Eur Biophys J (1991) 19:221–229

European Biophysics Journal

Invited review

Molten globule intermediates and protein folding

H. Christensen and R. H. Pain

Department of Biochemistry and Genetics, University of Newcastle upon Tyne, Newcastle upon Tyne NE2 4HH, United Kingdom

Received October 16, 1990/Accepted in revised form December 18, 1990

Abstract. The background to the concept of the term "molten globule" as a description of intermediates observed in the folding of globular proteins is discussed. These compact intermediates are characterised by certain properties including the presence of secondary structure and considerable conformational mobility compared to the native, functional state. Those intermediates that are thermodynamically stable under mild denaturing conditions have many features in common with the transient intermediates that accumulate significantly during the process of folding. Attention is drawn to cases where the two types are however distinguished on grounds of their Stokes radius, in which cases there is currently no direct evidence for the involvement of the stable intermediates on the folding pathway. Experimental evidence relating to the early stages in folding is reviewed and compared, highlighting the temporal relationship between general collapse of the polypeptide chain and the formation of secondary structure. The continued use of the term "molten globule" is recommended where the minimum essential structural criteria for this state are met.

Key words: Protein folding – Folding intermediate – Molten globule – Compact states

Introduction

At a conference organised by the British Biophysical Society on Protein Folding (Newcastle upon Tyne, 1990), considerable attention was focussed upon the nature of the early intermediate states in folding. This necessarily involved consideration of "compact intermediate" states and their relation to the "molten globule". In this paper, we present a view of the salient features of the current debate in the light of contributions from the conference. For recent reviews of the relevant literature the reader is referred to Kim and Baldwin (1990) and Kuwajima (1989).

Abbreviations: ANS, 8-anilino-1-naphthalenesulphonic acid; CD, circular dichroism; GEC, gel exclusion chromatography; NMR, nuclear magnetic resonance; UGGE, urea gradient gel electrophoresis; UV, ultraviolet

Offprint requests to: R. H. Pain

Background to the molten globule concept

Globular proteins were originally so defined on the basis of their intrinsic viscosity that showed them to be compactly folded and approximately spherical. A major contribution to the stability of this native state has been shown to be the hydrophobic interaction and three-dimensional structures confirm the general non-polar "inside" and polar "outside" feature of these "globules" (Dill 1990). The other readily identifiable feature of the native structure is the secondary structure present to a greater or lesser degree in nearly all globular proteins. The polarity of the faces of these secondary structure elements is complementary to the rest of the protein conformation and in particular to the protein-water interface at the surface, implying an interdependence between secondary structure and the globular state.

The kinetic attainment of the folded, native state cannot depend on a random search (Levinthal 1968) and much of the research on protein folding has been based on the assumption that a limited number of pathways exist involving kinetic intermediates between the fully unfolded and the native states, (Kim and Baldwin 1982). The major question concerns the nature of these intermediate states. Evidence is available from a variety of studies using different spectroscopic and hydrodynamic criteria. Two distinct approaches may be identified. One involves monitoring kinetic intermediate species associated with the pathway of folding or unfolding, where these pathways are not necessarily identical. The other involves the characterisation of thermodynamically stable intermediates present under given solution conditions, usually induced by denaturant conditions which bring about only partial unfolding from the native state. In this case, the thermodynamic definition requires that the same states exist under the given solution conditions, irrespective of the way in which these were established.

Secondary structure elements are an obvious feature of native proteins, and, in view of the high rate of the coil-to-helix transition in model compounds (Gruenewald et al. 1979), these elements have consequently been regarded as strong candidates for early intermediates in folding. It has

also been recognised however that further interactions would be required to stabilise the secondary structure elements owing to their relatively low intrinsic thermodynamic stability, especially at in vivo folding temperatures. A kinetic model that takes into account these two features is the framework model described by Ptitsyn (1973) and by Kim and Baldwin (1982). Secondary structure is proposed to form first and to associate to form the native state. In Ptitsyn's model association occurs to a compact globular state within which organisation of tertiary interactions occurs.

The interdependence of secondary structure and globular states can again be seen in Ptitsyn's model where the compact non-native globular conformation provides an environment of the appropriate polarity essential for stabilisation of secondary structures. Looking at the problem from the point of view of the occurrence of helices in folded proteins, a similar conclusion was reached and an intermediate species proposed that foreshadowed the experimentally based model of the "molten globule" (Robson and Pain 1971).

The term "molten globule", descriptive of a species whose conformation is intermediate between that of the native and fully unfolded states, has been applied almost entirely to the species found to be thermodynamically stable under mild denaturing conditions. Such states have been thoroughly characterised for the small number of proteins so far shown to exhibit such behaviour. The persistence of such states depends upon a fine balance of stabilising interactions such that the free energy difference between the native and intermediate states is less than that between the native and unfolded states i.e. when the structure is less than fully cooperative. The term "molten globule" has however been applied also to a state with similar but not identical properties, that has been shown to accumulate in a transient manner during folding of a larger number of proteins. Under these conditions there is as yet no evidence for significant accumulation of the structure found under mild denaturing conditions. In this review, attention is drawn to the distinction between these two types of species and emphasis is placed on the latter as being the one that is demonstrably relevant to the folding pathway of globular proteins.

More recently the term "molten globule" has been applied to a variety of less well characterised states extending from early intermediates to late, near-native states, leading to a degree of confusion and thus rendering the term less useful. Consequently attempts are being made to limit its use (Kim and Baldwin 1990). We wish to propose here that the term continue to have a useful role to play in describing a certain class of folding intermediates exhibiting particular properties. Its use should be limited however to states for which those properties have been well established while maintaining an open mind towards other potential candidates for the name.

Stable states of intermediate conformation

Much of the detailed characterisation of the so-called molten globule state has been carried out under mild denaturing conditions where a limited number of proteins exhibit states that are thermodynamically stable and whose conformation is intermediate between the native state, *N*, and the fully unfolded state, *U*. These states are found under acidic (*A*) and basic (*B*) conditions and in the presence of intermediate concentrations of urea or guanidinium chloride at neutral pH (*H*) as listed in Table 1.

The main characteristics of these stable intermediate states are:

Secondary structure is always present

Circular dichroism spectra show considerable amounts of secondary structure. For β-lactamase state H (neutral pH, 2 M urea), for example, the far UV ellipticity has an intensity 85% that of the native state (Carrey and Pain 1978). In the cases of α-lactalbumin (Baum et al. 1989) and cytochrome *c* (Roder et al. 1988), both at low pH, the use of NMR has shown that the helices that are present are located at positions in the sequence closely similar to those in the native state.

Apparent discrepancies exist between conclusions based on NMR and on CD results as to the precise secondary structure content of these intermediate states. Whatever the basis of these differences the general picture is one of a substantial fraction of stable secondary structure within the thermodynamically stable intermediate states.

Tertiary interactions are much reduced in stability

In these intermediate states, aromatic residues in general exhibit greater if not complete access to solvent (from difference spectroscopy e.g. Robson and Pain 1976) and loss of asymmetry in environment (from near UV CD e.g. Mitchinson and Pain 1985). High resolution NMR spectra show that a number of both aromatic and aliphatic residues are less ordered (Thomas et al. 1983). The stability of a number of non-secondary structure hydrogen bonds in the *A* state of cytochrome *c* is more than two orders of magnitude lower than those involved in helices in the same state (Feng et al. 1990). Only those considered to be contributing to stabilisation of the helices are more stable. Similar detailed NMR studies on the *A* state of α-lactalbumin (Baum et al. 1989) show that, while secondary structure is present and very stable compared with isolated peptide sequences, there is mobility in the hydrophobic core and considerable disorder in other regions. Similar conclusions have been drawn from the *A* state of apomyoglobin (Goto et al. 1990a, b).

The state is largely globular

Intuitively, the interactions required for the persistence of stable helices must lead to a more compact state than the unfolded state. Experimentally this has been shown to be true using X-ray diffraction (Damaschun et al. 1986), viscosity (Dolgikh et al. 1981), gel exclusion chromatography (Ptitsyn et al. 1990), light scattering (Gast et al.

Table 1. Stable and transient intermediate states

Protein	States[a]	Characterised by	References
β-lactamase	*A*	(Equilibrium studies) CD	Goto et al. (1990) Goto and Fink (1989)
	B		Goto and Fink (1989); Craig (1986)
	H	UGGE GEC	Robson and Pain (1976) Creighton and Pain (1980) Craig and Pain (1978)
		NMR, CD, viscosity	Thomas et al. (1983) Mitchinson and Pain (1985)
	I	UGGE GEC NMR CD	Creighton and Pain (1986) Zerovnik and Pain (1986) Ptitsyn et al. (1990) Mitchinson and Pain (1985)
	Kinetic intermediates	ANS binding	Ptitsyn et al. (1990); Fig. 2
Interleukin IL-1β	*I*	GEC	Ptitsyn et al. (1990) Christensen and Pain (in preparation)
Cytochrome *c*	*A*	CD, NMR, Viscosity	Ohgushi and Wada (1983)
	Kinetic intermediate	*H*-exchange NMR Stopped flow CD	Roder et al. (1988) Kuwajima et al. (1987)
Bovine carbonic anhydrase B	*I*	GEC, NMR, CD	Ptitsyn et al. (1989) Semisotnov et al. (1987) Dolgikh et al. (1984)
	H	GEC	Ptitsyn et al. (1990)
	Kinetic intermediate	ANS binding	Ptitsyn et al. (1990)
α-lactalbumin	*A*	NMR, CD	Baum et al. (1989) Dolgikh et al. (1985) Dolgikh et al. (1981)
	Kinetic intermediate	ANS binding CD Stopped flow CD	Ptitsyn et al. (1990) Kuwajima et al. (1985) Gilmanshin and Ptitsyn (1987)
β-lactoglobulin	Kinetic intermediate	Stopped flow CD ANS binding	Kuwajima et al. (1987) Ptitsyn et al. (1990)
Ribonuclease A	Kinetic intermediate	Stopped flow (Tyrosine absorbance)	Garel et al. (1976)
		H-exchange NMR	Udgaonkar and Baldwin (1988)
Octopine dehydrogenase	Kinetic intermediate	Stopped flow (protein fluorescence)	Teschner et al. (1987)
Lysozyme	Kinetic intermediate	Stopped flow (protein absorbance)	Kato et al. (1981)
		Stopped flow CD H-exchange NMR	Kuwajima et al. (1985) Miranker et al. (1991)
Tryptophan synthase	Kinetic intermediate	Stopped flow CD Monoclonal antibodies fluorescent probes	Murry-Brelier and Goldberg (1988) Goldberg et al. (1989)
Dihydrofolate reductase	Kinetic intermediate	Stopped flow fluorescence	Garvey et al. (1989)
Barnase	Kinetic intermediate	Stopped flow fluorescence *H*-exchange NMR	Matouschek et al. (1990) Bycroft et al. (1990)
Phosphoglycerate kinase	Kinetic intermediate	ANS binding	Ptitsyn et al. (1990)

GEC, gel exclusion chromatography; UGGE, urea gradient gel electrophoresis; ANS, Anilinonaphthalene sulphonic acid
[a] For definitions, see text

1986) and urea gradient gel electrophoresis (Creighton and Pain 1980).

Different proteins exhibit stable intermediate states whose equivalent hydrodynamic radii of gyration or Stokes radii range from approximately one to 2 times those of the respective native states. Where all other properties coincide with those for a compact or molten globule state, it is proposed that those proteins with larger volumes, e.g. β-lactamase (*H* state) (Carrey and Pain 1978) and apomyoglobin (*A* state) (Goto et al. 1990a) are mainly compact states that include the secondary structure, with extended random coil loops or chain termini.

Stable intermediate states tend to aggregate

Under conditions where such states are thermodynamically stable, they are frequently kinetically unstable owing to aggregation. This property is also reflected in the binding of the fluorescent hydrophobic probe anilino-naphthalene sulphonic acid (ANS) by these states, to a greater extent than either the native or unfolded states (Semisotnov et al. 1987). This may indicate the presence of substantial non-polar patches on the surface of the incompletely packed compact state or, alternatively, to intercalation of the probe into the globule. The latter would be consistent with the more mobile nature of the hydrophobic core referred to above and with complementary results on the heat capacity that suggest a degree of water inclusion between non-polar surfaces (Pfeil 1987).

The states behave thermodynamically as if "molten"

The transition between the native and stable intermediate states is frequently two state as demonstrated by the isobestic points for the transition (Goto and Fink 1989; Goto et al. 1990b) and by the lack of other populated states characterised by size (Creighton and Pain 1980; Ptitsyn et al. 1990). The thermodynamic stability of the stable equilibrium state is less than that of the native state, in the case of β-lactamase by 10 kJ · mol^{-1}. On the assumption that the transition is two state it is stable relative to the unfolded state again by 10 kJ · mol^{-1} (Mitchinson and Pain 1985).

Kim and Baldwin (1990) have discussed the finding that there is no difference in heat capacity between the compact intermediates and fully unfolded states (Pfeil et al. 1986). There is, in addition, evidence that the thermal unfolding of such a state under the solvent conditions in which it is stable at room temperature is non-cooperative. The *H* state of β-lactamase (neutral pH) exhibits a transition on isothermal unfolding to state *U* with increasing urea concentrations that has the appearance of co-operativity (Robson and Pain 1976; Mitchinson and Pain 1985). Theoretical considerations (Finkelstein and Shakhnovich 1989) suggest that distinction between a co-operative and a non-cooperative transition under these conditions is beyond the resolution of the experiment.

In summary, the "non-native, compact" intermediate has a structure best described as comprising a significant proportion of native-like secondary structure stabilised by interactions between secondary structure elements and/or interactions with a relatively well packed but mobile non-polar core or globule e.g. α-lactalbumin and cytochrome *c* (Baum et al. 1989; Roder et al. 1988). Other elements of structure may be less well ordered and may or may not be collapsed on to the central globular element. Thus this state has characteristics that are both "globular" and "molten". The degree to which it is truly molten depends upon further resolution of the thermodynamics and degree of co-operativity of the transition to the unfolded state.

The question arises as to whether these states lie on any kinetic folding pathway. If only because of their charge difference, those states observed under acid or basic conditions cannot do so. Whether the neutral pH (*H*-type) states are involved in folding will be addressed below. The relatively detailed characterisation of all these states has however helped in understanding those compact states that have been clearly demonstrated to accumulate during folding.

Kinetically accumulating compact states

The accumulation during folding of an intermediate nearly as compact as the native state has been demonstrated (Creighton and Pain 1980; Dolgikh et al. 1984). This intermediate (*I*-type) is formed rapidly ($t_{1/2} < 1$ s) from state *U* and folds more slowly to state *N*. Extrapolation of the kinetics of the latter phase to zero time together with examination at low temperature and low denaturant concentrations (to slow the transition to *N*) have led to the following characteristics for state *I* (summarised in Ptitsyn et al. 1990).

It is compact, as shown by urea gradient gel electrophoresis (Creighton and Pain 1980), viscosity (Dolgikh et al. 1984) and by gel exclusion chromatography (Ptitsyn et al. 1990); it contains secondary structure (Dolgikh et al. 1984; Mitchinson and Pain 1985) with about 85% of the ellipticity of the native state; the high resolution NMR spectra differ little from those for the stable *H*-type intermediate (Thomas et al. 1983; Ptitsyn et al. 1990). Compared with the native state, NMR spectra (Fig. 1) for state *I* show less detailed structure in the aromatic and high field methyl regions indicating far fewer stable tertiary interactions than in the native state, consistent with the structure described above for α-lactalbumin. The lack of native tertiary structure in state *I* is supported by CD measurements (Mitchinson and Pain 1985). Like the stable intermediates of enzymes, state *I* has no catalytic activity (Dolgikh et al. 1984; Mitchinson and Pain 1985) and it has non-polar residues that are less well packed than in the native state and that therefore bind ANS (Ptitsyn et al. 1990). The rate constant for accumulation of the native state is independent of whether the protein is folded from the acid, base, *U* or *H* states suggesting that *I* is a common intermediate (Craig 1986). The identity of the rate constants for this transition measured by different parameters suggests that there is only one significant energy barrier between it and the native state. It is essen-

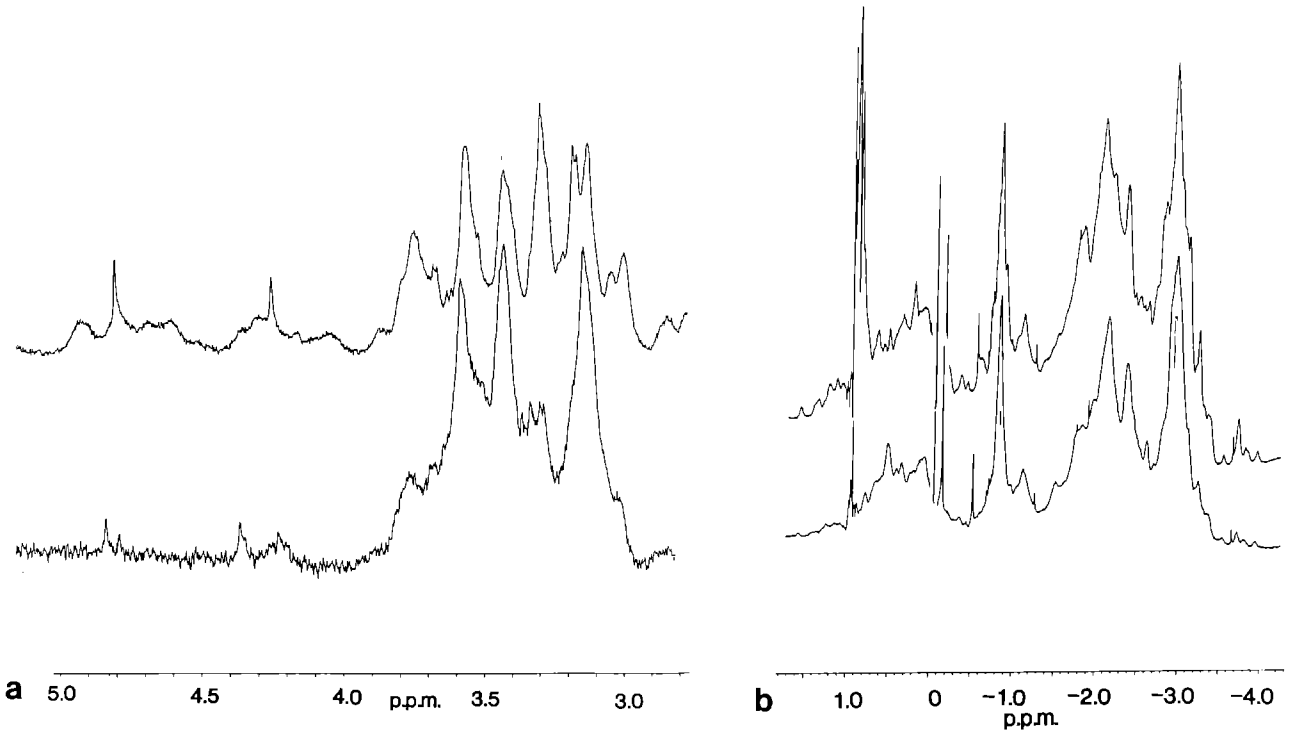

Fig. 1. High resolution spectra of the kinetic intermediate *I* (*upper*) and native state of β-lactamase in **a** aromatic region and **b** aliphatic region. Taken from Ptitsyn et al. (1990)

tially a transient species, there being no conditions under which folding to the native state will occur where state *I* is significantly populated at equilibrium.

Accumulation of this type of compact intermediate has been demonstrated directly by gel exclusion chromatography for interleukin IL-1β (Ptitsyn et al. 1990), a protein whose equilibrium unfolding transition is reported to be two state (Craig et al. 1987). Type I compact intermediates are therefore not restricted to proteins that exhibit an *H*-type stable intermediate. ANS binding also occurs to a number of other proteins during the early stages of folding (see next section) strongly supporting the involvement of the collapsed, type *I*, intermediate state.

For carbonic anhydrase and β-lactamase the type *H* and type *I* intermediates can be clearly distinguished in gel exclusion chromatography, elution volumes relative to *N* of 0.85 (*H*) and 0.90 (*I*) being typical (Ptitsyn et al. 1990). It has been found that, on folding the protein from the fully unfolded state, the stable type *H* state does not accumulate under any conditions where the native state is stable (Ptitsyn et al. 1990). It has been reported, on the basis of the similarity of their far UV CD spectra, that the two types of intermediate for α-lactalbumin are identical (Ikeguchi et al. 1986). The close similarity of the spectra for the transient and stable intermediates of carbonic anhydrase and β-lactamase, despite their discrimination on the basis of size, suggests that the proposed identity for α-lactalbumin must still be regarded as an open question.

It is becoming clear that, in many if not all globular proteins, folding occurs through a compact intermediate that is formed rapidly and that this intermediate exhibits very similar characteristics to the stable intermediate hitherto termed the molten globule state.

Formation of the transient compact intermediate

Stopped flow studies of ANS binding and its subsequent desorption have helped in characterising the early stages of protein folding (Semisotnov et al. 1987; Goto and Fink 1989; Ptitsyn et al. 1990; Goldberg et al. 1990). Stopped flow studies of ANS binding to tryptophan synthase β_2-subunit (Goldberg et al. 1990), β-lactamase (Fig. 2), interleukin IL-1β, carbonic anhydrase and phosphoglycerate kinases from yeast and *Thermus thermophilus* (Christensen, Pain and Varley, unpublished), all carried out at reduced temperatures, show that maximum fluorescence and hence binding is attained in 100 ms or less, in two or more kinetic phases (Fig. 2), and therefore prior to formation of the *I* state.

The fastest measurable phase has a half-life of 2 to 3 ms that is more or less independent of the protein and consistent with the timescale within which secondary structure forms.

These results suggest a restatement of earlier conclusions (Ptitsyn et al. 1990) to the effect that the kinetics of collapse and the kinetics of formation of secondary structure, at least for some proteins, may be of the same order. This will be so, if residues that transiently form secondary structure on a very fast (μs) timescale are stabilised by the collapse. It could also be explained by an initial collapse to a globular intermediate within which secondary structures form. Firm conclusions must await further technical developments in establishing rate constants for the formation and stabilisation of secondary structure.

In the folding of ribonuclease (Udgaonkar and Baldwin 1988), cytochrome *c* (Roder et al. 1988) and barnase, the RNase from *Bacillus amyloliquefaciens* (Bycroft et al.

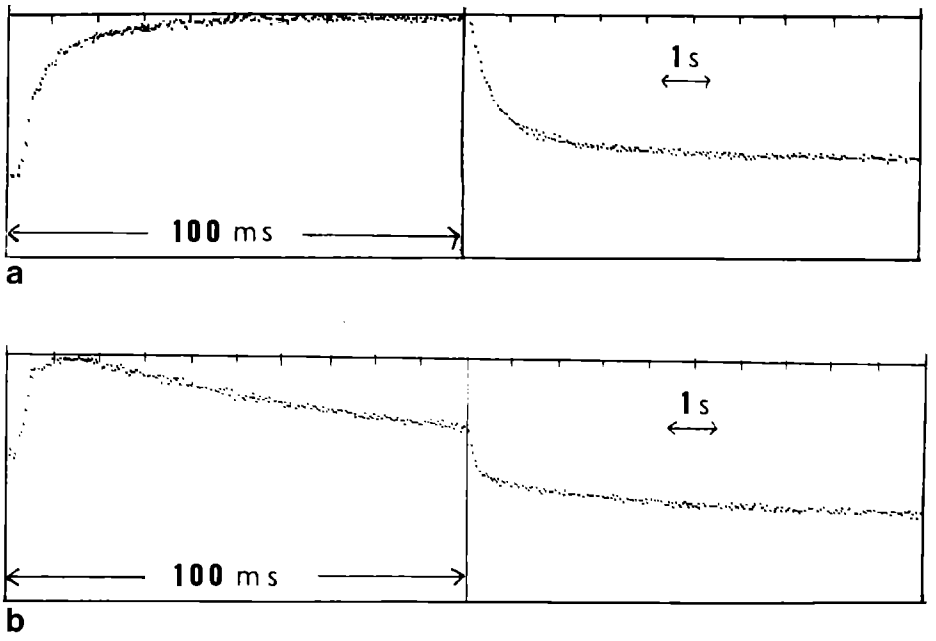

Fig. 2. Stopped flow kinetics showing ANS binding intermediates of β-lactamase at **a** 4 °C and **b** 23 °C. Experiments were carried out in 0.1 M phosphate buffer pH 7.0 and refolding was initiated by a 6 fold dilution of protein in 2 M guanadinium chloride. Final protein concentration was 6×10^{-6} M and ANS 10^{-4} M. At low temperature (4 °C) biphasic binding ($k_1 = 340\ s^{-1}$, $k_2 = 47\ s^{-1}$) and biphasic desorption ($k_3 = 1.3\ s^{-1}$, $k_4 = 0.2$) are observed. At 23 °C the equivalent desorption kinetics are $k_3 = 6.8\ s^{-1}$ and $k_4 = 0.3\ s^{-1}$. Fluorescence excitation was at 380 nm and emission was observed above 450 nm

1990), the deuterium entrapment technique suggests that a considerable proportion, though not all, of the amide hydrogens in the native β-sheet and helical structures become protected within a few tens of milliseconds. Fast reaction CD measurements on other proteins also show that secondary structure is formed and therefore stabilised within the same timescale (Gilmanshin and Ptitsyn 1987; Kuwajima et al. 1987; Goldberg et al. 1990).

Desorption of ANS takes place with several phases, all but the final one being on a timescale too fast to be seen under manual mixing conditions. Thus there is a significant number of intermediates, generally assumed to be consecutive, that form prior to the *I*-type intermediate which is defined above as the state to which the final first order phase of the folding process extrapolates at zero time. Most, if not all of these, contain secondary structure and the study of barnase indicates the progressive formation of a hydrophobic core (Matouschek et al. 1990). It is suggested that the late stage of folding involves docking of preformed elements. Whether this occurs within or independently of a molten globule is a significant question.

The question thus remains as to the extent of collapse of all these ANS binding, secondary structure containing intermediates. With the exception of the *I*-type intermediate it cannot be assumed without further supporting evidence that the initial ANS binding phase defines the formation of a collapsed, globular state.

A particularly good example of a folding pathway where a large number of states have been defined and characterised kinetically comes from the studies on the folding of the β_2 subunit of tryptophan synthase. In addition to ANS binding, three other spectroscopic probes were utilised; dansyl-Cys 170, the reduced Schiff base between pyridoxal phosphate and Lys87, and the intrinsic fluorescence of Trp177 (Goldberg et al. 1990; Blond-Elguindi and Goldberg 1990). Energy transfer monitors changes in distance between the tryptophan and the dansyl and pyridoxyl groups. Conformation specific monoclonal antibodies were also used to identify the formation of native and native-like epitopes during folding. Using these probes the following intermediates were identified kinetically, using a wide range of probes, and the rate constants of their formation measured:

$$\begin{array}{ccccccccc} \beta_u & \longrightarrow & \beta_{SS} & \longrightarrow & \beta'_{ANS} & \longrightarrow & \beta''_{ANS} & \longrightarrow & \beta'''_{ANS} \longrightarrow \beta_{MAB19} \longrightarrow \\ >100 & & 22 & & 4 & & 0.7 & & 0.065 \quad 0.02 \end{array}$$

$$\begin{array}{ccccccccc} \beta^{I1} & \longrightarrow & \beta^{I2} & \longrightarrow & \beta^{N'} & \longrightarrow & \beta_2^{N'} & \longrightarrow & \beta_2^{N''} \longrightarrow \beta_2 \quad \text{(native)} \\ 0.008 & & 0.001 & & & & \text{fast} & & 0.002 \end{array}$$

Scheme for the folding of the β_2 subunit of tryptophan synthase. β is the monomer comprising two domains; β_2 is the dimer; β_{SS} is the first intermediate seen with secondary structure; β'_{ANS}, β''_{ANS} and β'''_{ANS} represent the three kinetic phases for binding ANS; β_{MAB19} possesses the epitope in the F_1 domain (epitopes on the F_2 domain and the linking peptide are formed faster). Rate constants are in s^{-1}. Based on Murry-Brelier and Goldberg (1988); Blond-Elguindi and Goldberg (1990); and Goldberg et al. (1990).

This scheme demonstrates that secondary structure – 57% of the native ellipticity at 225 nm – is formed very rapidly and before the first detected binding of ANS (Fig. 3a and b). An important finding is that the conformational epitope forms within the intermediate that binds the maximum amount of ANS. On the basis that the compact intermediate is defined by ANS binding the authors have concluded that this epitope is formed within the "molten globule". It will be interesting to identify and characterise the states within which earlier epitopes are formed (Blond-Elguindi and Goldberg 1990). Thus kinetically, there seem to be 2 steps involved in the "freezing" of the molten globule: that associated with the formation of β^{I1} and that associated with the formation of $\beta^{N'}$ (Goldberg, personal communication).

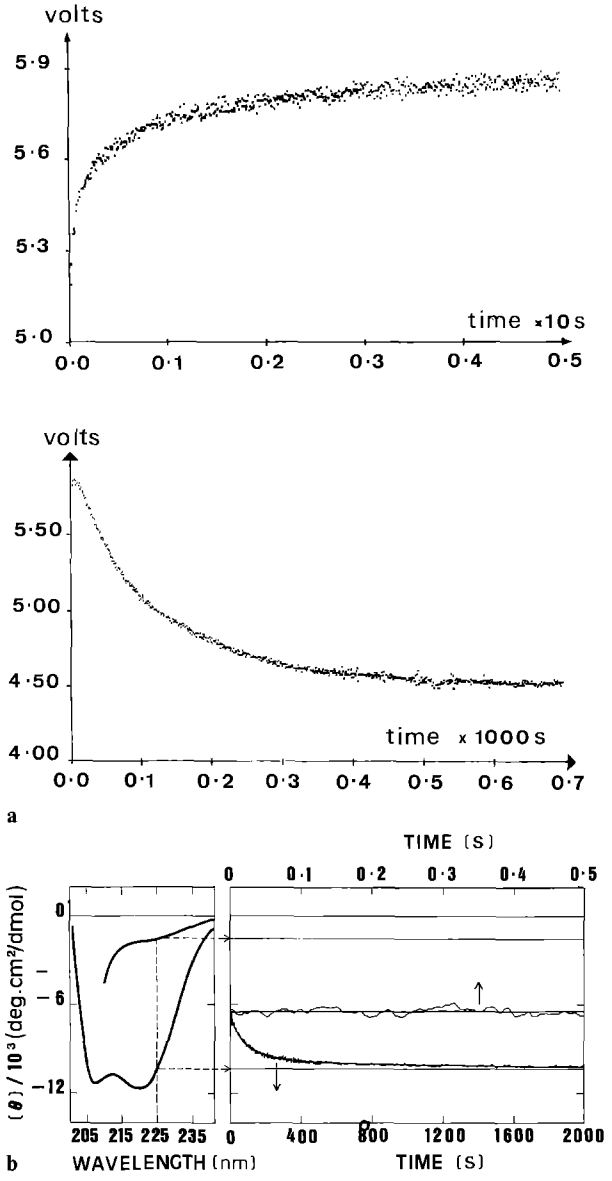

Fig. 3. **a** Kinetics of refolding of the β_2 subunit of Tryptophan synthase monitored by ANS fluorescence. *Upper part*: 5 s full scale; *lower part*: 700 s full scale. **b** Kinetics of refolding as observed by far UV CD. *Left*: far u.v. CD spectra for native and unfolded protein. *Right*: ellipticity as a function of time: 0.5 s full scale (*upper*) and 2 000 s full scale (*lower*). Taken from Goldberg et al. (1990)

Another important aspect illustrated by this work is that more than one kinetically defined intermediate binds ANS. Thus the six species from β'_{ANS} to β^{I2} all bind ANS, with maximum binding to β''_{ANS} and β_{MAB19} and 70% of maximum bound to β^{I1} and β^{I2}. The remaining 70% is desorbed with the formation of $\beta^{N'}$ or $\beta_2^{N'}$, which states are reckoned to have a native like packing density such that the globule is "frozen". The results discussed earlier are consistent with this picture (see Fig. 2), both adsorption and desorption of ANS being multiphasic processes. ANS binding therefore is exhibited during folding by a number of different conformers with different topologies. These undergo transitions to become more native-like as shown by the immunological markers.

For most of the ANS binding states in this and other systems, there exists no independent experimental evidence for their being globular. While the formation of secondary structure and longer range interactions, for which evidence is strong in the early intermediates, must involve a decrease in radius of gyration it cannot, solely on these grounds, be concluded that they are strictly globular or compact.

Conclusions

We have attempted to show that evidence is growing for the general involvement of a particular type of intermediate species that accumulates immediately before the rate determining step in the folding of globular proteins to their native state. This intermediate, originally termed "state *I*", (Creighton and Pain 1980; Dolgikh et al. 1984) is separated from the native or near native state by a transition that is often, but not always kinetically two state. Direct evidence exists for its compactness, regular secondary structure and low level of persistent tertiary interactions. Comparison with the stable intermediate species exhibited only by a limited number of proteins but which can be examined under equilibrium conditions has strengthened the proposal that both the stable and the transient accumulating intermediate (state *I*) have general structural features in common although usually differing in their degree of compactness.

While these intermediate species appear to contain some water, they are stable relative to the unfolded state presumably due to intramolecular interactions. They thus merit the term "globule". The evidence for a generally high level of internal mobility and non-cooperative thermal transitions to the unfolded state merit the term "molten". There is for many proteins therefore a distinctive species, distinct from the stable intermediate and from the fully unfolded state, namely the *I*-type intermediate, that accumulates during folding and that can appropriately and usefully be referred to as the *molten globule*.

A general scheme summarising present findings concerning the involvement of the type *I* molten globule in folding may be proposed as follows:

$$U \rightleftharpoons X_i \rightleftharpoons I \rightleftharpoons N'_j \rightleftharpoons N ,$$

where U is the fully unfolded protein, X_i are species with increasing amounts of stabilised interactions, and N'_j represent any possible intermediates that are essentially frozen but lack a few final detailed interactions such as adjustment between folding domains and/or subunits. The term "intermediates" should not necessarily be taken to imply a species on a consecutive reaction pathway between U and N (discussed by Kim and Baldwin 1990; Mitchinson and Pain 1985; Adams et al. 1980; Ikai and Tanford 1973). In this scheme therefore, I exhibits the properties of a molten globule. Species X_i contain secondary structure and thus may bind ANS more strongly than I. However, there is at present no evidence that the latter are compact and globular, and it is not therefore strictly valid to apply the term molten globule to these species.

Kim and Baldwin (1990) have drawn attention to some of the confusion implicit in the use of the term "molten globule". In our view the main problem lies in the use of the term "molten globule" and "collapsed form" for some of the X_i where evidence for collapse to a globule is presently lacking – such evidence would provide an answer to the important question as to whether any of X_i are molten globules (Kuwajima 1989). We would propose therefore that, until they are more fully characterised, the early intermediates X_i should be described in terms of their known properties only, and not in terms of any particular model.

Acknowledgements. This work has been supported by the Science and Engineering Research Council and the Medical Research Council. We thank Professor A. R. Fersht and Dr. H. Roder for providing access to their work prior to publication. The participants at the British Biophysical Society meeting on Protein Folding and Assembly for stimulating discussion and Dr. M. E. Goldberg for subsequent comments. Susan Lee and Janis Wilkinson are thanked for preparation of the manuscript.

References

Adams B, Burgess RJ, Carrey EA, Mackintosh IR, Mitchinson C, Thomas RM, Pain RH (1980) The role of folding units in the kinetic folding of globular proteins. In: Jaenicke R (ed) Protein folding. Elsevier/North Holland, Amsterdam pp 447–467

Baum J, Dobson CM, Evans PA, Hanley C (1989) Characterisation of a partly folded protein by NMR methods: studies on the molten globule state of guinea pig α-lactalbumin. Biochemistry 28:7–13

Blond-Elguindi S, Goldberg ME (1990) Kinetic characterisation of early immunoreactive intermediates during the refolding of guanidine unfolded *Escherichia coli* tryptophan synthase β_2 subunits. Biochemistry 29:2409–2417

Bycroft M, Matouschek A, Kellis JT, Serrano L, Fersht AR (1990) Detection and characterisation of a folding intermediate in barnase by NMR. Nature 346:488–490

Carrey EA, Pain RH (1978) Conformation of a stable intermediate on the folding pathway of *Staphylococcus aureus* penicillinase. Biochim Biophys Acta 553:12–22

Craig S (1986) Critical residues in the folding of β-lactamase. PhD Thesis, University of Newcastle upon Tyne

Craig S, Schmeissner U, Wingfield P, Pain RH (1987) Conformation, stability and folding of interleukin 1β. Biochemistry 26:3570–3576

Creighton TE, Pain RH (1980) Unfolding and refolding of *Staphylococcus aureus* penicillinase by urea gradient electrophoresis. J Mol Biol 137:431–436

Damaschun G, Germat Ch, Damaschun H, Bychkova VE, Ptitsyn OB (1986) Comparison of intramolecular packing of a protein in native and "molten globule" states. Int J Biol Macromol 8:226–230

Dill KA (1990) Dominant forces in protein folding. Biochemistry 29:7133–7155

Dolgikh DA, Gilmanshin RI, Brazhnikov EV, Bychkova VE, Semisotnov GV, Venyaminov SYu, Ptitsyn OB (1981) α-lactalbumin: compact state with fluctuating tertiary structure. FEBS Lett 136:311–315

Dolgikh DA, Kolomiets AP, Bolotina IA, Ptitsyn OB (1984) "Molten globule" state accumulates in carbonic anhydrase folding. FEBS Lett 165:88–92

Dolgikh DA, Abaturov LV, Bolotina IA, Brazhnikov EV, Bychkova VE, Bushuev VN, Gilmanshin RI, Lebedev YuO, Semisotnov GV, Tiktopulo EI, Ptitsyn OB (1985) Compact state of a protein molecule with pronounced small-scale mobility: bovine α-lactalbumin. Eur Biophys J 13:109–121

Feng M-F, Englander SW, Elöve GA, Wand AJ, Roder H (1990) Structural description of acid denatured cytochrome *c* by hydrogen exchange and 2D NMR. Biochemistry 29:10432–10437

Finkelstein AV, Shaknovich EI (1989) Theory of cooperative transitions in protein molecules. II. Phase diagram for a protein molecule in solution. Biopolymers 28:1681–1694

Garvey EP, Swank J, Matthews CR (1989) A hydrophobic cluster forms early in the folding of dihydrofolate reductase. Proteins: Struct Funct Genet 6:259–266

Gast K, Zirwer D, Welfle H, Bychkova VE, Ptitsyn OB (1986) Quasielastic light scattering from human α-lactalbumin: comparison of molecular dimensions in native and "molten globule" states. Int J Biol Macromol 8:231–236

Gilmanshin RI, Ptitsyn OB (1987) An early intermediate of refolding α-lactalbumin forms within 20 ms. FEBS Lett 223:327–329

Goldberg ME, Semisotnov GV, Friguet B, Kuwajima K, Ptitsyn OB, Sugai S (1990) An early immunoreactive folding intermediate of the tryptophan synthase β_2 subunit is a "molten globule". FEBS Lett 263:51–56

Goto Y, Fink AL (1989) Conformational states of β-lactamase: molten globule states at acidic and alkaline pH with high salt. Biochemistry 28:945–952

Goto Y, Calciano LJ, Fink AL (1990a) Acid-induced folding of proteins. Proc Natl Acad Sci USA 87:573–577

Goto Y, Takahashi N, Fink AL (1990b) Mechanism of acid-induced folding of proteins. Biochemistry 29:3480–3488

Gruenewald B, Nicola CU, Lustig A, Schwarz G, Klump H (1979) Kinetics of the helix-coil transition of a polypeptide with non-ionic side groups, derived from ultrasonic relaxation measurements. Biophys Chem 9:137–147

Ikai A, Tanford C (1973) Kinetics of unfolding and refolding of proteins. I. Mathematical analysis. J Mol Biol 73:145–164

Ikeguchi M, Kuwajima M, Sugai S (1986) Evidence for identity between the equilibrium unfolding intermediate and a transient folding intermediate: a comparative study of the folding reactions of α-lactalbumin and lysozyme. Biochemistry 25:6965–6972

Jeng M-F, Englander W, Elöve GA, Roder H (1991) Structural description of acid-denatured cytochrome *c* by NH exchange and 2D NMR (in press)

Kato S, Okamura M, Shimamoto N, Utiyama H (1981) Spectral evidence of a rapidly formed structural intermediate in the refolding kinetics of hen egg-white lysozyme. Biochemistry 20:1080–1085

Kim PA, Baldwin RL (1982) Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding. Ann Rev Biochem 51:459–489

Kim PS, Baldwin RL (1990) Intermediates in the folding of small proteins. Ann Rev Biochem 59:631–660

Kuwajima K (1977) A folding model of α-lactalbumin deduced from the three-state denaturation mechanism. J Mol Biol 114:241–258

Kuwajima K (1989) The molten globule state as a clue for understanding the folding and co-operativity of globular protein structure. Proteins: Struct Funct Genet 6:87–103

Kuwajima K, Hiraoka Y, Ikegushi M, Sugai S (1985) Comparison of the transient folding intermediates in lysozyme and α-lactalbumin. Biochemistry 24:874–881

Kuwajima K, Yamaya H, Miwa S, Sugai S, Nagamura T (1987) Rapid formation of secondary 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 Chem Phys 65:44–45

Matouschek A, Kellis JT, Serrano L, Bycroft M, Fersht AR (1990) Transient folding intermediates characterised by protein engineering. Nature 346:440–445

Miranker A, Radford SE, Karplus M, Dobsen C (1991) Demonstration by NMR of folding domains in lysozyme. Nature 349:633–636

Mitchinson C, Pain RH (1985) Effects of sulphate and urea on the stability and reversible unfolding of β-lactamase from *Staphylococcus aureus*. Implications for the folding pathway of β-lactamase. J Mol Biol 184:331–342

Murry-Brelier A, Goldberg ME (1988) Kinetics of appearance of an early immunoreactive species during the refolding of acid denatured *Escherichia coli* tryptophan synthase β2 subunit. Biochemistry 27:7633–7640

Ohgushi M, Wada A (1983) "Molten-globule state": a compact form of globular proteins with mobile side-chains. FEBS Lett 164:21–24

Ptitsyn OB (1973) Stepwise mechanism of self-organisation of protein molecules. Dokl Akad Nauk SSSR 213:473–475

Pfeil W (1987) Is thermally denatured protein unfolded? The example of α-lactalbumin. Biochim Biophys Acta 911:114–116

Pfeil W, Bychkova VE, Ptitsyn OB (1986) Physical nature of the phase transition in globular proteins. FEBS Lett 198:287–291

Ptitsyn OB, Pain RH, Semisotnov GV, Zerovnik E, Razgulyaev DI (1990) Evidence for a molten globule state as a general intermediate in protein folding. FEBS Lett 262:20–24

Robson B, Pain RH (1971) Analysis of the code relating sequence to conformation in proteins: possible implications for the mechanism of formation of helical regions. J Mol Biol 58:237–259

Robson B, Pain RH (1976) The mechanism of folding of globular proteins. Equilibria and kinetics of conformational transitions of penicillinase from *Staphylococcus aureus* involving a state of intermediate conformation. Biochem J 155:331–344

Roder H, Elöve GA, Englander SW (1988) Structural characterisation of folding intermediates in cytochrome *c* by *H*-exchange labelling and protein NMR. Nature 335:700–704

Semisotnov GV, Rodionova NA, Kutyshenko VP, Ebert B, Blanck J, Ptitsyn OB (1987) Sequential mechanism of refolding of carbonic anhydrase *B*. FEBS Lett 224:9–13

Teschner W, Rudolph R, Garel JR (1987) Intermediates on the folding pathway of octopine dehydrogenase from *Pecten jacobaeus*. Biochemistry 26:2791–2796

Thomas RM, Feeney J, Nicholson RB, Pain RH, Roberts GCK (1983) Identification by NMR spectroscopy of a stable intermediate in the unfolding of staphylococcal β-lactamase. Biochem J 215:525–529

Udgaonkar JB, Baldwin RL (1988) NMR evidence for an early framework intermediate on the folding pathway of ribonuclease A. Nature 335:694–699

Zerovnik E, Pain RH (1986) Refolding of β-lactamase followed by FPLC. Protein Eng 1:248