The burst phase in ribonuclease A folding and solvent dependence of the unfolded state

Submillisecond burst phase signals measured in kinetic protein folding experiments have been widely interpreted in terms of the fast formation of productive folding intermediates. Experimental comparisons with non-folding polypeptide chains show that, for ribonuclease A and cytochrome *c*, these signals in fact reflect a shift from one biased ensemble of the unfolded state to another as a function of change in denaturant concentration.

When proteins are diluted from concentrated denaturant at the beginning of a stopped-flow folding experiment, they very generally exhibit a fast chain contraction, on a submillisecond time scale, that can be detected by various probes. This 'burst phase' behavior has often been interpreted in terms of the fast formation of productive folding intermediates[1]. A prior test of this hypothesis with cytochrome *c* however reached a different conclusion. Sosnick *et al.*[2,3] found that truncated cytochrome *c* variants that cannot fold produce the same burst phase fluorescence and circular dichroism (CD) signals as the intact protein. This suggests that the burst phase does not reflect the formation of a distinct folding intermediate but rather some solvent-dependent modification of the still unfolded polypeptide chain. A similar conclusion has been reached by Chan *et al.*[4] on the basis of experiments with intact cytochrome *c*.

Ribonuclease A (RNase A) provides a good system to test the generality of this conclusion. RNase A exhibits clear burst phase behavior on dilution from concentrated denaturant, and this has been interpreted in terms of the fast formation of two productive folding intermediates[5]. A modified form of RNase A with its four disulfide bonds broken has long served as the paradigmatic unfolded polypeptide[6] and its unfolded nature is supported by the absence of hydrogen exchange protection under strongly native conditions[7]. This paper compares the burst phase signals produced by disulfide-intact and disulfide-broken RNase A.

The unfolded state

As has been observed for many proteins[8,9], thermally unfolded RNase A retains some residual ellipticity. Fig. 1*a* shows that the disulfide-broken polypeptide (triangles) matches the residual ellipticity of unfolded disulfide-intact RNase A and maintains this value essentially unchanged at temperatures ranging from 5–95 °C.

Fig. 1*b* shows the analogous behavior as a function of guanidinium chloride (GmCl) concentration. Again here RNase A exhibits residual ellipticity after two-state melting and this is matched by disulfide-broken RNase A. When GmCl is further increased, both unfolded chains lose their residual CD in an identical manner. When GmCl concentration is decreased, the protein cooperatively recovers its native structure but the disulfide-broken chain returns in a non-cooperative way to the value seen for the thermally unfolded chain at zero GmCl.

These results indicate that the disulfide-broken polypeptide provides a good model for the unfolded protein. It matches the CD of unfolded RNase A at high denaturant, and the somewhat different CD of unfolded RNase A at high temperature, and it maintains the same unfolded condition even at low temperatures.

Although the unfolded conformation is insensitive to temperature, it does change with denaturant. Denaturants are well known to affect the parameters of unfolded polypeptides due to main chain solvation effects[8,10] and the solublization of apolar side chains that act to condense the unfolded chain[11]. These properties provide the physical basis for the burst phase behavior addressed here.

Burst phase kinetics of unfolded state transitions

Both forms of RNase A show a burst phase increase in CD on dilution from high GmCl (arrow in Fig. 2). The disulfide-broken polypeptide reaches its new equilibrium condition in the burst phase and shows no further change in time. The disulfide-intact protein shows the same burst phase change and then folds much more slowly to the native state.

Fig. 3 shows the ellipticity reached in the burst phase as a function of final denaturant concentration. The disulfide-broken polypeptide (open triangles) reaches its equilibrium value (Fig. 1*b*) in the burst phase and changes no further. The disulfide-intact protein shows an identical burst phase (closed cir-

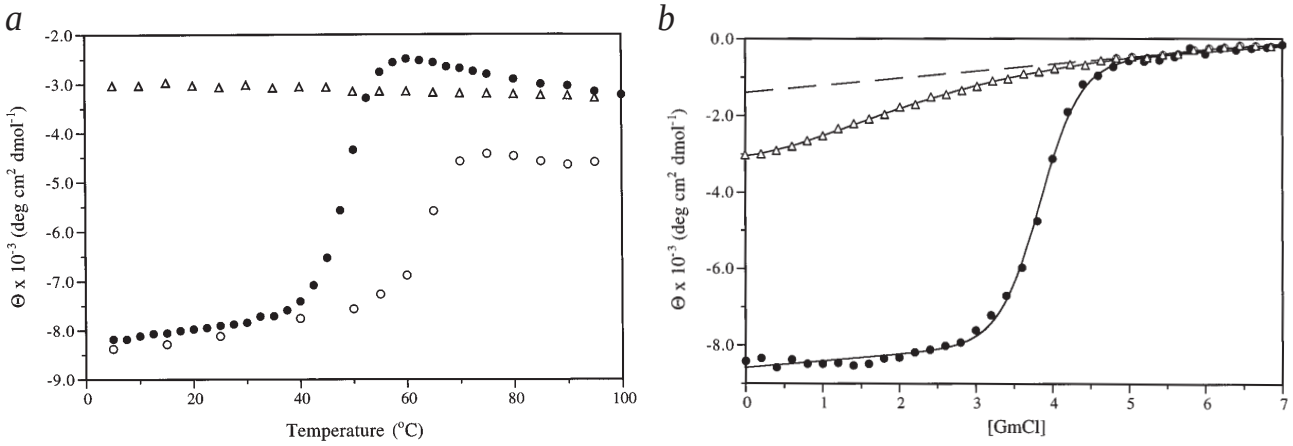

Fig. 1 Melting behavior of native RNase A (closed circles) and the disulfide-broken polypeptide (open triangles). ***a***, Thermal melting at pH 3 and ***b***, denaturant melting at pH 7 (5 °C) are fully reversible. (Thermal melting at pH 7 is not; open circles.) Disulfide-broken RNase A matches the residual CD of the unfolded protein above the melting transition and defines the CD of the unfolded state at more native conditions. It is noteworthy that the CD of the unfolded state at low denaturant is not well predicted by linear extrapolation from high denaturant (dashed line).

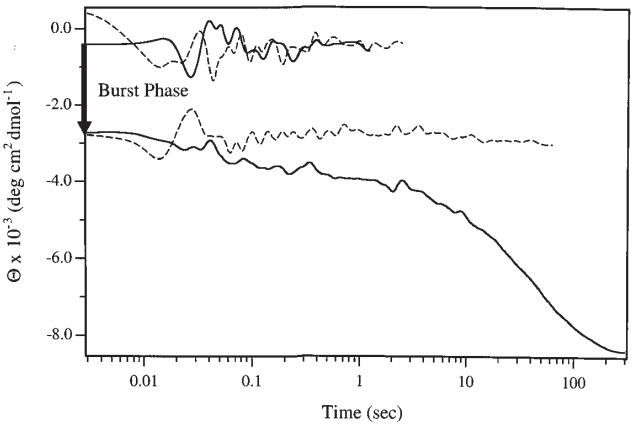

Fig. 2 Kinetic traces of refolding by stopped-flow CD. Disulfide-intact and disulfide-broken RNase A initially at 5.5 M GmCl were diluted into 5.5 M GmCl (top reference traces) or 0.7 M GmCl (lower traces) at pH 7 and 5 °C. The arrow indicates the burst phase. Disulfide-broken RNase A (dashed line) attains its equilibrium value in the burst phase. Intact RNase A (solid line) reaches the same burst phase ellipticity and then more slowly folds to the native state.

cles), with the same dependence on GmCl concentration, and then slowly folds to its final native value (Fig. 1*b*).

Fig. 3 also includes previously published burst phase CD data[5] for intact and disulfide-broken RNase A (diamonds). Houry and Scheraga did note the near equivalence of the protein burst amplitude to the polypeptide equilibrium CD but ascribed it to coincidence and considered that the CD values reached at the two final GmCl concentrations studied (Fig. 3) reflect two different folding intermediates. The quantitative match with the unfolded polypeptide CD (burst phase and equilibrium) over the whole range of GmCl concentrations suggests that this agreement is more than coincidental.

Houry and Scheraga[5] measured hydrogen exchange protection in disulfide-intact RNase A just after the burst phase by pulse labeling. The hydrogen labeling measured for 21 amide NHs was between 80% and 40%. A maximum of ~80% labeling rather than 100% occurred because 23% of the chain reached the native state in the 6 ms dead time prior to the labeling pulse[5]. The spread of labeling to lower values for the various NHs is due to: (i) the competition between folding to the protected native state (26.4 s^{-1}) and labeling rates for unprotected amides (20–830 s^{-1}at pH 9, 5 °C)[7]; (ii) the limited pulse length used (20 ms; 1.3 folding lifetimes); and (iii) protection against labeling in the burst phase forms. Houry and Scheraga assumed that folding proceeds through two burst phase intermediates, detected as in Fig. 3[12] and ref. 13, and calculated sizeable hydrogen exchange protection factors, from 5 to more than 100 at seven of the 21 NHs that could be measured and 1.5–5 at four others. This calculation assumes that the protected kinetic intermediate (I_ϕ) is in rapid equilibrium with an earlier intermediate (I_u) that has no protection and accounts for much of the exchange labeling observed. A recalculation of the same data that assumes only a general burst phase ensemble yields protection factors of 4 and 2.7 for the most protected NHs and 0.8–1.8 for the rest. A straightforward test of these alternatives can be performed with disulfide-broken RNase A at equilibrium (zero GdmCl, 5 °C). The results demonstrate the absence of hydrogen exchange protection, indicating that distinct hydrogen-bonded structure is absent[7]. Finally it can be noted that the difference in proline isomers in these experiments is not expected to affect the largely unstructured forms dealt with here (Fig. 3).

In summary, the burst phase CD signals seen for intact RNase A are quantitatively duplicated by the demonstrably unfolded disulfide-broken polypeptide over the whole range of denaturant concentration. Earlier results[7] show the absence of protection in disulfide-broken RNase A and the exchange labeling data obtained in burst phase experiments with intact RNase A are consistent with this result (recalculated from ref. 12). Similar experiments with cytochrome *c*[2,3] also showed nearly identical burst phase amplitudes (CD and fluorescence quenching) for the intact protein and two non-folding variants and minor HX protection in the collapsed burst phase protein[14]. These results point to the absence of distinct structure formation in the burst phase. The burst phase signals appear to reflect a fast readjustment from the unfolded ensemble characteristic at high denaturant to a more compact but still unfolded ensemble characteristic for low denaturant (~50 μs[4,15]). Similar behavior will undoubtedly contribute to burst phase observations in other proteins.

It is important to distinguish the submillisecond burst phase behavior considered here from the multi-millisecond structure formation sometimes referred to by the same term. For example, specific structure formation during the experimental dead time, documented by strong hydrogen exchange protection in apomyoglobin[16] and RNase H[17], has been referred to operationally as burst phase behavior. In these cases the dead times were 5 ms and 12 ms respectively, thus represent a time scale longer than the burst phase dealt with here by 100-fold. In other tested cases, the submillisecond burst phase produces either no hydrogen exchange protection (ubiquitin[18,19]) or very weak protection, less than 10-fold (cytochrome *c*[14], thermophilic phosphoglycerokinase[20], CD2.D1[21], lysozyme[18]).

Initial events in folding

The submillisecond changes seen in denaturant dilution experiments have often been interpreted in terms of the formation of an on-pathway folding intermediate, as in U' → I → N[1]. This view is tenable in the sense that the condition reached at the end

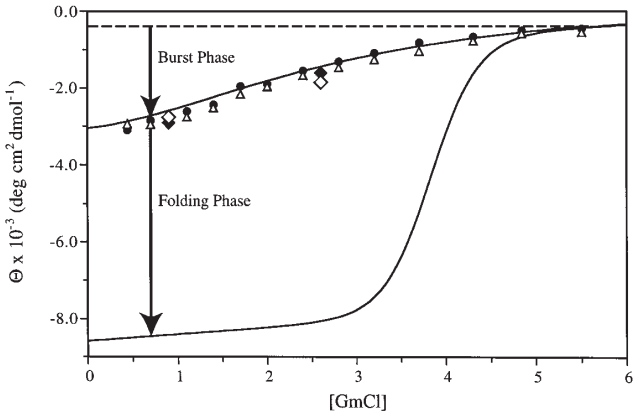

Fig. 3 Dependence of kinetic amplitudes on final GmCl concentration. Stopped-flow experiments start at 5.5 M GmCl (horizontal line). The burst phase (arrows) resets the CD for both intact (circles) and disulfide-broken RNase A (open triangles) to the equilibrium unfolded curve defined in Fig. 1*b* (upper solid line). At longer times intact RNase A folds to its equilibrium native curve (lower solid line, Fig. 1*b*). The diamonds show burst phase CD data for intact and disulfide-broken RNase A (from Houry and Scheraga[5]).

of the burst phase is obligatory and represents an altered and perhaps even a productive structural format. The present results indicate that the I form reached in fact represents the unfolded state characteristic for the ambient conditions used, leading to the more appropriate description of this behavior as U' → U → N. The U' to U transition observed is wholly due to the design of the experiment itself which starts at the high denaturant U' condition.

This issue becomes more than semantic when one wishes to address the specific mechanisms and the time scales that proteins utilize to initiate productive folding. Earlier work showed that the rate of cytochrome *c* folding is ultimately limited by an initial long range conformational search to find a large scale topologically native-like transition state nucleus[2]. It appears that the folding of other normally occurring globular proteins is most often limited by a similar long range search. We suggested that folding can reach submillisecond rates only under two conditions: (i) when finding the initial nucleation transition state depends on a short range search; (ii) when the folding events measured occur in partially structured molecules that have already passed the nucleation process.

The present work was stimulated by numerous reports of submillisecond burst phase behavior that does not meet either of these conditions. The results for RNase A described here and analogous results for cytochrome *c*[2–4] now indicate that the submillisecond burst phase seen for these proteins reflects a solvent-dependent readjustment of the unfolded state ensemble rather than the rapid formation of distinct folding intermediates. These results refocus the issue of initial events in protein folding. Nucleated structure formation that depends on short range searching — in model helices[22] or β structures[23] or even in engineered proteins[24] — can proceed on a time scale much faster than milliseconds. Naturally occurring globular proteins in general fold much more slowly. We continue to suggest[2] that the initial step in globular protein folding is a long range conformational search to find a topologically native-like transition state nucleus (see also ref. 25). The long range nature of the search appears generally to require a millisecond time scale. Intermediate formation and native state acquisition can go more slowly when subsequent misfold-reorganization barriers intervene[26] but they generally do not go much faster.

Methods

Reagents were from Sigma Chemical Company (iodoacetamide, DTT, Tris buffer) and Fisher Scientific Company (HCl, NaOH, potasium phosphate). Ultra pure guanidine hydrochloride was from ICN Biochemicals. RNase A (bovine pancreas, type XII-A) from Sigma was used without further purification. Reduced and carboxamidomethylated RNase A was produced as described by Houry *et al.*[12] and used for equilibrium melting experiments. Kinetic experiments used disulfide-reduced RNase A produced *in situ* by incubation with 30 mM dithiothreitol for over 1 h in the starting 5.5 M GmCl. Protein concentrations were determined using extinction coefficients of 9,800 M^{-1} cm^{-1} (intact RNase A) and 8,160 M^{-1} cm^{-1} (disulfide-broken RNase A) at 275 nm[27].

Equilibrium CD was recorded at 1 nm resolution using 1 mm (50 μM protein) and 1 cm (5 μM protein) pathlength cells in an Aviv 62DS spectrometer. The CD scale was calibrated frequently using 1S-(+)-10-camphor sulfonic acid[28]. Stopped-flow CD experiments averaged up to 35 kinetic traces using a Biologic SFM4 module interfaced with a Jasco model 715 spectrometer set for 1–2 nm resolution with an 0.8 mm observation cell. Kinetic runs were initiated by dilution from 5.5 M GdmCl to a final protein concentration of 35–100 μM. Instrumental dead-time was measured at ~1 ms.

Acknowledgments

This work was supported by NIH research grants to S.W.E. and T.R.S.

Phoebe X. Qi[1], Tobin R. Sosnick[2] and S. Walter Englander[1]

[1]The Johnson Research Foundation, Department of Biochemistry and Biophysics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104-6059, USA. [2]Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, Illinois 60637-1432, USA.

Correspondence should be addressed to P.X.Q. *email: Phoebe@HX2.Med.UPenn.Edu*

Received 2 June, 1998; accepted 4 August, 1998.

1. Roder, H. & Colon, W. *Curr. Opin. Struct. Biol.* **7**, 15–28 (1997).
2. Sosnick, T.R., Mayne, L. & Englander, S.W. *Proteins:* **24**, 413–426 (1996).
3. Sosnick, T.R., Shtilerman, M.D., Mayne, L. & Englander, S.W. *Proc. Natl. Acad. Sci. USA* **94**, 8545–8550 (1997).
4. Chan, C.K. *et al. Proc. Natl. Acad. Sci. USA* **94**, 1779–1784 (1997).
5. Houry, W.A. & Scheraga, H.A. *Biochemistry* **35**, 11734–11746 (1996).
6. Harrington, W.F. & Schellman, J.A. *C. R. Trav. Lab Carlsberg Ser. Chim.* **30**, 21 (1956).
7. Bai, Y., Milne, J.S., Mayne, L. & Englander, S.W. *Proteins* **17**, 75–86 (1993).
8. Privalov, P.L. *et al. J. Mol. Biol.* **205**, 737–750 (1989).
9. Dill, K.A. & Shortle, D. *Annu. Rev. Biochem.* **60**, 795–825 (1991).
10. Krimm, S. & Tiffany, M.L. *Israel J. Chem.* **12**, 189–200 (1974).
11. Nozaki, Y. & Tanford, C. *J. Biol. Chem.* **245**, 1648–1652 (1970).
12. Houry, W.A., Rothwarf, D.M. & Scheraga, H.A. *Biochemistry* **35**, 10125–10133 (1996).
13. Houry, W.A., Rothwarf, D.M. & Scheraga, H.A. *Biochemistry* **33**, 2516–2530 (1994).
14. Sauder, J. M. & Roder, H. *Folding & Design* **3**, 293–301 (1998).
15. Shastry, M.C.R. & Roder, H. *Nature Struct. Biol.* **5**, 385–392 (1998).
16. Jennings, P.A. & Wright, P.E. *Science* **262**, 892–896 (1993).
17. Raschke, T.M. & Marqusee, S. *Nature Struct. Biol.* **4**, 298–304 (1997).
18. Gladwin, S.T. & Evans, P.A. *Folding & Design* **1**, 407–417 (1996).
19. Khorasanizadeh, S., Peters, I.D. & Roder, H. *Nature Struct. Biol.* **3**, 193–205 (1996).
20. Hosszu, L.L. *et al. Nature Struct. Biol.* **4**, 801–804 (1997).
21. Parker, M.J., Dempsey, C.E., Lorch, M. & Clarke, A.R. *Biochemistry* **36**, 13396–13405 (1997).
22. Thompson, P.A., Eaton, W.A. & Hofrichter, J. *Biochemistry* **36**, 9200–9210 (1997).
23. Munoz, V., Thompson, P.A., Hofrichter, J. & Eaton, W.A. *Nature* **390**, 196–199 (1997).
24. Burton, R.E., Huang, G.S., Dougherty, M.A., Fullbright, P.W. & Oas, T.G. *J. Mol. Biol.* **263**, 311–322 (1996).
25. Abkevich, V.I., Gutin, A.M. & Shakhnovich, E.I. *Biochemistry* **33**, 10026–10036 (1994).
26. Sosnick, T.R., Mayne, L., Hiller, R. & Englander, S.W. *Nature Struct. Biol.* **1**, 149–156 (1994).
27. Rothwarf, D.M. & Scheraga, H.A. *Biochemistry* **32**, 2671–2679 (1993).
28. Johnson, W.C. *Annu. Rev. Biophys. Biophys. Chem.* **17**, 145–166 (1988).