Proc. Natl. Acad. Sci. USA
Vol. 94, pp. 8545–8550, August 1997
Biophysics

Ultrafast signals in protein folding and the polypeptide contracted state

T. R. SOSNICK*, M. D. SHTILERMAN, L. MAYNE, AND S. W. ENGLANDER†

Johnson Research Foundation, Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, PA 19104-6059

Contributed by S. W. Englander, June 4, 1997

ABSTRACT To test the significance of ultrafast protein folding signals (≪1 msec), we studied cytochrome *c* (Cyt c) and two Cyt *c* fragments with major C-terminal segments deleted. The fragments remain unfolded under all conditions and so could be used to define the unfolded baselines for protein fluorescence and circular dichroism (CD) as a function of denaturant concentration. When diluted from high to low denaturant in kinetic folding experiments, the fragments readjust to their new baseline values in a "burst phase" within the mixing dead time. The fragment burst phase reflects a contraction of the polypeptide from a more extended unfolded condition at high denaturant to a more contracted unfolded condition in the poorer, low denaturant solvent. Holo Cyt *c* exhibits fluorescence and CD burst phase signals that are essentially identical to the fragment signals over the whole range of final denaturant concentrations, evidently reflecting the same solvent-dependent, relatively nonspecific contraction and not the formation of a specific folding intermediate. The significance of fast folding signals in Cyt *c* and other proteins is discussed in relation to the hypothesis of an initial rate-limiting search-nucleation-collapse step in protein folding [Sosnick, T. R., Mayne, L. & Englander, S. W. (1996) *Proteins Struct. Funct. Genet.* 24, 413–426].

Many recent publications describe submillisecond folding-related signals (1) and interpret these signals in terms of the rapid formation of productive folding intermediates. Our earlier results lead us to question this interpretation. We found that the folding rate of cytochrome *c* (Cyt *c*) is limited by an initial, time-requiring, nucleated polypeptide chain collapse (2). The nucleated collapse represents the rate-limiting step for overall folding when folding is two-state and it limits the rate at which intermediates can populate when folding is more than two-state, i.e., when an even larger misfold-reorganization barrier is encountered in subsequent folding (3). The nucleated collapse itself appears to be limited by a large scale diffusion-dependent conformational search, culminating in the assembly of a transition state nucleus that can support forward folding steps in an energetically downhill manner. For Cyt *c* the initial search-nucleation step requires about 1 msec under the most favorable conditions, and results available for other small proteins when starting from the fully unfolded state seemed to us to be generally consistent with a similar requirement (2). However, this view of the strategy for the initiation of folding now appears to be contradicted by many recent reports of much faster folding events, both in Cyt *c* and other proteins.

This paper considers the significance of ultrafast protein folding signals. The results show that when unfolded Cyt *c* is diluted from high denaturant in folding experiments, it experiences a fast polymer chain contraction from a more extended form of the unfolded state to a more contracted form of the unfolded state. The optical signals produced by the chain contraction can misleadingly mimic the formation of a distinct folding intermediate. Rather, the supposed intermediate is the unfolded state itself.

An assessment of available fast folding observations in this light shows them to be consistent with the view that folding is limited by an obligatory, initial, time-requiring (≈1 msec) search-nucleation step.

MATERIALS AND METHODS

Horse heart Cyt *c* (highest available grade) was from Sigma Chemical Company. Cyt *c* fragments were produced by partially selective cyanogen bromide cleavage (4) at the methionine residues, Met-65 and -80, purified by reverse phase HPLC (C18 column, acetonitrile gradient), and verified by mass spectrometry. F1–65 was virtually pure. The F1–80 preparation was contaminated (≈10%) with a species close to Cyt *c* in mass. Molar concentration of the fragments was measured by absorbance using an extinction coefficient at 401 nm of 137,000 $M^{-1}\cdot cm^{-1}$ in denaturing guanidinium chloride (GdmCl), obtained by comparison with holo Cyt *c*.

Equilibrium CD spectra were recorded at 1 nm resolution using an Aviv 62DS spectrometer in a 1–2 mm pathlength cell. Hydrogen-tritium exchange data were obtained as described (5). Predicted H–T exchange curves (6–8) for the fragments include the information that the NHs of residues 14, 15, and 18 are protected from exchange by a factor greater than 10 due to residual structure in the heme-associated peptide (9, 10). Hydrogen exchange (HX) experiments were performed at pH 3.9 and 0°C where exchange of even freely exposed peptide group NHs is slow enough ($\tau \approx 20$ min) to be accurately measured. The peptide fragments exhibit identical CD spectra at this condition and at the pH 4.9, 10°C condition used for the stopped-flow experiments reported here.

Refolding experiments were initiated by dilution from the denatured state (4.3 M GdmCl, pH 4.9 or pH 2) into native conditions (pH 4.9, 0.15 to 0.3 M sodium acetate, 10°C, variable GdmCl). Protein concentrations were at 25 μM for CD and ranged from 4 to 400 μM for fluorescence. Experiments used a Biologic (Grenoble, France) SFM3 or SFM4 stopped-flow apparatus with 0.8 mm pathlength cuvet. Kinetic fluorescence data were obtained with a spectrometer equipped with a 200 W xenon/mercury arc lamp. Kinetic CD data were obtained using an SFM4 interfaced with a Jasco (Easton, MD) model 715 spectrometer at 1–2 nm resolution and averaged up to 30 kinetic traces. The instrumental dead-time was measured using a dye quenching reaction.

The CD spectrum of F1–80 showed no change from 7 to 50 μM. At 0.7 M GdmCl, the burst phase amplitudes of F1–80 and

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: Cyt *c*, cytochrome *c*; CD, circular dichroism; HX, hydrogen exchange; GdmCl, guanidinium chloride.

*Present address: Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, IL 60637.
†To whom reprint requests should be addressed. e-mail: Walter@HX2.Med.UPenn.edu.

the holoprotein were concentration-independent between 3–44 μM and 4–400 μM, respectively. These results indicate that aggregation does not contribute to the burst phase signals seen.

RESULTS

Fragment Identity. The identity of the fragments, F1–65 and F1–80, was confirmed by matrix-assisted laser desorption ionization mass spectrometry (Fig. 1*A*). F1–65 (Fig. 1*A, Bottom*) was essentially pure; the F1–80 preparation (Fig. 1*A, Middle*) was contaminated with a native-like species. This and other observations (below) confirm the contaminant as a separate species (minor component in melting curves; slow component in HX curves), and estimate its level at≈10%.

F1–80 lacks residues 81–104, which form the C-terminal helix in the native protein and are essential for an early folding intermediate (9, 11). F1–65 additionally lacks residues that form the short 70s helix and the major 60s helix. Both fragments retain the single tryptophan of Cyt *c*, Trp-59. and the heme group, covalently held by Cys-14 and -17, which provide a fluorescence quenching probe that is sensitive to molecular extent. The tryptophan fluorescence is quenched by Förster resonance energy transfer to the heme ($r_o^6/[r^6 + r_o^6]$). Since r_o is 32 Å and r is 32 Å for Cyt *c* and its fragments in high GdmCl (12), the fluorescence changes most sharply between 32 and 20 Å and so is sensitive to the initial stages of molecular contraction.

Fragment Structure by CD. Thermally unfolded Cyt *c* retains half the ellipticity of the native protein (Fig. 2*A*; at 222 nm). Proteins often retain considerable ellipticity in the thermally unfolded state, even though they lack hydrogen exchange protection (e.g., ref. 13). In native proteins, CD is well understood as an indicator of regular secondary structure. In the unfolded state, the distribution of main chain dihedral angles can produce ellipticity that responds to solvent interactions (14–16) and contraction effects (17).

Fig. 2*A* shows that the ellipticity of the Cyt *c* fragments matches the thermally unfolded protein and remains at that level from 0 to 100°C. Fig. 2*B* shows that the fragment CD (room temperature) identically matches the entire CD spectrum of thermally unfolded Cyt *c* and also its sensitivity to GdmCl. By CD criteria, the fragments exist in the unfolded state.

Fragment Structure by HX. The possible presence of H-bonded structure in the fragments at strongly native conditions

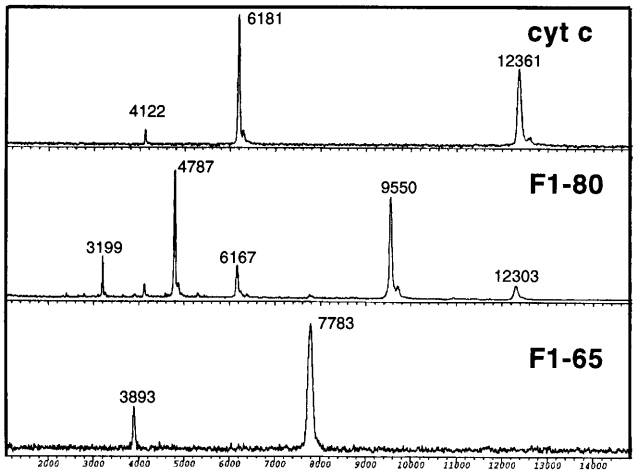

FIG. 1. The fragment preparations. Traces from matrix-assisted laser desorption ionization mass spectrometry showing the holoprotein (*Top*), the F1–80 preparation with a minor native-like contaminant (*Middle*), and the F1–65 preparation (*Bottom*). For each species, z on the m/z axis is 1, 2, and 3 from right to left.

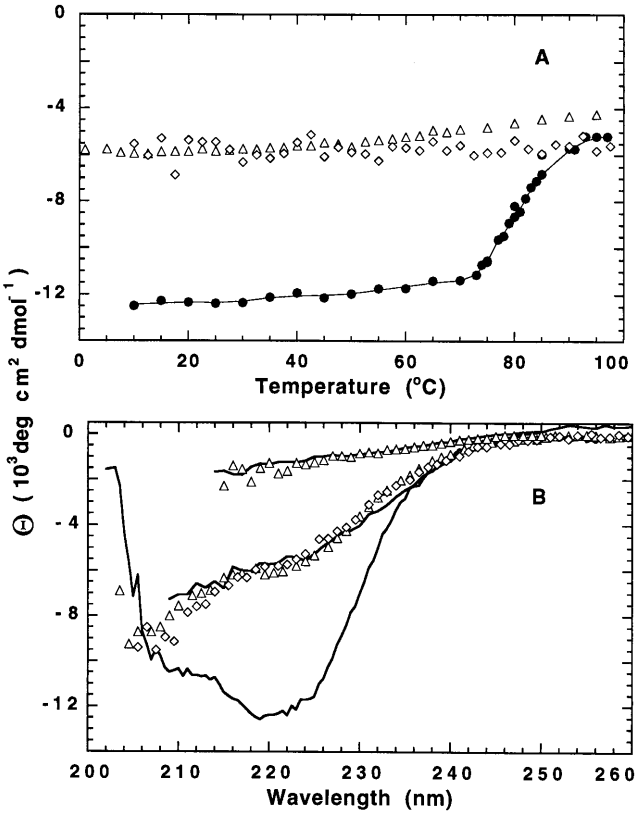

FIG. 2. Ellipticity of the fragments (△, F1–80; ◇, F1–65), and holo Cyt *c* (● and/or —). (*A*) Ellipticity (222 nm) as a function of temperature. (*B*) CD spectra: native Cyt *c* (*Bottom*); the fragments at 22°C and Cyt *c* thermally denatured at 97°C (*Middle*); Cyt *c* and F1–80 in 4.4 M GdmCl (*Top*).

was evaluated by tritium exchange measurements at conditions chosen to maximally stabilize residual structure (0°C, no GdmCl). Tritium exchange can measure the number of hydrogens exchanged as a function of exchange time with an accuracy close to 1% (18). Therefore a comparison of measured exchange with the HX curve expected for the unstructured polypeptide (6–8) can sensitively detect the presence of H-bonded hydrogens. It can be noted that the H-bonded NH hydrogens in native Cyt *c* (79 out of 100 NHs total) are well distributed through the protein (70% in 1–80, 60% in 1–65) (19).

HX data for F1–65 (Fig. 3*A*) match the curve expected for no HX protection. Fig. 3*A* includes predicted curves for some situations that might account for the residual ellipticity observed. The fragment ellipticity in the absence of GdmCl (Fig. 2*A*) corresponds to the signal expected for an alpha-helical content of 20% (20). One possibility is that some specific subset, ≈20% of the peptide NHs present, is protected by helical H-bonding (- - - in Fig. 3*A*). Another possibility is that a helical content of 20% is time-shared over the chain leading to a distributed HX protection factor of 1.2 (dotted curve). The presence of any sizable number of protected NHs in the fragments is ruled out by the measured data, indicating that F1–65 does not contain stably H-bonded regions of structure, even at strongly stabilizing conditions. These data do not rule out the presence of a number of short helical segments, which would protect many fewer NH hydrogens than the residues involved.

For F1–80 the data at longer exchange times detect approximately six slowly exchanging NHs (Fig. 3*B*). This is consistent with the estimate of 10% native-like Cyt *c* contamination in the F1–80 preparation, as shown by the predicted curve for this

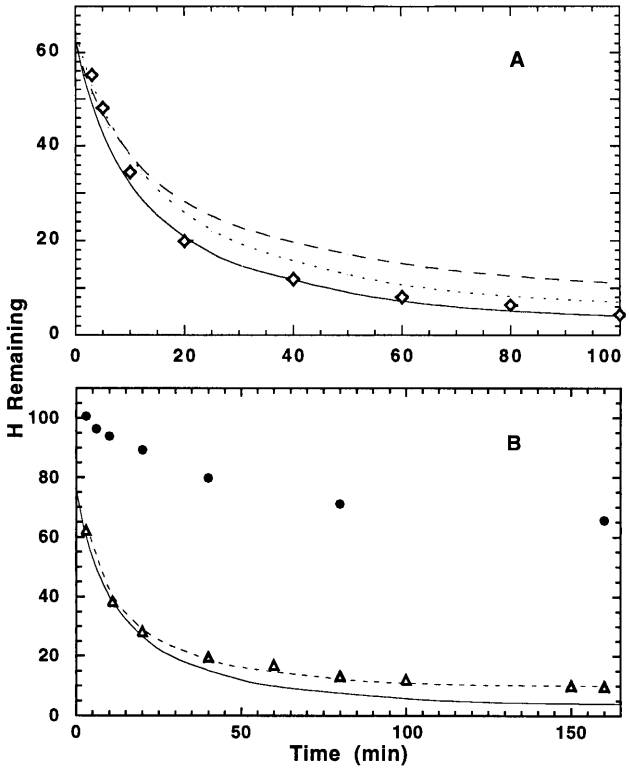

FIG. 3. H-T exchange of the fragments. (*A*) Measured data for F1–65 (◇) and the curve predicted for an unprotected polypeptide with the same amino acid sequence (—). Predicted curves are also shown for the fragment with 20% of its NHs blocked by H-bonding (— — —), and for all the NHs time-sharing 20% of H-bonding (···). (*B*) Measured data for F1–80 (△) compared with the random coil prediction (—) and for a random coil preparation with 10% contamination by Cyt *c* (- - -; 0.9× random coil curve plus 0.1× Cyt *c* curve). HX of Cyt *c* at the same condition is shown (●). See *Materials and Methods* for details.

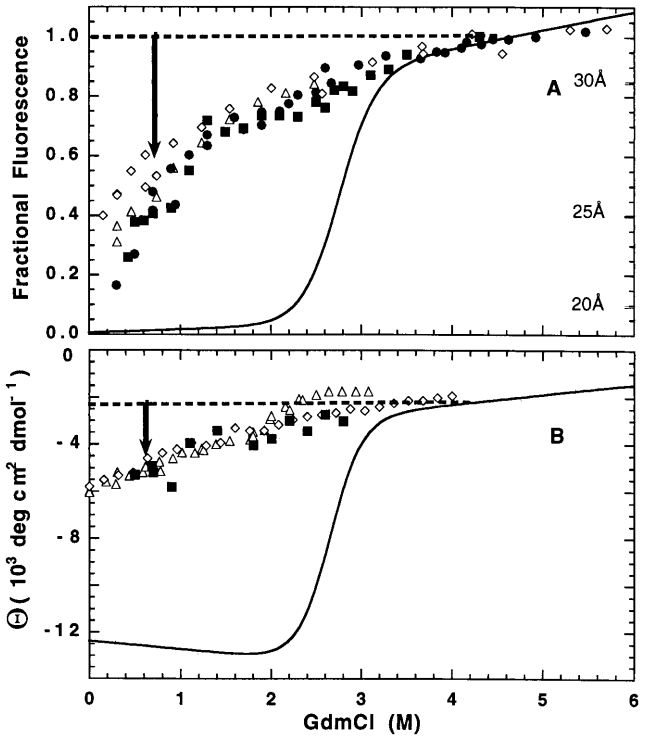

FIG. 4. The unfolded baseline and the Cyt *c* burst phase. The solid curves show the equilibrium behavior of Cyt *c*. The equilibrium fluorescence and CD of the (unfolded) fragments (△ and ◇) match the unfolded holo Cyt *c* baseline at high GdmCl and define the continuation of the unfolded baseline to lower GdmCl concentrations. The horizontal dashed line shows the initial fluorescence and CD in the stopped-flow experiments (4.3 M GdmCl). The solid symbols indicate the fluorescence (*A*) and the ellipticity (*B*; 222 nm) reached by holo Cyt *c* in the burst phase upon dilution into lower (or higher) GdmCl, as suggested by the arrows (compare Fig. 5) [starting from either pH 2 (●) or pH 4.9 (■)]. These comparisons are made on an absolute, per molecule basis. Förster-averaged distance (Trp-59 to heme) is at the right of *A*.

condition (- - -), obtained by proportionately adding the curve measured for holo Cyt *c* at this condition.

In summary, HX demonstrates the absence of discrete H-bonding in the fragments.

Equilibrium Behavior in the Unfolded State. Although fragment conformation is essentially independent of temperature, it varies with GdmCl concentration (Fig. 2). Fig. 4*A* (open symbols) shows the equilibrium fluorescence of the fragments as a function of GdmCl obtained by a kinetic method. When the fragments are diluted from 4.3 M GdmCl into lower GdmCl, their fluorescence falls within the experimental dead time ($\tau < 1$ msec) to a new level that does not change at longer times (see Fig. 5*A*). These experiments define the equilibrium fluorescence for the fragments at each final GdmCl concentration, as plotted in Fig. 4*A*, and also show that this level is reached in the kinetic burst phase. The equilibrium ellipticity of the fragments as a function of GdmCl concentration was obtained by direct static measurements (open symbols in Fig. 4*B*).

Fig. 4 also shows equilibrium curves for the cooperative folding of holo Cyt *c*. In denaturing GdmCl, the fragments match the baseline fluorescence and the baseline ellipticity of unfolded holo Cyt *c*. In decreasing GdmCl, the equilibrium fluorescence and ellipticity of the fragments change in a noncooperative manner. Fragment ellipticity increases faster than the linear extrapolation of the unfolded baseline (Fig. 4*B*). Fluorescence quenching increases even more steeply due to chain contraction in low GdmCl and the sharp dependence of fluorescence quenching on the Trp-59 to heme distance (right axis). (Intrinsic tryptophan fluorescence is unaffected by GdmCl; ref. 21.)

Kinetic Burst Phase Behavior. The dashed lines in Fig. 4 indicate the level of fluorescence and ellipticity at the initial condition (4.3 M GdmCl) of the stopped-flow experiments. When the fragments are diluted into lower (or higher) GdmCl, their fluorescence and ellipticity change in a burst phase to their new equilibrium values (arrows in Fig. 5 *A* and *B*). The burst phase represents a fast readjustment from the initial unfolded baseline to the new unfolded baseline.

The burst phase values reached by holo Cyt *c* are indicated by the filled symbols in Fig. 4 (see also ref. 22). The holo Cyt *c* burst phase readjusts its fluorescence and CD essentially quantitatively to the equilibrium baseline values defined by the fragments over the whole range of final GdmCl concentrations. (These data also show that the measured CD, expressed as mean residue ellipticity, contains no dominant contribution from the heme group or particular residues, which would be apparent in systematic differences between the different chains when divided by 104, 80, and 65.) The holo protein subsequently folds to the native state on a much slower time scale (Fig. 5).

DISCUSSION

Chain Contraction in the Unfolded State. Clearly the fragments used here remain in the unfolded state under all conditions. They lack segments that provide important parts of

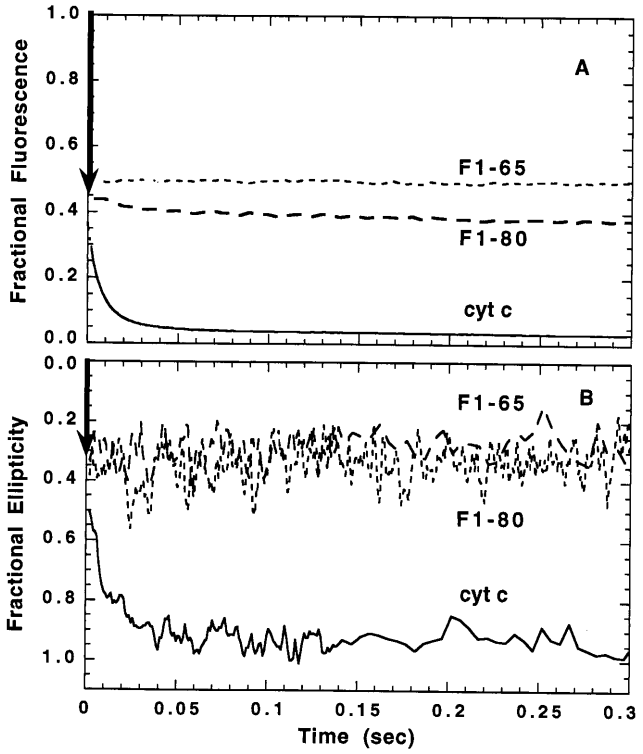

FIG. 5. Kinetic refolding measured by fluorescence (*A*) and ellipticity (*B*; 222 nm). The arrows indicate the fast burst phase change on dilution from 4.3 M GdmCl to 0.7 M GdmCl (pH 4.9, 10°C).

the native Cyt *c* structure including the stable C-terminal helix, a crucial component of an early folding intermediate (9, 11). They identically duplicate the CD spectrum of thermally unfolded Cyt *c*. This fragment CD remains unchanged between 0–100°C. The molecular extent (fluorescence quenching) and CD of the fragments shows the same sensitivity to GdmCl as the unfolded holoprotein. They have no discrete H-bonded structure even at zero GdmCl and low temperature.

The unfolded fragments provide an unusual opportunity to define the unfolded Cyt *c* baseline in the region below the unfolding transition, as shown in Fig. 4. The increase in fluorescence quenching at low denaturant reflects a contraction of the still unfolded polypeptide chain (decrease in Trp-59 to heme distance), perhaps due to the increasingly stable nonspecific association of hydrophobic side chains which are less soluble at low GdmCl (23). This behavior is analogous to the well-known contraction of polymer chains when moved into a poorer solvent. The increase in ellipticity in the contracted chain at low denaturant may represent a secondary effect of the chain compaction (17, 24, 25) or a solvent effect on the distribution of main chain phi and psi angles (14–16).

When the fragments are diluted from high GdmCl in kinetic experiments, they rapidly contract (fluorescence quenching) and reach their new unfolded baseline values in the stopped-flow dead time (Fig. 5; $\tau < 1$ msec). The holoprotein when similarly diluted also acquires new fluorescence and CD values in the dead time (Fig. 5). These values duplicate the baseline measured for the unfolded fragments over the whole range of GdmCl concentrations (Fig. 4).

It has been suggested that the burst phase fluorescence (26) and CD (27) signals of Cyt *c* reflect the fast formation of a productive folding intermediate that is in rapid denaturant-dependent equilibrium with the unfolded form. If so, then the Cyt *c* fragments studied here must form a corresponding intermediate with the same equilibrium constant and the same dependence on GdmCl. However the present results demonstrate that the fragments remain in the unfolded state. The same conclusion must hold for the holoprotein. The burst phase signals for Cyt *c*, like the fragments, reflect a rapid readjustment of the unfolded polypeptide chain from the random coil in high GdmCl, a good solvent, to a more contracted condition in the poorer, low GdmCl solvent. A similar conclusion was suggested before (ref. 2; see also ref. 28).

Burst Phase and Chain Contraction in Other Proteins. Results reported for other proteins suggest that this interpretation may be fairly general. Cold shock protein B when diluted from denaturing GdmCl resets its fluorescence in a burst phase to the value obtained by extrapolation of its sharply sloping denatured baseline (29). When disulfide-intact RNase A in denaturing GdmCl is diluted, its CD increases in a burst phase to equal the CD of the structureless disulfide-cleaved protein (see Table 2 in Houry *et al.*, ref. 30). Further, HX pulse labeling immediately after the RNase A burst phase produces a labeling pattern (40–80% labeling) that is as expected for the absence of protecting H-bonding (figure 3 in Houry and Sheraga, ref. 31). The low HX pulse intensity used (31) makes the result sensitive to even low levels of HX protection. Similarly, a fast variant of the HX pulse labeling experiment demonstrates the absence of structural protection in ubiquitin (32) after the ubiquitin burst phase (33) is completed. Thus it appears that the supposed burst phase intermediates for RNase A and ubiquitin in fact reflect the contracted unfolded state characteristic at low denaturant.

The contracted condition of polypeptide chains in zero denaturant can similarly account for the compact unstructured "pre-molten globule state" of Uversky and Ptitsyn (34), and the "molten coil state" of Ferrer *et al.* (35), and may well contribute to measurements of apparent structure in unfolded proteins (36). Chain contraction also provides a promising explanation for the submillisecond burst phase behavior reported for barstar (see ref. 37) and lysozyme (38–42).

In summary, available information suggests that many observations in the protein folding literature may find an explanation in terms of the polymer-like chain contraction characteristic for unfolded polypeptides in the relatively poor aqueous solvent (low denaturant).

Two Kinds of Chain Condensation. It is important to distinguish two different kinds of chain condensation—the fast, energetically downhill, relatively nonspecific chain contraction in the unfolded state, documented here, and the time-requiring, energetically uphill, relatively native-like nucleated collapse suggested before to initiate protein folding (2).

A computer simulation due to Gutin *et al.* (43) captures this distinction. When simulated interactions are strong, as in our low denaturant condition, a fast formation of non-native interactions is seen, resembling the nonspecific energetically downhill hydrophobic zipper behavior of Chan and Dill (44, 45). Folding to the native form occurs later. When the simulated interactions are set to be weaker, as in higher denaturant, the fast random association stage is suppressed and folding follows a slower all-or-none transition to the native form.

Similarly, unfolded Cyt *c* and other unfolded proteins experience a fast, relatively nonspecific compaction on dilution from high into low denaturant, consistent with strong hydrophobic interactions (23, 46). As GdmCl in the refolding buffer is increased, hydrophobic interactions are weakened and polymer-like contraction is suppressed. The burst phase amplitude decreases and disappears, presumably because random association interactions can no longer overwhelm unfavorable loop closure entropy and the uncompensated burial of polar groups (2).

In moderate denaturant where nonspecific collapse is suppressed, a condensation step that encompasses the rate-limiting process for folding can be clearly isolated and characterized (2). This more specific collapse process occurs at the

very beginning of the folding sequence. The process is energetically uphill, buries a great deal of surface, and generally requires ≈1–10 msec or longer. We inferred that the time-requiring process involves a conformational search for some relatively specific transition state that is able to support molecular collapse (2). This is by definition a nucleation process. Other workers have also found evidence for a nucleation process, both on experimental grounds (47, 48) and from computer simulations with simple models (49–51) (reviewed in ref. 52). The transition state for nucleated collapse is able to support subsequent forward folding steps in an energetically downhill manner. It therefore seems likely to resemble the native state topology (2). Since the time-requiring nucleated collapse depends on an initial diffusion-dependent conformational search, it appears to represent an irreducible rate-limiting step for productive folding, although this rate is subject to some manipulation (see below).

It is interesting that the nonspecific collapse and the so-called search-nucleation-collapse may interact. At low denaturant where the chain relaxes into a condensed form, the dependence of folding rate on denaturant (*m* value) often decreases and the rate becomes slower than expected, producing a rollover in the rate vs. denaturant curve. This can be seen as a consequence of the prior formation of the randomly condensed state, from both thermodynamic and kinetic points of view.

Ultrafast Folding. A number of folding experiments have detected ultrafast folding signals that do not represent denaturant-dependent chain collapse in the unfolded state.

Photo-initiated experiments with Cyt *c* (53–56) and fast-mixing continuous flow experiments (57, 58) have detected optical signals with a ≈50 μsec time constant. These signals represent heme ligation reactions and not conformational folding events (54, 57).

Phillips *et al.* (59) used a laser temperature jump to heat solutions of RNase A and observed changes in infrared absorbance after 3.5 nsec of unfolding, pointing to a similar refolding rate (since $K_{unf} \approx 1$). The signals observed are much faster than the authentic 2-state folding of RNase A (msec; ref. 60) and are believed to reflect fraying and hydration events in the already structured protein rather than events that access the unfolded state (R. M. Hochstrasser, personal communication).

Ballew *et al.* (61, 62) used a laser temperature jump (−5°C to +15°C) to promote the folding of a small fraction of cold denatured apomyoglobin and observed fluorescence changes with 250 nsec and 3.5 μsec time constants. It seems possible that these observations, resulting from a large temperature increase, may register the expected temperature-dependent promotion of the nonspecific, hydrophobic contraction in the still unfolded state (see ref. 33 for a precedent). Alternatively, as for RNase A (59), these signals may reflect temperature-dependent changes in already structured molecules, either in the large "native" fraction or in the still partially helical "denatured" fraction. Cold denatured apomyoglobin at −5°C exhibits high ellipticity, ≈70% of the native value at 222 nm (61). The persistence of helical structure in cold denatured proteins appears to be widespread (63–65). Molecules that start with preformed structure, past the initial nucleation step, can be expected to fold much faster than the millisecond nucleation rate (2).

Nölting *et al.* (66, 67) used a laser temperature jump (2–10°C) to fold a small fraction of cold denatured barstar, and measured a 0.3 msec fluorescence phase followed by much slower folding to the native state. These observations may provide an example of folding at the upper limit of rates that one might expect, but are also subject to the questions just raised. Cold denatured barstar contains some marginally stable native-like helix (65) and mutations that most affect the temperature-induced folding (high phi-analysis values) are in those segments (66).

Finally, it is noteworthy that helix nucleation, which depends on short range interactions, can be very fast. Williams *et al.* (68) and Eaton *et al.* (45) find a time constant about 100 nsec for structure formation in a short, independently stable alanine-based helix. However, proteins typically nucleate on a much slower time scale. This suggests that fast helix nucleation does not generally determine protein folding rates, either because naturally occurring helices are usually not independently stable or because they are not exploited even when they are reasonably stable. The latter case is observed for the S-peptide in RNase A (69), the E-helix in RNase H (70), and the fully helical coiled-coil of GCN4 (71). The same observations suggest caution in attributing a necessary early folding role to residual structure observed in the unfolded state.

How Fast Is Fast Enough? We previously considered (2) that amino acid encounters, which represent the primary step in any long range conformational search, are likely to require about 1 usec, up or down by a factor of 10, estimated as ($\approx 10^8$ $M^{-1}\cdot s^{-1}$) × (≈10 mM). This assumes 10^8 $M^{-1}\cdot s^{-1}$ as a reasonable value for encounter of non-neighboring amino acids in a polypeptide chain (compare 10^{10} $M^{-1}\cdot s^{-1}$ for small molecules in free diffusion) and ≈10 mM for the relative concentration of residues in a polypeptide (72). Hagen *et al.* (73) present a sophisticated calculation that reaches a similar result. Polymeric searching and collapse based on short range interactions can go much faster (e.g., helix nucleation, as above) but there appears to be no evolutionary pressure to do so, apparently because the >1 msec time scale necessary for a long-range search-limited folding is fast enough for biological purposes.

It may well be possible to engineer proteins to fold even more rapidly. The 80-residue fragment of lambda repressor (residues 6–85) may be such a case (74, 76). The fragment folds at $(0.2 \text{ msec})^{-1}$ at 37°C. Engineered variants are able to fold faster, perhaps as fast as $(10\ \mu\text{sec})^{-1}$, although this rate represents an extrapolation from high denaturant concentration where the fastest rate so far measured is as shown $(0.2 \text{ msec})^{-1}$.

Conclusions. The experiments and analysis described here suggest that known ultrafast folding events (≪1 msec) represent either a fast solvent response prior to nucleation or a postnucleation event in the already partially structured protein. It will undoubtedly be possible to press the apparent millisecond limit by manipulating folding conditions. The time required can be shortened by increased temperature, decreased molecular size, and the strategic placement of hydrophobic residues that can speed the conformational search, perhaps by promoting short range interactions (44, 51, 75–78). However it continues to appear that the folding of naturally occurring small proteins typically requires a time scale in the range of 1 msec and often more. The important point is that this rate limitation stems from a large scale energetically uphill search-nucleation-collapse process that proteins use to initiate their folding sequence (2).

We thank J. J. Englander for assistance in the tritium exchange experiments and valuable advice on the manuscript; D. Yphantis and J. Larry for conducting the ultracentrifugation experiments; Y. Xu for assistance in their interpretation; G. D. Rose for discussion on the unfolded state; R. W. Woody and E. Stellwagen for discussion on CD in the unfolded state; and P. Speros of Jasco, Inc. for assistance with kinetic CD measurements. Mass spectrometry was provided by W. T. Moore of the Protein Chemistry Laboratory, University of Pennsylvania, supported by core grants of the Diabetes and Cancer Centers (National Institutes of Health Grants DK19525 and CA16520 to J. D. Lambris). This work was supported by National Institutes of Health Research Grants GM31847 (S.W.E.), GM54838 (T.R.S.), and a structural biology training grant (GM08275) (M.S.).

1. Roder, H. & Colon, W. (1997) *Curr. Opin. Struct. Biol.* **7,** 15–28.
2. Sosnick, T. R., Mayne, L. & Englander, S. W. (1996) *Proteins Struct. Funct. Genet.* **24,** 413–426.
3. Sosnick, T. R., Mayne, L., Hiller, R. & Englander, S. W. (1994) *Nat. Struct. Biol.* **1,** 149–156.
4. Gross, E. (1967) *Methods Enzymol.* **11,** 238–255.
5. Englander, S. W. & Englander, J. J. (1972) *Methods Enzymol.* 406–413.
6. Molday, R. S., Englander, S. W. & Kallen, R. G. (1972) *Biochemistry* **11,** 150–158.
7. Bai, Y., Milne, J. S., Mayne, L. & Englander, S. W. (1993) *Proteins Struct. Funct. Genet.* **17,** 75–86.
8. Connelly, G. P., Bai, Y., Jeng, M.-F., Mayne, L. & Englander, S. W. (1993) *Proteins Struct. Funct. Genet.* **17,** 87–92.
9. Bai, Y., Sosnick, T. R., Mayne, L. & Englander, S. W. (1995) *Science* **269,** 192–197.
10. Elöve, G. A., Bhuyan, A. K. & Roder, H. (1994) *Biochemistry* **33,** 6925–6935.
11. Roder, H., Elöve, G. A. & Englander, S. W. (1988) *Nature (London)* **335,** 700–704.
12. Jeng, M.-F. & Englander, S. W. (1991) *J. Mol. Biol.* **221,** 1045–1061.
13. Robertson, A. D. & Baldwin, R. L. (1991) *Biochemistry* **30,** 9907–9914.
14. Krimm, S. & Tiffany, M. L. (1974) *Isr. J. Chem.* **12,** 189–200.
15. Wilson, G., Hecht, L. & Barron, L. D. (1996) *Biochemistry* **35,** 12518–12525.
16. Park, S.-H., Shalongo, W. & Stellwagen, E. (1997), in press.
17. Chan, H. S. & Dill, K. A. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 6388–6392.
18. Englander, S. W. & Englander, J. J. (1978) *Methods Enzymol.* **49,** 24–39.
19. Bushnell, G. W., Louie, G. V. & Brayer, G. D. (1990) *J. Mol. Biol.* **213,** 585–595.
20. Gans, P. J., Lyu, P. C., Manning, M. C., Woody, R. W. & Kallenbach, N. R. (1991) *Biopolymers* **31,** 1605–1614.
21. Tsong, T. Y. (1975) *Biochemistry* **14,** 1542–1547.
22. Chan, C. K., Hofrichter, J. & Eaton, W. A. (1996) *Science* **274,** 628–629.
23. Nozaki, Y. & Tanford, C. (1970) *J. Biol. Chem.* **245,** 1648–1652.
24. Hao, M.-H., Pincus, M. R., Rackovsky, S. & Scheraga, H. A. (1993) *Biochemistry* **32,** 9614–9631.
25. Gregoret, L. M. & Cohen, F. E. (1991) *J. Mol. Biol.* **219,** 109–122.
26. Colon, W. & Roder, H. (1996) *Nat. Struct. Biol.* **3,** 1019–1025.
27. Elöve, G. A., Chaffotte, A. F., Roder, H. & Goldberg, M. E. (1992) *Biochemistry* **31,** 6876–6883.
28. Tanford, C. (1968) *Adv. Protein Chem.* **23,** 121–282.
29. Schindler, T., Herrler, M., Marahiel, M. A. & Schmid, F. X. (1995) *Nat. Struct. Biol.* **2,** 663–673.
30. Houry, W. A., Rothwarf, D. M. & Scheraga, H. A. (1996) *Biochemistry* **35,** 10125–10133.
31. Houry, W. A. & Scheraga, H. A. (1996) *Biochemistry* **35,** 11734–11746.
32. Gladwin, S. T. & Evans, P. A. (1996) *Folding Design* **1,** 407–417.
33. Khorasanizadeh, S., Peters, I. D., Butt, T. R. & Roder, H. (1993) *Biochemistry* **32,** 7054–7063.
34. Uversky, V. N. & Ptitsyn, O. B. (1996) *J. Mol. Biol.* **255,** 215–28.
35. Ferrer, M., Barany, G. & Woodward, C. (1995) *Nat. Struct. Biol.* **2,** 211–7.
36. Smith, L. J., Fiebig, K. M., Scwhalbe, H. & Dobson, C. M. (1996) *Folding Design* **1,** 95–107.
37. Agashe, V. R., Shastry, M. C. & Udgaonkar, J. B. (1995) *Nature (London)* **377,** 754–757.
38. Kiefhaber, T. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 9029–9033.
39. Chaffotte, A. F., Guillou, Y. & Goldberg, M. E. (1992) *Biochemistry* **31,** 9694–9703.
40. Goldberg, M. E. & Guillou, Y. (1994) *Protein Sci.* **3,** 883–887.
41. Denton, M. E., Rothwarf, D. M. & Scheraga, H. A. (1994) *Biochemistry* **33,** 11225–11236.
42. Itzhaki, L. S., Evans, P. A., Dobson, C. M. & Radford, S. E. (1994) *Biochemistry* **33,** 5212–5220.
43. Gutin, A. M., Abkevich, V. I. & Shakhnovich, E. I. (1995) *Biochemistry* **34,** 3066–3076.
44. Dill, K. A., Fiebig, K., M. & Chan, H. S. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 1942–1946.
45. Eaton, W. A., Munoz, V., Thompson, P. A., Chan, C.-K. & Hofrichter, J. (1997) *Curr. Opin. Struct. Biol.* **7,** 10–14.
46. Gutin, A. M., Abkevich, V. I. & Shakhnovich, E. I. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 1282–1286.
47. Fersht, A. R. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 10869–10873.
48. Itzhaki, L. S., Otzen, D. E. & Fersht, A. R. (1995) *J. Mol. Biol.* **254,** 260–288.
49. Abkevich, V. I., Gutin, A. M. & Shakhnovich, E. I. (1994) *Biochemistry* **33,** 10026–10036.
50. Thirumalai, D. & Guo, Z. (1995) *Biopolymers* **35,** 137–140.
51. Shakhnovich, E., Abkevich, V. & Ptitsyn, O. (1996) *Nature (London)* **379,** 96–98.
52. Fersht, A. R. (1997) *Curr. Opin. Struct. Biol.* **7,** 3–9.
53. Jones, C. M., Henry, E. R., Hu, Y., Chan, C. K., Luck, S. D., Bhuyan, A., Roder, H., Hofrichter, J. & Eaton, W. A. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 11860–11864.
54. Chan, C.-K., Hofrichter, J. & Eaton, W. A. (1996) *Science* **274,** 628–630.
55. Pascher, T., Chesick, J. P., Winkler, J. R. & Gray, H. B. (1996) *Science* **271,** 1558–1560.
56. Mines, G. A., Pascher, T., Lee, S. G., Winkler, J. R. & Gray, H. B. (1996) *Chem. Biol.* **3,** 491–497.
57. Yeh, S.-R., Takahashi, S., Fan, B. & Rousseau, D. L. (1997) *Nat. Struct. Biol.* **4,** 51–56.
58. Chan, C.-K., Hu, Y., Takahashi, S., Rousseau, D. L., Eaton, W. A. & Hofrichter, J. (1997) *Proc. Natl. Acad. Sci. USA*, **94,** 1779–1784.
59. Phillips, C. M., Mizutani, Y. & Hochstrasser, R. M. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 7292–7296.
60. Houry, W. A., Rothwarf, D. M. & Scheraga, H. A. (1995) *Nat. Struct. Biol.* **2,** 495–503.
61. Ballew, R. M., Sabelko, J. & Gruebele, M. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 5759–5764.
62. Ballew, R. M., Sabelko, J. & Gruebele, M. (1996) *Nat. Struct. Biol.* **3,** 923–926.
63. Nash, D., Lee, B. S. & Jonas, J. (1996) *Biochim. Biophys. Acta* **1297,** 40–48.
64. Griko, Y. V. & Kutyshenko, V. P. (1994) *Biophys. J.* **67,** 356–363.
65. Wong, K. B., Freund, S. M. & Fersht, A. R. (1996) *J. Mol. Biol.* **259,** 805–818.
66. Nölting, B., Golbik, R., Neira, J. L., Soler-Gonzalez, A. S., Schreiber, G. & Fersht, A. R. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 826–830.
67. Nölting, B., Golbik, R. & Fersht, A. R. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 10668–10672.
68. Williams, S., Causgrove, T. P., Gilmanshin, R., Fang, K. S., Callender, R. H., Woodruff, W. H. & Dyer, R. B. (1996) *Biochemistry* **35,** 691–697.
69. Udgaonkar, J. B. & Baldwin, R. L. (1995) *Biochemistry* **34,** 4088–4096.
70. Chamberlain, A. K., Handel, T. M. & Marqusee, S. (1996) *Nat. Struct. Biol.* **3,** 782–787.
71. Sosnick, T. R., Jackson, S., Wilk, R. M., Englander, S. W. & DeGrado, W. F. (1996) *Proteins Struct. Funct. Genet.* **24,** 427–432.
72. Peng, Z. Y., Wu, L. C. & Kim, P. S. (1995) *Biochemistry* **34,** 3248–3252.
73. Hagen, S. J., Hofrichter, J., Szabo, A. & Eaton, W. A. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 11615–11617.
74. Huang, G. S. & Oas, T. G. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 6878–6882.
75. Otzen, D. E., Itzhaki, L. S., elMasry, N. F., Jackson, S. E. & Fersht, A. R. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 10422–10425.
76. Milla, M. E., Brown, B. M., Waldburger, C. D. & Sauer, R. T. (1995) *Biochemistry* **34,** 13914–13919.
77. Colon, W., Elöve, G., Wakem, L. P., Sherman, F. & Roder, H. (1996) *Biochemistry* **35,** 5538–5549.
78. Burton, R. E., Huang, G. S., Dougherty, M. A., Fullbright, P. W. & Oas, T. G. (1996) *J. Mol. Biol.* **263,** 311–322.