R. A., Takei, J., Barchi, J. J., Jr & Bai, Y. (1999). Relationship between native-state hydrogen exchange and the folding pathway of barnase. *Biochemistry*, **38**, 14119–14124.
27. Takei, J., Chu, R. A. & Bai, Y. (2000). Absence of stable intermediates on the folding pathway of barnase. *Proc. Natl Acad. Sci. USA*, **97**, 10796–10801.
28. Chu, R.-A. & Bai, Y. (2002). Lack of definable nucleation sites in the rate-limiting transition state of barnase under native conditions. *J. Mol. Biol.* **315**, 759–770.
29. Bycroft, M., Matouschek, A., Kellis, J. T., Serrano, L. & Fersht, A. R. (1990). Detection and characterization of a folding intermediate in barnase by NMR. *Nature*, **346**, 488–490.
30. Matouschek, A., Serrano, L., Meiering, E. M., Bycroft, M. & Fersht, A. R. (1992). The folding of an enzyme. V. H/^{2}H exchange-nuclear magnetic resonance studies on the folding pathway of barnase:

complementarity to and agreement with protein engineering studies. *J. Mol. Biol.* **224**, 837–845.
31. Oliveberg, M. & Fersht, A. R. (1996). Formation of electrostatic interactions on the protein folding pathway. *Biochemistry*, **35**, 2726–2737.
32. Fersht, A. R. (2000). A kinetically significant intermediate in the folding of barnase. *Proc. Natl Acad. Sci. USA*, **97**, 14121–14126.
33. Otzen, D. E., Kristensen, O., Proctor, M. & Oliveberg, M. (1999). Structural changes in the transition state of protein folding: alternative interpretations of curved chevron plots. *Biochemistry*, **38**, 6499–6511.
34. Chan, C.-K., Hu, Y., Takahashi, S., Rousseau, D. L., Eaton, W. A. & Hofrichter, J. (1997). Submillisecond protein folding kinetics studied by ultrarapid mixing. *Proc. Natl. Acad. Sci. USA*, **94**, 1779–1784.
35. Krantz, B. A., Moran, L. B., Kentsis, A. & Sosnick, T. R. (2000). D/H amide kinetic isotope effects reveal when hydrogen bonds form during protein folding. *Nature Struct. Biol.* **7**, 62–71.
36. Krantz, B. A., Srivastava, A. K., Nauli, S., Baker, D., Sauer, R. T. & Sosnick, T. R. (2002). Understanding protein hydrogen bond formation with kinetic H/D amide isotope effects. *Nature Struct. Biol.* **9**, 458–463.
37. Bhuyan, A. K. & Udgaonkar, J. B. (1999). Relevance of burst phase changes in optical signals of polypeptides during protein folding. In *Perspectives in Structural Biology* (Vijayar, M., Yathindra, N. & Kolaskar, A. S., eds), pp. 295–303, Universities Press, Hyderabad.
38. Segel, D. J., Eliezer, D., Uversky, V., Fink, A. L., Hodgson, K. O. & Doniach, S. (1999). Transient dimer in the refolding kinetics of cytochrome *c* characterized by small-angle X-ray scattering. *Biochemistry*, **38**, 15352–15359.
39. Lyubovitsky, J. G., Gray, H. B. & Winkler, J. R. (2002). Mapping the cytochrome *c* folding landscape. *J. Am. Chem. Soc.* **124**, 5481–5485.
40. Tezcan, F. A., Findley, W. M., Crane, B. R., Ross, S. A., Lyubovitsky, J. G., Gray, H. B. & Winkler, J. R. (2002). Using deeply trapped intermediates to map the cytochrome *c* folding landscape. *Proc. Natl Acad. Sci. USA*, **99**, 8626–8630.
41. Akiyama, S., Takahashi, S., Ishimori, K. & Morishima, I. (2000). Stepwise formation of alpha-helices during cytochrome *c* folding. *Nature Struct. Biol.* **7**, 514–520.
42. Sauder, J. M. & Roder, H. (1998). Amide protection in an early folding intermediate of cytochrome *c*. *Fold. Des.* **3**, 293–301.
43. Bai, Y. & Englander, S. W. (1994). Hydrogen bond strength and beta-sheet propensities: the role of a side-chain blocking effect. *Proteins: Struct. Funct. Genet.* **18**, 262–266.
44. Milne, J. S., Mayne, L., Roder, H., Wand, A. J. & Englander, S. W. (1998). Determinants of protein hydrogen exchange studied in equine cytochrome *c*. *Protein Sci.* **7**, 739–745.
45. Bieri, O. & Kiefhaber, T. (2001). Origin of fast and non-exponential kinetics of lysozyme folding measured in pulsed hydrogen exchange experiments. *J. Mol. Biol.* **310**, 919–935.
46. Parker, M. J., Dempsey, C. E., Lorch, M. & Clarke, A. R. (1997). Acquisition of native beta-strand topology during the rapid collapse phase of protein folding. *Biochemistry*, **36**, 13396–13405.
47. Hosszu, L. L., Craven, C. J., Parker, M. J., Lorch, M., Spencer, J., Clarke, A. R. & Waltho, J. P. (1997). Structure of a kinetic protein folding intermediate by equilibrium amide exchange. *Nature Struct. Biol.* **4**, 801–804.
48. Shastry, M. C. R. & Roder, H. (1998). Evidence for barrier-limited protein folding kinetics on the microsecond time scale. *Nature Struct. Biol.* **5**, 385–392.
49. Hagen, S. J. (2002). Exponential decay kinetics in downhill protein folding. *Proteins: Struct. Funct. Genet.* In press.
50. Bolhuis, P. G., Dellago, C. & Chandler, D. (2000). Reaction coordinates of biomolecular isomerization. *Proc. Natl Acad. Sci. USA*, **97**, 5877–5882.
51. Lorch, M., Mason, J. M., Clarke, A. R. & Parker, M. J. (1999). Effects of core mutations on the folding of a beta-sheet protein: implications for backbone organization in the I-state. *Biochemistry*, **38**, 1377–1385.
52. Lorch, M., Mason, J. M., Sessions, R. B. & Clarke, A. R. (2000). Effects of mutations on the thermodynamics of a protein folding reaction: implications for the mechanism of formation of the intermediate and transition states. *Biochemistry*, **39**, 3480–3485.
53. Parker, M. J. & Marqusee, S. (2001). A kinetic folding intermediate probed by native state hydrogen exchange. *J. Mol. Biol.* **305**, 593–602.
54. Englander, S. W. (2000). Protein folding intermediates and pathways studied by hydrogen exchange. *Annu. Rev. Biophys. Biomol. Struct.* **29**, 213–238.
55. Elöve, G. A., Bhuyan, A. K. & Roder, H. (1994). Kinetic mechanism of cytochrome *c* folding: involvement of the heme and its ligands. *Biochemistry*, **33**, 6925–6935.
56. Roder, H., Elöve, G. A. & Englander, S. W. (1988). Structural characterization of folding intermediates in cytochrome *c* by H-exchange labeling and proton NMR. *Nature*, **335**, 700–704.
57. Akiyama, S., Takahashi, S., Kimura, T., Ishimori, K., Morishima, I., Nishikawa, Y. & Fujisawa, T. (2002). Conformational landscape of cytochrome *c* folding studied by microsecond-resolved small angle X-ray scattering. *Proc. Natl Acad. Sci. USA*, **99**, 1329–1334.
58. Gorski, S. A., Capaldi, A. P., Kleanthous, C. & Radford, S. E. (2001). Acidic conditions stabilise intermediates populated during the folding of Im7 and Im9. *J. Mol. Biol.* **312**, 849–863.
59. Capaldi, A. P., Kleanthous, C. & Radford, S. E. (2002). Im7 folding mechanism: misfolding on a path to the native state. *Nature Struct. Biol.* **9**, 209–216.
60. Bieri, O., Wildegger, G., Bachmann, A., Wagner, C. & Kiefhaber, T. (1999). A salt-induced kinetic intermediate is on a new parallel pathway of lysozyme folding. *Biochemistry*, **38**, 12460–12470.
61. Schultz, D. A., Schmid, F. X. & Baldwin, R. L. (1992). *Cis* proline mutants of ribonuclease A. II. Elimination of the slow-folding forms by mutation. *Protein Sci.* **1**, 917–924.
62. Udgaonkar, J. B. & Baldwin, R. L. (1995). Nature of the early folding intermediate of ribonuclease A. *Biochemistry*, **34**, 4088–4096.
63. Laurents, D. V., Bruix, M., Jamin, M. & Baldwin, R. L. (1998). A pulse-chase-competition experiment to determine if a folding intermediate is on or off-pathway: application to ribonuclease A. *J. Mol. Biol.* **283**, 669–678.
64. Honig, B. & Yang, A. S. (1995). Free energy balance in protein folding. *Advan. Protein Chem.* **46**, 27–58.
65. Rose, G. D. (2002). Unfolded proteins. *Advan. Protein Chem.* In press.
66. Rumbley, J. N., Hoang, L. & Englander, S. W. (2002). Recombinant equine cytochrome *c* in *E. coli*: high

level expression, characterization, and folding and assembly mutants. *Biochemistry*. In press.
67. Matthews, C. R. (1987). Effects of point mutations on the folding of globular proteins. *Methods Enzymol.* **154**, 498–511.
68. Clarke, J. & Fersht, A. R. (1996). An evaluation of the use of hydrogen exchange at equilibrium to probe intermediates on the protein folding pathway. *Fold. Des.* **1**, 243–254.

Edited by C. R. Matthews

(Received 2 April 2002; received in revised form 13 September 2002; accepted 17 September 2002)