PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Does the Molten Globule have a Native-Like Tertiary Fold?

Zheng-Yu Peng, Lawren C. Wu, Brenda A. Schulman and Peter S. Kim

Phil. Trans. R. Soc. Lond. B 1995 **348**, 43-47
doi: 10.1098/rstb.1995.0044

Does the molten globule have a native-like tertiary fold?

ZHENG-YU PENG, LAWREN C. WU, BRENDA A. SCHULMAN AND PETER S. KIM

Howard Hughes Medical Institute, Whitehead Institute for Biomedical Research, Department of Biology, Massachusetts Institute of Technology, Nine Cambridge Center, Cambridge, Massachusetts 02142 U.S.A.

SUMMARY

One of the mysteries in protein folding is how folding intermediates direct a protein to its unique final structure. To address this question, we have studied the molten globule formed by the α-helical domain of α-lactalbumin (α-LA) and demonstrated that it has a native-like tertiary fold, even in the absence of rigid, extensive side chain packing. These studies suggest that the role of molten globule intermediates in protein folding is to maintain an approximate native backbone topology while still allowing minor structural rearrangements to occur.

A central issue in protein folding is to understand how a protein can fold quickly and efficiently to a unique native structure, despite the immense number of conformations accessible to the unfolded polypeptide (Levinthal 1968). This dilemma suggests that protein folding follows specific pathways, since an exhaustive search of all conformations is not possible on a physiological timescale. Thus, in order to understand protein folding, it is crucial to characterize the intermediates on these pathways.

Molten globules are partially folded forms of proteins proposed to be general intermediates in protein folding (Ptitsyn *et al.* 1990). Molten globules are characterized (table 1) by near-native levels of secondary structure but very little rigid, specific tertiary packing (for reviews, see Ptitsyn 1987, 1992; Kuwajima 1989; Christensen & Pain 1991). Equilibrium molten globules have spectroscopic properties and stabilities similar to that of early kinetic folding intermediates (Kuwajima *et al.* 1985; Ikeguchi *et al.* 1986; Jennings & Wright 1993). Yet, despite numerous studies, no high resolution structure of a molten globule is known, and the significance of molten globules to protein folding remains unclear.

The backbone topology of the molten globule (i.e., the relative orientations of secondary structure elements to one another) is a key unresolved issue. It has been proposed that molten globules correspond to either (i) non-specific collapsed polypeptides or (ii) expanded native-like proteins (figure 1). Answering this question is of vital importance for understanding the role of molten globules in protein folding. If molten globules are non-specific collapsed polypeptides, then they would not contain much information for protein folding. On the other hand, if molten globules have a native-like topology, then they provide an approximate solution to the folding problem in which substantial information transfer has occurred.

Table 1. *Common characteristics of molten globules*

substantial level of secondary structure, often comparable with that of the native protein
absence of well-defined side-chain packing
lack of a cooperative thermal unfolding transition
compactness

α-Lactalbumin (α-LA) forms the best studied molten globule (Kuwajima *et al.* 1976, 1985; Nozaka *et al.* 1978; Dolgikh *et al.* 1981, 1985; Ikeguchi *et al.* 1986; Baum *et al.* 1989; Xie *et al.* 1991; Ewbank & Creighton 1991, 1993*a*, *b*; Alexandrescu *et al.* 1993; Creighton & Ewbank 1994; Peng & Kim 1994; Wu *et al.* 1995). The α-LA molten globule can be obtained under a variety of conditions, including low pH (the A-state) and in the presence of low concentrations of denaturants.

α-LA is a two-domain protein (figure 2). The helical domain consists of residues 1–37 and 85–123 and contains all four α-helices in α-LA. The β-sheet domain consists of residues 38–84 and contains a small antiparallel β-sheet and several irregular structures. α-LA has four disulphide bonds. Two of the disulphide bonds (6–120 and 28–111) are in the α-helical domain, one (61–77) is in the β-sheet domain, and one (73–91) connects the two domains.

We decided to study the molten globule of α-LA by 'protein dissection,' removing parts of the protein molecule deemed extraneous (Peng & Kim 1994). The resulting molecule is called α-Domain and consists exclusively of residues from the helical domain of α-LA (figure 2; for simplicity, we use the same numbering system as used for α-LA). We focused on the α-helical domain for several reasons. First, the helical content of the α-LA molten globule is similar to that of native α-

Phil. Trans. R. Soc. Lond. B (1995) **348**, 43–47
Printed in Great Britain

LA, as judged by circular dichroism (CD) studies (Kuwajima *et al.* 1976; Nozaka *et al.* 1978; Dolgikh *et al.* 1981, 1985). Second, nuclear magnetic resonance (NMR) studies show that some of the native helices, but little of the β-sheet domain, are structured in the molten globule of α-LA (Baum *et al.* 1989; Alexandrescu *et al.* 1993; Chyan *et al.* 1993). Finally, the helical domain of hen egg-white lysozyme, a protein that is structurally homologous to α-LA, has been shown to fold prior to the β-sheet domain (Miranker *et al.* 1991, 1993; Radford *et al.* 1992).

α-Domain (figure 2) with the native disulphide pairings between residues 28–111 and 6–120 (called α-Domain^{ox}) exhibits the characteristics of a molten globule and is a good model system for molten globule studies (Peng & Kim 1994). α-Domain^{ox}is a compact monomer at low concentrations (pH 8.5, no additional salt), contains substantial helical secondary structure, and lacks rigid side-chain packing. α-Domain^{ox} and native α-LA have approximately the same number of residues in an α-helical conformation, as measured by far-UV CD spectroscopy. The near-UV CD and one-dimensional proton NMR spectra of α-Domain^{ox} are very similar to that of the A-state molten globule of α-LA. α-Domain^{ox} also exhibits a non-cooperative thermal transition similar to that of the molten globule of intact α-LA.

To investigate the backbone topology of α-Domain, we performed equilibrium disulphide exchange experiments (Peng & Kim 1994). The cysteines in α-Domain are located in disparate parts of the molecule. Thus, the relative populations of disulphide-bonded species reflect the probability of forming different backbone topologies. If disulphide bond formation is random, then the relative populations can be calculated based on a random-walk model (Kauzmann 1959), which predicts that only 7% of the fully oxidized molecules should have the native disulphide bonds. Strikingly, we observe that ~90% of the fully oxidized molecules have native disulphide bonds (i.e. α-Domain^{ox}) under native conditions (figure 3). In contrast, under denaturing conditions, ~8% of the molecules were found to have the native disulphide bonds, in agreement with the value predicted by the random walk

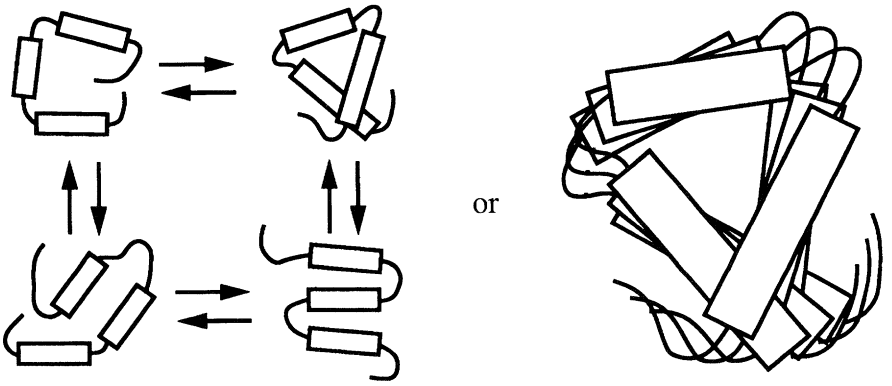

Figure 1. A central question of molten globule structure and its role in protein folding is whether the molten globule has a native-like tertiary fold.

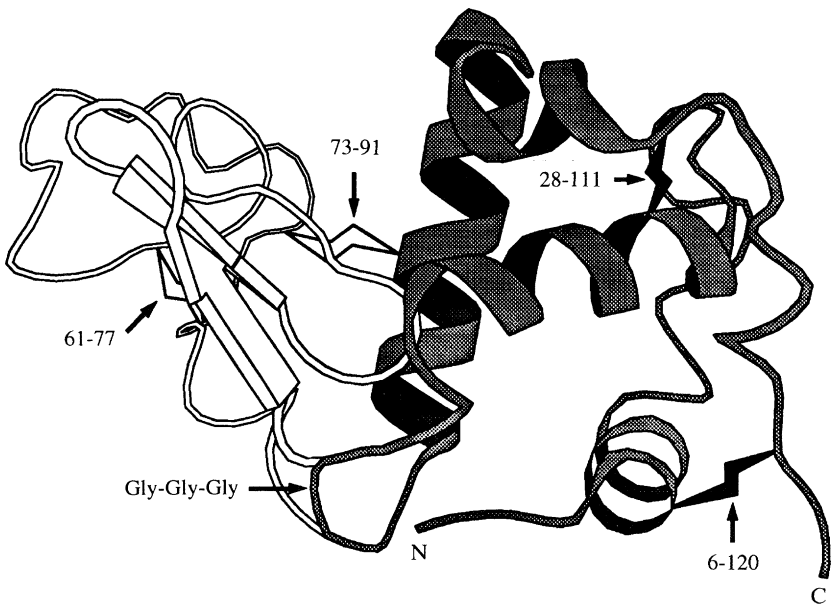

Figure 2. Schematic representation of α-LA. The recombinant α-Domain (shaded) consists of residues 1–39 and 81–123 of human α-LA connected by a short linker of three glycines and preceded by an N-terminal methionine. The two disulphide bonds in α-Domain (6–120 and 28–111) are shown in black. Cys 91, which forms an inter-domain disulphide bond in α-LA, has been changed to alanine to avoid unwanted thiol-disulphide reactions.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

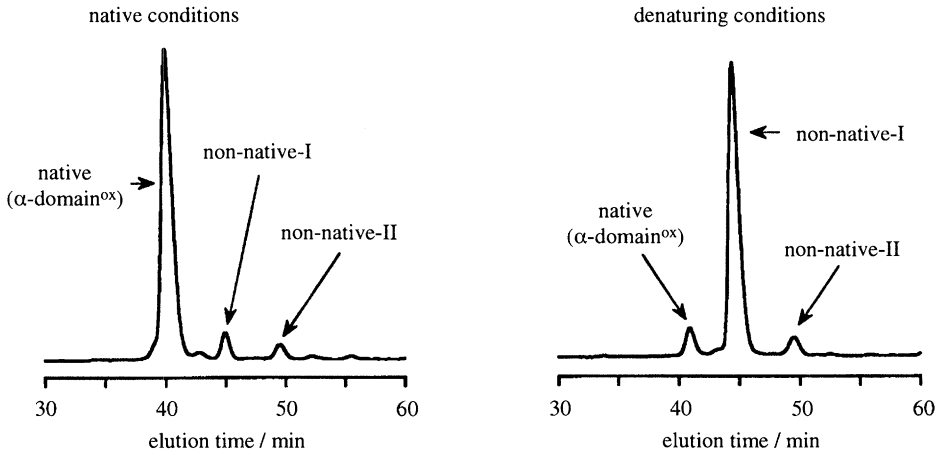

Figure 3. HPLC analysis of the disulphide exchange in α-Domain at pH 8.5 (Peng & Kim 1994). The 'native' isomer (also called α-Domain^{ox}) contains disulphide bonds 28–111 and 6–120. The 'non-native-I' isomer contains disulphide bonds 6–28 and 111–120, and the 'non-native-II' isomer contains disulphide bonds 6–111 and 28–120. Under native conditions, the ratio of native : non-native-I : non-native-II is 90:6:4. Under denaturing conditions (6 M GuHCl), the ratio of native : non-native-I : non-native-II is 8:85:7. For comparison, the ratio predicted for a random walk model is 7:88:5, where the probability of forming an intramolecular disulphide bond is proportional to $n^{-3/2}$ and $n-1$ is the number of intervening non-cysteine residues in the loop.

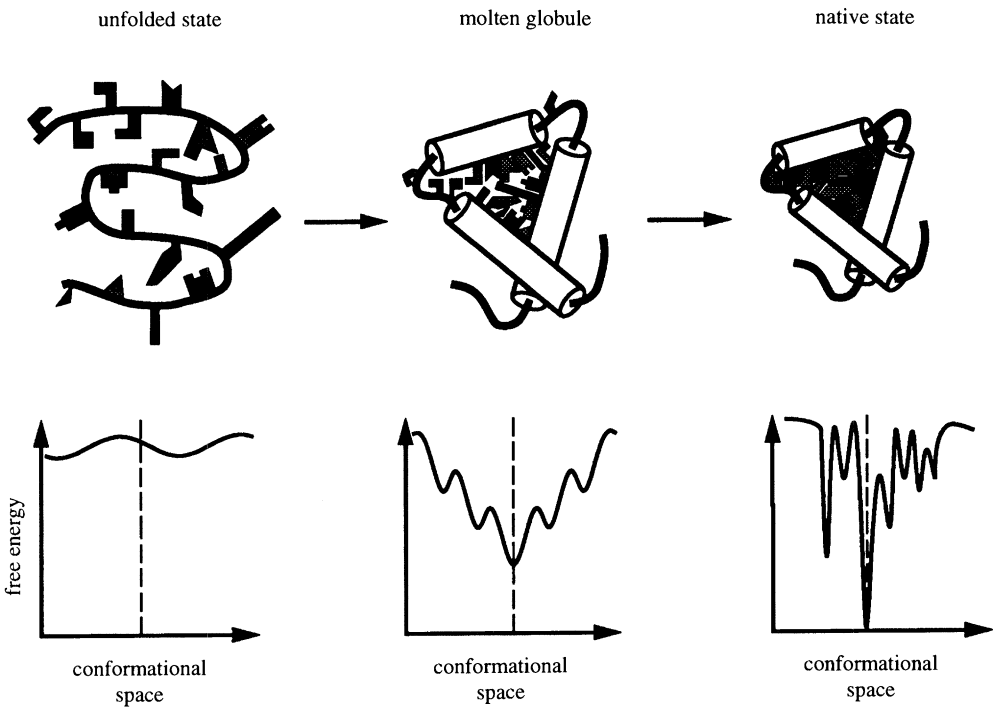

Figure 4. A two-step model of protein folding. The hypothetical structures of a protein in the unfolded, molten globule, and native states are shown along with the free energy landscapes at different stages of the folding reaction.

model. Circular dichroism studies indicate that α-Domain^{ox} has substantially more helical secondary structure than reduced α-Domain (Peng & Kim 1994). In contrast, both non-native disulphide bond isomers ((6–28; 111–120) and (6–111; 28–120)) contain significantly less helical secondary structure than reduced α-Domain. Taken together, these data indicate that the polypeptide backbone of α-Domain prefers a native-like topology and that non-native topologies imposed by non-native disulphide bonds are inconsistent with the high level of secondary structure found in the molten globule of intact α-LA.

Disulphide exchange studies of the molten globule of the entire α-LA molecule with all eight cysteines intact failed to show a strong preference for the species with native disulphide pairings (Ewbank & Creighton 1991, 1993*a*). Instead, many disulphide bond isomers were significantly populated, and only average properties could be examined, since individual disulphide species could not be studied separately. These results were interpreted to indicate that molten globules are actually non-specific collapsed polypeptides, in apparent contradiction to the studies of α-Domain. Two explanations of this discrepancy are apparent.

First, it is possible that the three glycine residues in α-Domain, substituting for 41 residues of the β-sheet domain, constrain the flexibility of the α-Domain backbone, thereby providing a bias toward a topology

with native disulphide pairings. Alternatively, the β-sheet domain may be unstructured in the molten globule of intact α-LA. If this were the case, then those α-LA molecules with native disulphide pairings in the helical domain would be spread among three equally probable β-sheet domain and interdomain disulphide pairings. In addition, interdomain disulphide exchange would further obscure the preference of the helical domain for native disulphide pairings. A rough calculation of the expected disulphide populations can be made, assuming, as a crude approximation, that those species of α-LA with native disulphide pairings in the α-helical domain are favoured twenty-fold over all other species. Then, the populations of α-LA molecules with native disulphide bonds in the α-helical domain would be spread over three peaks, each containing $\sim 12\%$ of the population, with the rest of the α-LA population contained in the other 102 disulphide isomers.

Recent experimental results indicate that the molten globule form of intact α-LA has a bipartite structure (Wu *et al.* 1995). The α-helical domain strongly favours the native backbone topology, while the β-sheet domain is largely unstructured. This finding provides a likely resolution to the apparent discrepancy between the studies of α-Domain (Peng & Kim 1994) and intact α-LA (Ewbank & Creighton 1991, 1993*a*) molten globules, as outlined above. In addition, these results demonstrate that molten globule properties need not encompass the entire polypeptide chain and can be achieved independently by individual domains.

Molten globules can form very quickly (typically $\leqslant 20$ ms; see, for example, Kuwajima *et al.* 1987; Gilmanshin & Ptitsyn 1987). Late folding intermediates containing extensive tertiary interactions, including the so-called 'highly ordered molten globule' (Redfield *et al.* 1994; Feng *et al.* 1994), are known to be native-like. Our results suggest that even early folding intermediates, such as molten globules (in the traditional sense), have a native-like tertiary fold, providing a quick and approximate solution to the Levinthal paradox.

Our results suggest a two-step model for protein folding in which early formation of the molten globule achieves much of the information transfer from one- to three-dimensions (figure 4). Thus, the role of molten globules in protein folding is to maintain an approximate native-like structure, thereby greatly decreasing the conformational space to be searched by the polypeptide chain and preventing global misfoldings. The subsequent search for a unique folded conformation is facilitated by the flexibility of molten globules, which reduces the energy barriers for side chain rearrangements.

How do molten globules achieve the native tertiary fold? One extreme is that molten globules contain specific tertiary interactions that are not detectable by near-UV CD and NMR studies. The other extreme is that more global features, such as the pattern of hydrophilic and hydrophobic residues, side-chain volumes, and local secondary structure propensities largely determine the tertiary fold of a protein. Further experiments should resolve this important question.

This work was supported by the NIH (GM-41307).

REFERENCES

Alexandrescu, A.T., Evans, P.A., Pitkeathly, M., Baum, J. & Dobson, C.M. 1993 Structure and dynamics of the acid-denatured molten globule state of α-lactalbumin: a two-dimensional NMR study. *Biochemistry* **32**, 1707–1718.

Baum, J., Dobson, C.M., Evans, P.A. & Hanley, C. 1989 Characterization of a partly folded protein by NMR methods: Studies on the molten globule state of guinea pig α-lactalbumin. *Biochemistry* **28**, 7–13.

Christensen, H. & Pain, R.H. 1991 Molten globule intermediates and protein folding. *Eur. Biophys. J.* **19**, 221–229.

Chyan, C.L., Wormald, C., Dobson, C.M., Evans, P.A. & Baum, J. 1993 Structure and stability of the molten globule state of guinea-pig alpha-lactalbumin: a hydrogen exchange study. *Biochemistry* **32**, 5681–5691.

Creighton, T.E. & Ewbank, J.J. 1994 Disulfide rearranged molten globule state of α-lactalbumin. *Biochemistry* **33**, 1534–1538.

Dolgikh, D.A., Abaturov, L.V., Bolotina, I.A. *et al.* 1985 Compact state of a protein molecule with pronounced small-scale mobility. *Eur. Biophys. J.* **13**, 109–121.

Dolgikh, D.A., Gilmanshin, R.I., Brazhnikov, E.V. *et al.* 1981 α-Lactalbumin: compact state with fluctuating tertiary structure. *FEBS Lett.* **136**, 311–315.

Ewbank, J.J. & Creighton, T.E. 1991 The molten globule protein conformation probed by disulphide bonds. *Nature, Lond.* **350**, 518–520.

Ewbank, J.J. & Creighton, T.E. 1993*a* Pathway of disulfide-coupled unfolding and folding of bovine alpha-lactalbumin. *Biochemistry* **32**, 3677–3693.

Ewbank, J.J. & Creighton, T.E. 1993*b* Structural characterization of the disulfide folding intermediates of bovine alpha-lactalbumin. *Biochemistry* **32**, 3694–3707.

Feng, Y.-q., Sligar, S.G. & Wand, A.J. 1994 Solution Structure of Apo-cytochrome b562. *Nature Struct. Biol.* **1**, 30–35.

Gilmanshin, R.I. & Ptitsyn, O.B. 1987 An early intermediate of refolding α-lactalbumin forms within 20 ms. *FEBS Lett.* **223**, 327–329.

Ikeguchi, M., Kuwajima, K., Mitani, M. & Sugai, S. 1986 Evidence for identity between the equilibrium unfolding intermediate and a transient folding intermediate: A comparitive study of the folding reactions of α-lactalbumin and lysozyme. *Biochemistry* **25**, 6965–6972.

Jennings, P.A. & Wright, P.E. 1993 Formation of a molten globule intermediate early in the kinetic folding pathway of apo-myoglobin. *Science, Wash.* **262**, 892–896.

Kauzmann, W. 1959 Relative probabilities of isomers in cystine-containing randomly coiled polypeptides. In *Sulfur in proteins* (ed. R. Benesch, R.E. Benesch, P.D. Boyer *et al.*), pp. 93–108. New York: Academic Press.

Kuwajima, K. 1989 The molten globule state as a clue for understanding the folding and cooperativity of globular-protein structure. *Proteins Struct. Funct. Genet.* **6**, 87–103.

Kuwajima, K., Hiraoka, Y., Ikeguchi, M. & Sugai, S. 1985 Comparison of the transient folding intermediates in lysozyme and α-lactalbumin. *Biochemistry* **24**, 874–881.

Kuwajima, K., Nitta, K., Yoneyama, M. & Sugai, S. 1976 Three-state denaturation of α-lactalbumin by guanidine hydrochloride. *J. molec. Biol.* **106**, 359–373.

Kuwajima, K., Yamaya, H., Miwa, S., Sugai, S. & Nagamura, T. 1987 Rapid formation of secondary

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

structure framework in protein folding studied by stopped-flow circular dichroism. *FEBS Lett.* **221**, 115–118.

Levinthal, C. 1968 Are there pathways for protein folding? *J. Chim. Phys.* **85**, 44–45.

Miranker, A., Radford, S.E., Karplus, M. & Dobson, C.M. 1991 Demonstration by NMR of folding domains in lysozyme. *Nature, Lond.* **349**, 633–636.

Miranker, A., Robinson, C.V., Radford, S.E., Alpin, R.T. & Dobson, C.M. 1993 Detection of transient protein folding populations by mass spectrometry. *Science, Wash.* **262**, 896–900.

Nozaka, M., Kuwajima, K., Nitta, K. & Sugai, S. 1978 Detection and characterization of the intermediate on the folding pathway of human α-lactalbumin. *Biochemistry* **17**, 3753–3758.

Peng, Z.-y. & Kim, P.S. 1994 A protein dissection study of a molten globule. *Biochemistry* **33**, 2136–2141.

Ptitsyn, O.B. 1987 Protein folding: hypotheses and experiments. *J. Protein Chem.* **6**, 273–293.

Ptitsyn, O.B. 1992 The molten globule state. In *Protein folding* (ed. T.E. Creighton), pp. 243–300. New York: W.H. Freeman and Co.

Ptitsyn, O.B., Pain, R.H., Semisotnov, G.V., Zerovnik, E. & Razgulyaev, O.I. 1990 Evidence for a molten globule state as a general intermediate in protein folding. *FEBS Lett.* **262**, 20–24.

Radford, S.E., Dobson, C.M. & Evans, P.A. 1992 The folding of hen lysozyme involves partially structured intermediates and multiple pathways. *Nature, Lond.* **358**, 302–307.

Redfield, C., Smith, R.A.G. & Dobson, C.M. 1994 Structural characterization of a highly-ordered 'molten globule' at low pH. *Nature Struct. Biol.* **1**, 23–29.

Wu, L.C., Peng, Z.-y. & Kim, P.S. 1995 Bipartite structure of the α-lactalbumin molten globule. *Nature Struct. Biol.* **2**, 281–286.

Xie, D., Bhakuni, V. & Freire, E. 1991 Calorimetric determination of the energetics of the molten globule intermediate in protein folding: Apo-α-lactalbumin. *Biochemistry* **30**, 10673–10678.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES