INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
I.C.T.P., P.O. BOX 586, 34100 TRIESTE, ITALY. CABLE CENTRATOM TRIESTE

H4.SMR/775-14

COLLEGE IN BIOPHYSICS: EXPERIMENTAL AND THEORETICAL ASPECTS OF BIOMOLECULES

26 September - 14 October 1994

Miramare - Trieste, Italy

Protein Dynamics III

William A. Eaton
National Institutes of Health
Bethesda, Maryland - USA

6

The Molten Globule State

OLEG B. PTITSYN

1 INTRODUCTION

There are two main problems in the physics of proteins: how amino acid sequences encode the three-dimensional (3D) structures of native proteins and how protein chains fold into their native structures in spite of the astronomically large number of alternatives.

It is very likely that the key to both of these problems is the hierarchy of structural levels in native proteins. The most important generalization that follows from the 30 years of X-ray studies of protein structures is probably that they can be divided into three levels: *secondary structure* (the positions of α-helices and β-strands along the chain), *"protein fold"* or *"folding pattern"* (the mutual positions of the α-helices and β-strands in 3D-space), and *tertiary structure* (a set of coordinates of the protein atoms). It is extremely important that only the third level is really unique for each protein, while the first and second ones can be similar or even identical for many related proteins. In fact, there are typical folds ("globin fold," "immunoglobulin fold," "trypsin fold," and so on) that are common for protein families, in spite of large differences in their amino acid sequences. This suggests that only the third level is determined by all the details of the amino acid sequence, while the codes for the first and second levels are highly degenerate. If this is the case, the problem of the "physical code" that connects protein 3D structures with their amino acid sequences can be solved in two steps: a general physical theory to predict the secondary structure and the tertiary fold, followed by conformational analysis to predict the atomic coordinates (tertiary structure) within the framework of this tertiary fold (Ptitsyn and Finkelstein, 1980a, b; Ptitsyn, 1985; Finkelstein and Ptitsyn, 1987; Ptitsyn et al., 1989).

The best experimental approach to study the levels of protein structure may be to "switch off" part of the interactions in the native protein in order to understand which type of structure can be preserved by the remaining interactions. To this end it is useful to study the physical states of a protein under different *denaturing conditions*, to identify and to investigate possible *equilibrium intermediates* between the native and the completely unfolded states.[1] A related approach is to investigate the *kinetics of protein folding*, to identify and to study *kinetic intermediates* in protein folding, and to compare them with the equilibrium intermediates.

The discovery of the equilibrium molten globule state of protein molecules (Dolgikh et al., 1981,1983,1985) and the identification of this state as a general kinetic intermediate in protein folding (Dolgikh et al., 1984; Semisotnov et al., 1987; Ptitsyn et al., 1990) should be helpful in this respect. This chapter is devoted to the description of the molten globule state, which is also reviewed by Ptitsyn (1987), Kuwajima (1989), and Dill and Shortle (1991). Section 2 of this chapter outlines briefly the history of the idea of protein equilibrium intermediates. Section 3 describes the properties of the equilibrium molten globule state. A model of the molten globule that attempts to comprise all the available experimental data is presented in Section 4. Section 5 describes the transitions between three main conformational states of protein molecules: native, molten globule, and unfolded. Section 6 contains a list of proteins and the conditions under which something like the molten globule state has been observed. Section 7 describes the role of the molten globule state in protein folding and the properties of the kinetic molten globule state, which are similar to those of the equilibrium molten globule. In Section 8, the possible role of the molten globule state in the living cell is discussed briefly.

2 BACKGROUND

More than 20 years ago, Tanford (1968) summarized all available data on nonnative (or denatured) physical states of protein molecules. He came to the conclusion that at high concentrations of strong denaturants (GdmCl or urea) proteins are more or less completely unfolded, although even in these cases some fluctuating structures cannot be excluded. (See also Dill and Shortle, 1991.) Other denaturing conditions often lead to only "partly unfolded" states. Denaturation of a number of proteins by high temperatures

[1]The term *native* is used here for the rigid state of a protein under physiological conditions in which the majority of the atoms are more or less fixed. The term *completely unfolded* is less definite, because an unfolded, statistical coil molecule can have a large number of conformations that depend upon the solvent and the temperature (Flory, 1953). The term *completely unfolded* will be used for all these states to distinguish them from *partly unfolded* states in which some residual structure can be observed.

and/or low pH (as well as by adding different salts) leads to remarkably small changes in their hydrodynamic and optical parameters, as compared to their denaturation by high concentrations of GdmCl.

Some of the differences between proteins denatured by different agents can be explained simply by the solvent- or temperature-dependencies of the physical properties of unfolded molecules, but many of them clearly need other explanations. For example, the denaturation of serum albumin by moderate pH (3.5) leads to a very small increase in its intrinsic viscosity, from 3.7 to 4.5 cm^3/g (Foster, 1960), which suggests that the protein remains globular under these conditions. The most definitive evidence for the presence of residual structure in acid- and temperature-denatured proteins is that they can undergo another cooperative transition under the influence of GdmCl (Aune et al., 1967) or urea (Brandts and Hunt, 1967). Thus, the existence of "partly folded" denatured states of proteins was well established as early as the late 1960's, although their nature was not known at that time.

Later, the existence of residual structure in temperature- and pH-denatured proteins was questioned by Privalov (1979) by a very convincing argument that the changes of enthalpy, entropy, and heat capacity upon denaturation are basically the same for temperature-, pH-, or GdmCl-induced denaturations. These data led him to the conclusion that temperature- and pH-denatured proteins do not contain stable residual structure and are nearly as unfolded as GdmCl- or urea-denatured proteins (Privalov, 1979). Especially informative is the change in heat capacity of denaturation, which is due mainly to the exposure of nonpolar side chains to water. (See Privalov, 1979, Chapter 3.) The careful measurements of Spolar et al. (1989) and especially of Privalov and Makhatadze (1990) have shown that the changes in heat capacity upon temperature- and acid-denaturation of apomyoglobin, cytochrome *c*, RNase A, and lysozyme are those expected if all their nonpolar groups are exposed to water, as would be the case for completely unfolded chains (Chapter 3).

The temperature melting of small proteins (Privalov, 1979) and protein domains (Privalov, 1982) is well known from the work of Privalov and his collaborators to be an "all-or-none" transition, i.e., a transition without any intermediate states apparent (Chapter 3). Together with the evidence for temperature-denatured proteins being unfolded, it suggested that there are no stable states of protein molecules that are intermediate between the native and the unfolded states. This point of view became popular in spite of hydrodynamic and optical data that showed that temperature-denatured proteins are far from being completely unfolded and that they can undergo another cooperative transition when GdmCl or urea is added.

Strong evidence for the existence of equilibrium intermediates between the native and unfolded states came from the studies of protein

unfolding by GdmCl, especially on the unfolding of bovine (Kuwajima et al., 1976) and human (Nozaka et al., 1978) α-lactalbumins. These proteins undergo two different conformational transitions when GdmCl is added. The first transition (at lower concentrations of the denaturant) causes a drastic decrease of CD in the "aromatic" (near-UV) region, i.e., destruction of the rigid environment of the aromatic groups in the fully folded state. The second transition (at higher concentrations of the denaturant) produced a drastic decrease of CD in the "peptide" (far-UV) region, i.e., destruction of the secondary structure. Similar data had been obtained even earlier for bovine carbonic anhydrase B (Wong and Tanford, 1973), and for growth hormones (Holladay et al., 1974). These observations demonstrated the existence of one or more equilibrium intermediate states that differ from the native state by the absence of a rigid environment of aromatic side chains, but differ from the unfolded state by the presence of secondary structure. Moreover, the optical properties of bovine growth hormone (Burger et al., 1966), bovine carbonic anhydrase B (Wong and Hamlin, 1974), and bovine (Kuwajima et al., 1976) and human (Nozaka et al., 1978) α-lactalbumins at acid pH are similar to the states observed at intermediate concentrations of GdmCl. The absence of rigid tertiary structure and the presence of secondary structure in these intermediates lead Kuwajima (1977) to model these intermediates as unfolded, noncompact molecules with local secondary structure.

The equilibrium intermediates described in those papers seemed to be of key importance for the whole problem of protein structure and protein folding. Therefore, we have studied the intermediate forms of bovine and human α-lactalbumins (Dolgikh et al., 1981,1985; Gilmanshin et al., 1982; Ptitsyn et al., 1983,1986; Pfeil et al., 1986; Damaschun et al., 1986; Gast et al., 1986; Timchenko et al., 1986; Bychkova et al., 1990; Semisotnov et al., 1991a) and bovine carbonic anhydrase B (Dolgikh et al., 1983; Ptitsyn et al., 1983; Brazhnikov et al., 1985; Rodionova et al., 1989; Semisotnov et al., 1989,1991a) by a wide variety of physical methods. These studies revealed a novel physical state of protein molecules that has been named the "molten globule" by Ohgushi and Wada (1983). Similar states have been observed later in a number of other proteins. (See Section 6.)

3 PROPERTIES OF THE MOLTEN GLOBULE

The physical properties of the molten globule state will be described using bovine and human α-lactalbumins and bovine carbonic anhydrase B as examples, because these three proteins have been studied more carefully than have others. Other proteins for which data are not as complete will be described briefly in Section 6.

3.1 Compactness

One of the most important physical properties of the molten globule state is that it is almost as compact as the native protein. Table 6–1 summarizes the corresponding data for α-lactalbumins obtained from their diffusion and sedimentation coefficients. The very good coincidence between three different methods, after correction for small amounts of aggregation, indicates that the hydrodynamic radius of α-lactalbumin increases in the molten globule state by 14 ± 2% as compared to the native state; this corresponds to a volume increase of 50 ± 8%. For comparison the hydrodynamic radii of these proteins in the unfolded state (with intact S-S bridges) are increased by 49 ± 5%, which corresponds to a volume 3.3 ± 0.3 times larger than that of the native molecule. The large difference between the volume expansion in the molten globule state (1.5) and in the unfolded state (3.3) clearly shows that the molten globule state is relatively compact.

Data on the intrinsic viscosities of α-lactalbumins (Dolgikh et al., 1981,1985) and of carbonic anhydrase (Wong and Hamlin, 1974; Dolgikh et al., 1983), as well as on the radius of gyration of bovine α-lactalbumin (Dolgikh et al., 1981,1985; Izumi et al., 1983; Timchenko et al., 1986), are also consistent with a small expansion of protein molecules in the molten globule state.

3.2 Internal Water

The approximately 50% increase in a protein's volume in the molten globule state suggests that water can penetrate inside; indeed, there must be several hundred such molecules in the case of α-lactalbumin. Several experimental data support this point of view:

Partial specific volume. If the molten globule were "empty" (i.e., did not contain water) its partial specific volume would be much greater (~50%) than that of the native (or of the unfolded) protein. Although no careful comparison of the partial specific volumes of molten globule, unfolded, and native proteins has been carried out, there is little doubt that this is not the case.

Heat capacity. As mentioned in Section 2, the change in protein heat capacity upon denaturation reflects mainly the exposure to aqueous solvent of nonpolar side chains. No careful comparison of the heat capacities of molten globule and other denatured proteins has been performed. There is no doubt, however, that the heat capacity of a molten globule is larger than that of native proteins, which suggests that some nonpolar groups buried in the native protein become exposed to water in the molten globule state.

TABLE 6–1. Diffusion and Sedimentation Constants of α-lactalbumins in Different States

	State	Condition	M_{app}[a]/M	$D_{20,w} \times 10^7$, cm^2/sec[c] QLS	$D_{20,w} \times 10^7$, cm^2/sec[c] PI	$s_{20,w}$, S[e]	R/R_N[f] $D_{20,w}$ QLS	R/R_N[f] $D_{20,w}$ PI	R/R_N[f] $s_{20,w}$
Bovine α-lactalbumin	Native	pH 7.5; 20°C	1.05	—	11.8 (11.6)[d]	1.63 (1.67)			
	Molten globule	pH 2.0; 20°C	0.96	—	10.1	1.40	—	1.17	1.17
	Unfolded	6.4 M GdmCl pH 7.5; 20°C	—	—	7.7	—	—	1.53	—
Human α-lactalbumin	Native	pH 7.5; 20°C	1.04	12.3 (12.1)	12.5 (12.3)	1.62 (1.64)	—	—	—
	Molten globule	pH 2.0; 20°C	1.16	10.8 (11.0)	11.1 (10.5)	1.40 (1.56)	1.14	1.13	1.11
	Molten globule	Ca^{2+}-free (10 mM EDTA) pH 7.5; 50°C	1.40[b]	10.9 (9.2)	—	—	1.13	—	—
	Unfolded	6.0 M (QLS) or 6.4 M (PI) GdmCl; pH 7.5; 20°C	—	8.6	8.3	—	1.43	1.51	—

[a] M_{app}: apparent molecular weights measured by equilibrium sedimentation (Gilmanshin et al., 1982); M: molecular weight calculated from amino acid content.

[b] Evaluated from M_{app} of Ca^{2+}-free form at 20°C and temperature-dependence of light scattering (Gast et al., 1986).

[c] Diffusion coefficients measured by quasi-elastic light scattering QLS (Gast et al., 1986) and by polarization interferometer PI (Bychkova et al., 1990).

[d] Values of diffusion and sedimentation coefficients have been corrected for slight aggregation, using a method analogous to that published earlier (Gast et al., 1986). Noncorrected values are shown in brackets. No corrections have been made for unfolded proteins in 6.0–6.4 M GdmCl, because proteins do not aggregate under these conditions.

[e] Sedimentation coefficients measured by Gilmanshin et al. (1982).

[f] R/R_N — ratio of the hydrodynamic radii to its value in the nature (N) state obtained from diffusion (D) and sedimentation (s) coefficients as $R_D/(R_D)_N = D_N/D$ and $R_s/(R_s)_N = s_N/s$. The correction for a probable change of asymmetry of α-lactalbumins in the molten globule state (Timchenko et al., 1986) is small and does not change the results significantly.

3.3 Core

There is clear evidence, on the other hand, that the molten globule has a relatively dense core. In native proteins, many atom pairs are in van der Waals contact at a distance of ~4.5 Å; consequently, diffuse X-ray scattering by these proteins has a maximum corresponding to this Bragg distance (Echols and Anderegg, 1960; Grigoryev et al., 1971; Gernat et al., 1986). This maximum is due in part to van der Waals contacts in secondary structure but is mainly due to long-range contacts in the tertiary structure (Fedorov and Ptitsyn, 1977); it is absent or weak in unfolded or helical polypeptides (Damaschun et al., 1986; Gernat et al., 1986).

Figure 6–1 demonstrates that this maximum is present also in the acid molten globule state of human α-lactalbumin (Damaschun et al., 1986). The only difference is that it is shifted from 4.50 to 4.65 Å, i.e., by 3% to 4%. This small shift is similar to the greater contact distances that are present in typical liquids, as compared to crystals. A similar result has been observed in the heat-denatured molten globule state of this protein (Ptitsyn et al., 1986; Ptitsyn, 1987). This indicates that the nonpolar groups in the protein core remain in contact in the molten globule, although their packing is not as tight as in the native protein. There is also NMR evidence for the presence of a cluster of aromatic groups in the molten globule state of guinea pig α-lactalbumin (Baum et al., 1989). (See Section 3.5.)

3.4 Secondary Structure

Circular dichroism spectra of α-lactalbumins (Kuwajima et al., 1976; Nozaka et al., 1978; Dolgikh et al., 1981, 1985) and carbonic anhydrases (Wong and Tanford, 1973; Wong and Hamlin, 1974; Jagannadham and Balasubramanian, 1985; Bolotina, 1987; Rodionova et al., 1989) in the far-UV region are very pronounced in various molten globule states (acid pH, high temperature, or ~2 M GdmCl) and suggest a high content of secondary structure. The CD spectra are not the same as those of the native proteins, but they usually are even more pronounced. For example, in human carbonic anhydrase B the negative molar ellipticity at 210 nm is nearly four times greater in the acid molten globule than in the native state (Jagannadham and Balasubramanian, 1985). This does not necessarily mean a change in secondary structure, however, as far-UV CD spectra can be influenced by aromatic side chains. These side chains can contribute to CD spectra not only in the near-UV, but also in the far-UV region (Sears and Beychok, 1973; Manning and Woody, 1989). This contribution may be especially large for clusters of aromatic groups.

Aromatic groups lose their rigid environment in the molten globule state (see Section 3.6), so the contribution of these groups to far-UV CD must

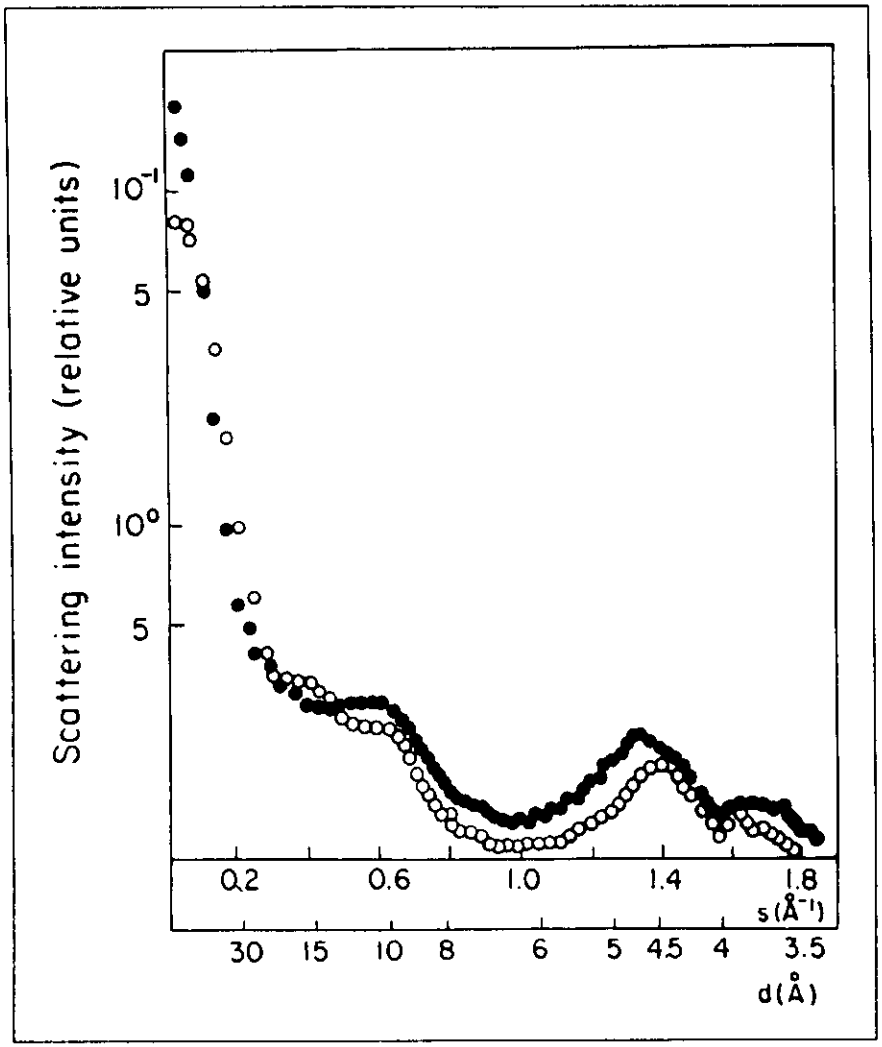


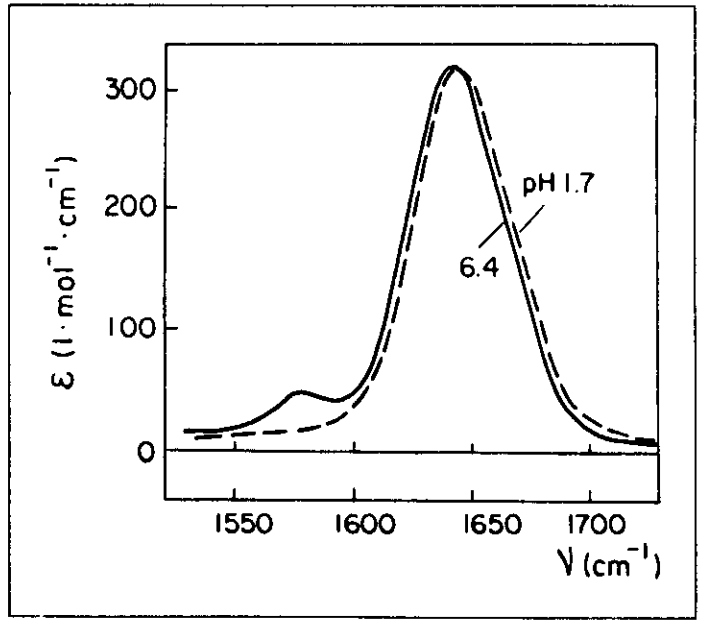
FIGURE 6–1. Diffuse X-ray scattering curves of human α-lactalbumin in the native (○) and in the acid-denatured, molten globule (●) states, measured at protein concentrations of ~10% (w/v). s = (4π/λ) sinΘ is the value of scattering vector (wavelength, λ = 1.54 Å, 2Θ-scattering angle), d = 2π/s is the Bragg distance. The shift of the maximum from s = 1.40 Å^{-1} for the native protein to s = 1.35 Å^{-1} for the molten globule corresponds to an increase of d from 4.49 to 4.65 Å. (Adapted from Damaschun et al., 1986)

vanish or at least be greatly reduced, which can lead to the large changes of far-UV optical properties (Kronman et al., 1966). It is worthwhile to note that aromatic groups can have both positive and negative bands in the far-UV region (Brahms and Brahms, 1980; Manning and Woody, 1989); therefore, far-UV CD spectra of the molten globule state may be either more or less pronounced than those of the native protein.

An approximate method for the decomposition of far-UV CD spectra into the contributions of peptide and aromatic groups (Bolotina and Lugauskas, 1985) suggests that these spectra of the molten globule states of α-lactalbumins and carbonic anhydrase differ from those of the native proteins mainly by the contribution of aromatic groups (Bolotina, 1987).

More precise information can be obtained from infrared spectra, as in this case the contribution of side chains can be measured and subtracted from the total absorption (Chirgadze et al., 1975; Venyaminov and Kalnin, 1991). Figure 6–2 shows that the infrared spectra of bovine α-lactalbumins in the amide I region are almost identical for the native and molten globule states (Dolgikh et al., 1985). Very small differences between these spectra exclude any changes in content of α-helices and β-structure of greater than

FIGURE 6–2. Infrared spectra of bovine α-lactalbumin in the amide I region of the native (———) and acid-denatured molten globule (– – –) states (after the subtraction of the side chain contributions). (Adapted from Dolgikh et al., 1985)



10% and 5%, respectively. Infrared spectra of native and of acid-denatured bovine carbonic anhydrase B have different shapes, but their decomposition into the contributions of β- and unfolded structures gives nearly identical β-content: 39% for the native form and 37% for the molten globule state (Dolgikh et al., 1983; see also Brazhnikov et al., 1985; Ptitsyn, 1987). The difference in the shapes of the spectra can be due to the broadening of the main β-band, which suggests some disordering of the β-structure.

3.5 Native-like Structural Organization

Definitive evidence that at least some α-helices of α-lactalbumin are located in their native positions along the polypeptide chain in the molten globule came from NMR studies of hydrogen exchange (Baum et al., 1989· Dobson et al., 1991). In these experiments, the protein was first allowed to exchange in the molten globule state for a given time; it was then transformed into the native state, and 2D-NMR was used to identify those NH protons that were protected from exchange in the molten globule state. In this way, Dobson and his collaborators (Baum et al., 1989; Dobson et al., 1991) have shown that the NH-protons of the B- and C-helices in guinea pig α-lactalbumin (see Figure 6–3) are protected in the acid molten globule state. It is almost certain that protected protons are involved in intramolecular hydrogen bonds and are shielded from solvent (e.g., Udgaonkar and Baldwin, 1988). Therefore, at least two α-helices of native α-lactalbumin (B, residues 23 to 34, and C, residues 86 to 99) are present also in the molten globule state. These α-helices contribute to a common nonpolar core in the native protein (Acharya et al., 1989), and there is the intriguing possibility that this may also be the case in the molten globule (Dobson et al., 1991). On the other hand, no evidence for a native-like structure has been found thus far for the other subdomain of this protein, which has predominantly β-structure in the native state.

In a similar way, Baldwin's group has shown the existence of at least three native α-helices in the molten globule state of sperm whale apomyoglobin (Hughson et al., 1990). They have measured individual exchange rates for many protons and have shown that the NH-protons of helices A, G, and H exchange 5 to 200 times more slowly than expected for the unfolded state (although much faster than in the native state). The NH-protons of helix B exchange only 2 to 10 times more slowly than in the unfolded protein, which suggests that this helix is only partly folded or is unstable. No traces of protection have been observed for helix E, which therefore seems to be unfolded. The NH-protons of helix F and of the short helices C and D could not be studied. These observations lead to the attractive idea that the compact unit formed by helices A, G, and H in native myoglobin may remain in the molten globule state (Hughson et al., 1990).

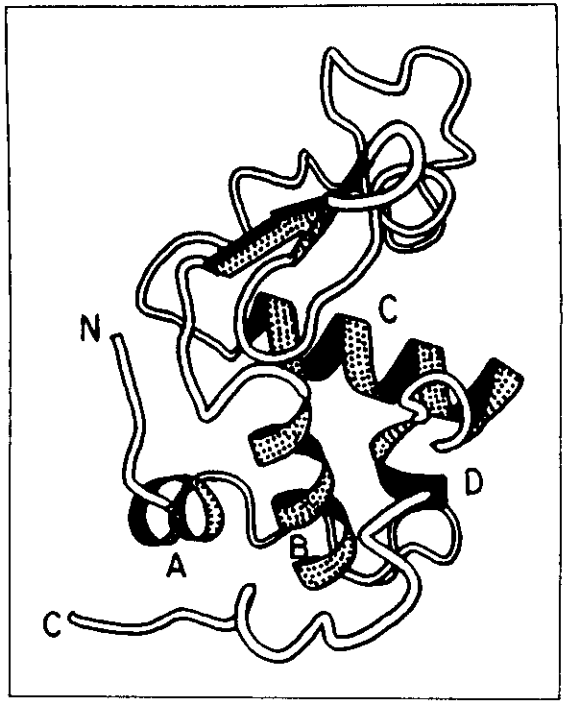


FIGURE 6–3. The structure of α-lactalbumin, showing the locations of the major α-helices and β-strands. (Adapted from Acharya et al., 1989)

Similar results have been obtained for cytochrome *c* (Jeng et al., 1990). In this protein, the main parts of all three large native α-helices (residues 6 to 14, 60 to 69, and 87 to 102) (Bushnell et al., 1990) remain protected from hydrogen exchange in the molten globule state at acid pH and high ionic strength.[2] This protection is, of course, much weaker than in the native state. On the other hand, many NH groups that are involved in hydrogen bonds in reverse turns and in the tertiary structure of the native protein are not protected substantially in the molten globule state.

A direct approach to the study of the tertiary fold of the molten globule would be by 2D-NMR spectra. NMR spectra of the molten globule state are very poorly resolved (see Section 3.6), however, and cannot be

[2]The existence of the molten globule state of cytochrome *c* at acid pH and high ionic strength was first claimed by Ohgushi and Wada (1983). Later it was suggested (Ptitsyn, 1987) that this state is native-like (differing only by the change of the heme state) as it melts cooperatively upon heating (Potekhin and Pfeil, 1989). However, recent experimental data presented by Jeng et al. (1990) and by Goto et al. (1990a) suggest that the acid form of cytochrome *c* at high ionic strength really has many properties typical of the molten globule state.

assigned by conventional methods. In Dobson's group, such assignments have been partly achieved by magnetization transfer. If the rate of interconversion between the native and the molten globule state in the transition region is faster than the nuclear relaxation rates, magnetic saturation of resonances of a given residue in the native state can be transferred to the resonances of the same residue in the molten globule. This can permit the "transfer" of the assignment of the NMR spectrum of the native state to that of the molten globule. Using this approach, Dobson's group (Baum et al., 1989; Dobson et al., 1991) has succeeded in correlating the strongly perturbed aromatic resonances of guinea pig α-lactalbumin in the acid molten globule state of those in the native state. The result was that the resonances of four aromatic residues, which are assigned tentatively to Trp 26, Phe 31, Tyr 103, and Trp 104, have the largest NMR chemical shifts in both the native and the molten globule states. The corresponding residues (two of them belong to α-helix B, residues 23 to 34, and two others are between α-helices C, residues 86 to 99, and D, residues 105 to 109; see Figure 6–3) are close together in the native structure of baboon α-lactalbumin (Acharya et al., 1989). It is very likely that this cluster of aromatic groups exists also in the molten globule state.

Recently, the molten globule state has been obtained for ubiquitin in 60% methanol at pH 2 (Harding et al., 1991). In this case, the NMR spectrum of the molten globule state was relatively sharp and permitted assignment of a number of resonances. As a result, it was shown that three β-strands of the β-sheet of the native protein are also present in the molten globule state, both in their positions in the polypeptide chain and in their mutual positions in space. The single α-helix of ubiquitin is also present and probably has a native-like position relative to the β-sheet. The other parts of the native conformation of this small protein, including two small additional β-strands, could not be detected in the molten globule state.

3.6 Fluctuating Environment of Side Chains

The data just presented may create an impression that a protein in the molten globule state is "almost native." The other properties of the molten globule state, however, are quite different from those of the native protein and are similar to those of the unfolded state.

^{1}H-NMR spectra of acid- and temperature-denatured forms of bovine (Ptitsyn et al., 1983; Dolgikh et al., 1985; Ikeguchi et al., 1986; Kuwajima et al., 1986) and guinea pig (Baum et al., 1989; Dobson et al., 1991) α-lactalbumins are quite different from those of the native proteins and are more similar to the spectra of the unfolded proteins. This is illustrated by Figure 6–4, which shows that the NMR spectrum of the molten globule state is much simpler than that of the native protein; the number of perturbed

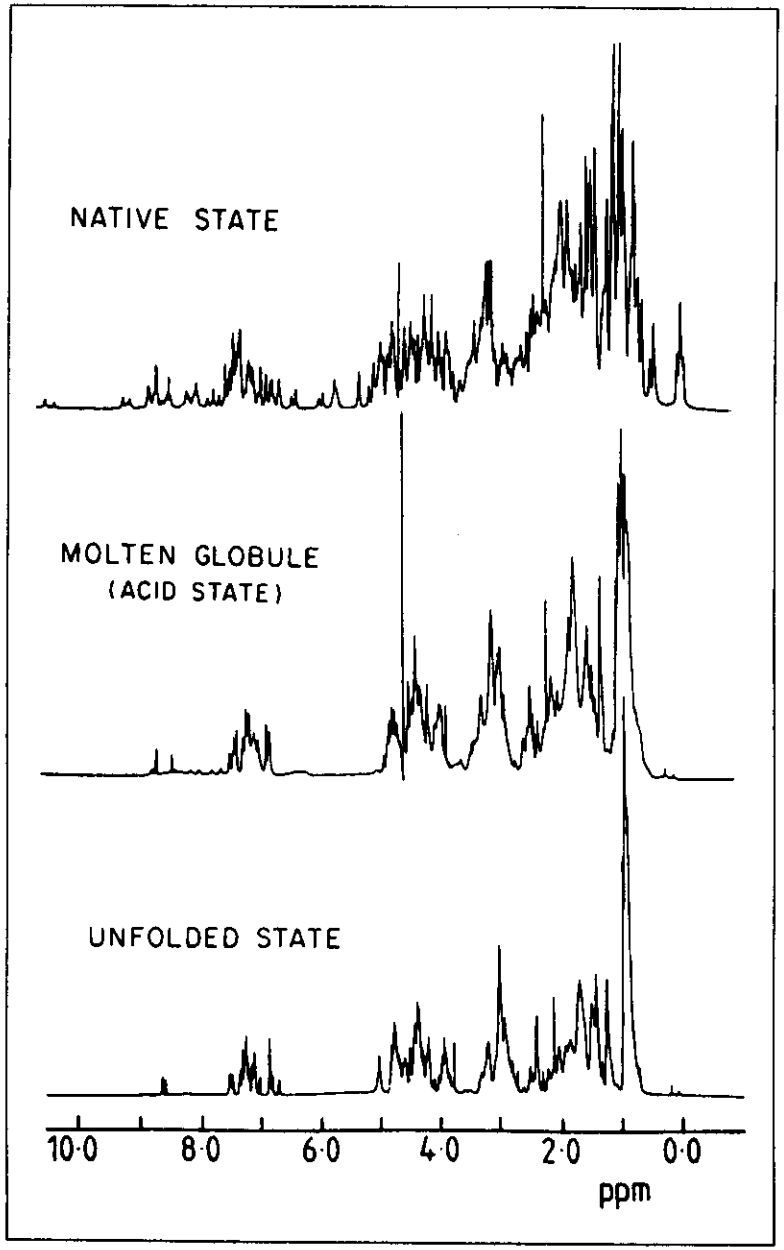


FIGURE 6–4. Five hundred–MHz ^{1}H-NMR spectra of guinea pig α-lactalbumin in the native (pH 5.4), acid (pH 2.0), and unfolded (in 9-M urea) states recorded at 52°C. (Adapted from Baum et al., 1989)

resonances is much smaller, and the overall picture is much more similar to the unfolded state. Note particularly that some pronounced high field resonances (at <1 ppm) of the native protein almost completely disappear in the molten globule state. This shows that the rigid mutual arrangement of aromatic and aliphatic side chains in the protein core is largely destroyed. On the other hand, a number of resonances are still remarkably perturbed, which reflects traces of structure and distinguishes the NMR spectrum of the molten globule state from that of the unfolded state. These are the perturbed resonances that have been used by Baum et al., (1989) for a tentative assignment of the molten globule NMR spectrum and for the conclusion of the existence of the native-like aromatic cluster in the molten globule state. (See Section 3.5.)

^{1}H-NMR spectra of the acid molten globule states of bovine carbonic anhydrase B (Rodionova et al., 1989; Semisotnov et al., 1989; Ptitsyn, 1987) and human retinol-binding protein (Bychkova et al., 1992) are also much more similar to those of the unfolded than of the native state.

In a similar way, the near-UV CD of α-lactalbumins (Kuwajima et al., 1976; Nozaka et al., 1978; Dolgikh et al., 1981,1985), carbonic anhydrase (Wong and Tanford, 1973; Wong and Hamlin, 1974) and other proteins is practically absent or greatly reduced in the molten globule state, suggesting that the local environment of the aromatic side chains is much more flexible.

3.7 Side Chain Movements

NMR and near-UV CD measurements show that the environment of many side chains is much less rigid in the molten globule than in the native state. This means that the fluctuations of side chains in the molten globule state are greatly increased as compared with the native state. This conclusion has been confirmed by direct experiments on spin-lattice and spin-spin relaxation times (T_1 and T_2) for bovine carbonic anhydrase B in the native, molten globule, and unfolded states (Semisotnov et al., 1989).

The spin-spin relaxation time decreases with an increase in the ratio of molecular movements. Figure 6–5A illustrates that the spin-spin relaxation on time T_2 of methyl groups in the molten globule state coincides with that of the unfolded state (0.085 sec) while it is quite different for the native state (0.021 sec).[3]

The molecular volumes of the molten globule and native states are similar (Wong and Hamlin, 1974), so the main reason for the difference in their relaxation times must be intramolecular movements. The similarity of the spin-spin relaxation times for the molten globule and for the unfolded

[3]Spin-spin relaxation time decreases with the increase of times of molecular movements.

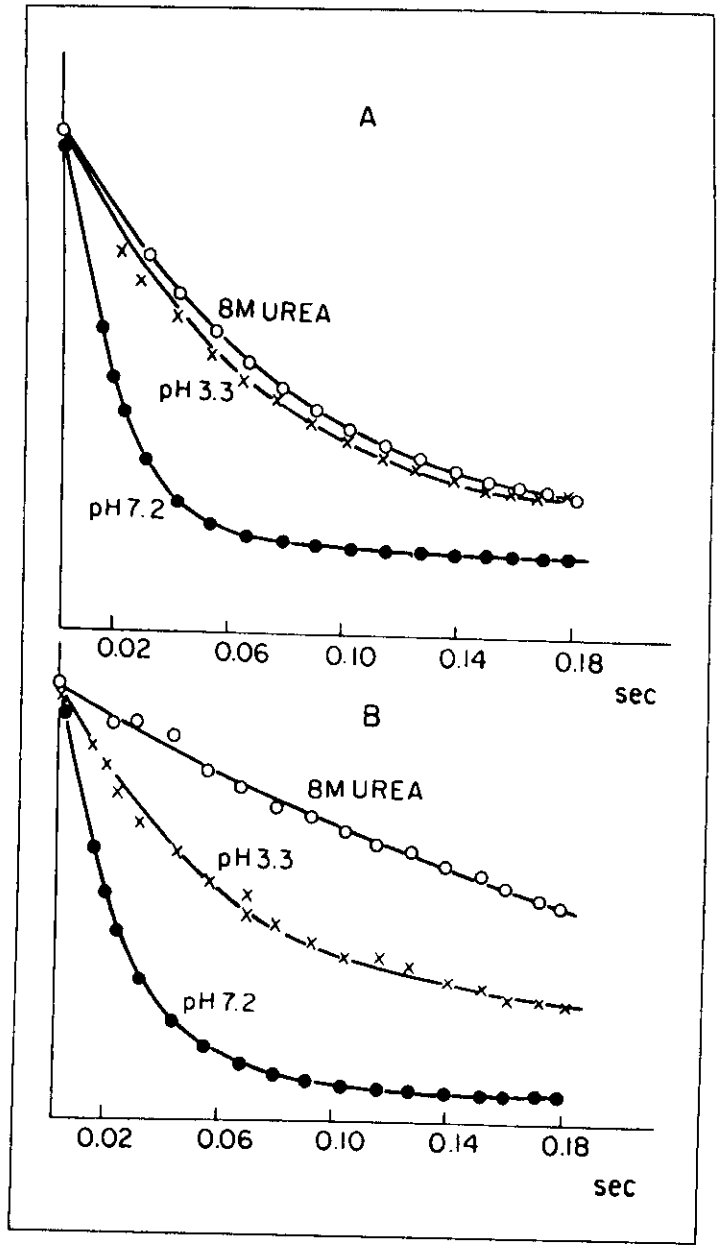


FIGURE 6–5. Time-dependence of the spin echo amplitudes for protons of the methyl (A) and aromatic (B) groups of bovine carbonic anhydrase B in the native (pH 7.2), molten globule (pH 3.3), and unfolded (in 8 M urea) states. Spin echo of the methyl groups was measured at 0.89 ppm and of the aromatic groups at 7.1 ppm. (Adapted from Semisotnov et al., 1989)

states suggests that the intramolecular movements for methyl nonpolar groups in the molten globule and in the unfolded states are practically the same.

On the other hand, spin echo curves for aromatic groups in the molten globule state are intermediate between those in the native and in the unfolded states (Fig. 6–5B), and the spin-spin relaxation time for these protons in the molten globule state (0.062 sec) is much closer to those in the native state (0.031 sec) than those in the unfolded protein (0.22 sec). Therefore, intramolecular movements of aromatic side chains (unlike aliphatic ones) are much more hindered in the molten globule than in the unfolded state, although not as hindered as in the native state. It appears that there is sufficient space inside the molten globule state for the "free" movements of small and symmetric aliphatic groups, but not enough for the movements of larger and less symmetric aromatic groups (see also Shakhnovich and Finkelstein, 1989).

Hindering of the motion of aromatic side chains in the molten globule state is confirmed by the study of urea denaturation of bovine carbonic anhydrase B (Rodionova et al., 1989). The proton resonances of the aliphatic side chains change at the first stage of denaturation (together with near-UV CD), i.e., at the transition from the native to the molten globule state. The proton resonances of aromatic side chains change at the second stage of denaturation (together with the decrease in far-UV CD and compactness), i.e., at the transition from the molten globule to the unfolded state. (See Figure 6–7B.)

Polarization of luminescence of Trp residues also shows that intramolecular mobilities of the Trp indole rings in the molten globule states of bovine and human α-lactalbumins (Dolgikh et al., 1981,1985) and in bovine carbonic anhydrase (Rodionova et al., 1989) are nearly as restricted as in the native state, while the restriction in the unfolded state is much less.

3.8 Large-scale Movements

Restricted large-scale fluctuations in the molten globule state are reflected in the field-dependent broadening of individual resonances in ^{1}H-NMR spectra. This broadening is especially apparent for aromatic groups and has been observed in the acid molten globule state of guinea pig α-lactalbumin (Baum et al., 1989). This is consistent with interconversion of different local conformations of the molten globule at rates slower than ~10^3 sec^{-1}.

Another approach to the study of large-scale fluctuations is the accessibility of the internal parts of protein molecules to the solvent or to other molecules. The most common technique used in these studies is hydrogen-deuterium exchange. Measurements of the overall rate of this exchange for the NH-protons of the backbone of bovine α-lactalbumin (Dolgikh et al., 1981,1985) and of bovine carbonic anhydrase B (Dolgikh et al., 1983) have shown that deuterium exchange is much faster in the molten globule than in the native protein and is closer to that in unfolded chains. More detailed studies of deuterium exchange using 2D-NMR (see Section 3.5) have shown that NH groups of many α-helices are still protected in the molten globule state (although much less than in the native state), while the NH groups of loops are not protected (Baum et al., 1989; Hughson et al., 1990; Jeng et al., 1990; Dobson et al., 1991; Baldwin, 1991).

The accessibility of a protein molecule to proteases also increases in the molten globule state. For example, digestion of channel peptide of colicin E1 by papain and bromelain is much faster in the molten globule state than in the native one (Merrill et al., 1990).

3.9 Stability

The small increase in protein volume in the molten globule should be sufficient to destroy the tight packing of side chains, which should lead to a large decrease in their van der Waals attractions. Even though being packed more loosely ("liquid-like"), nonpolar groups still preserve their attraction in an aqueous environment. It was suggested (Shakhnovich and Finkelstein, 1982; Gilmanshin et al., 1982; Dolgikh et al., 1985), therefore, that the molten globule state is stabilized mainly by the liquid-like interactions of nonpolar groups in water that are often called hydrophobic interactions.

This idea has been confirmed recently by site-directed mutagenesis of apomyoglobins (Hughson and Baldwin, 1989; Hughson et al., 1991; Baldwin, 1991). These experiments have shown that an increase in side chain hydrophobicity (replacements of Cys → Leu, Ala → Leu, Phe → Trp, Ser → Leu, Ser → Phe) stabilizes the molten globule state against unfolding, while the same mutations almost always destabilize the native state. This suggests that the molten globule state is stabilized mainly by less specific hydrophobic interactions (i.e., liquid-like interactions of nonpolar groups), while specific tight packing is important for the native state.

It is important to note that although the molten globule state is already "molten" (as it has lost the tight packing of side chains—see also Sections 4 and 5), it still preserves many elements of structure that can be melted by the influence of stronger denaturants. First, α-helices and other types of secondary structure certainly are preserved in the molten globule (see Sections 3.4 and 3.5) and the enthalpy of the helix-coil transition in water is as large as ~1 kcal/mol (Finkelstein et al., 1991; Baldwin, 1991). Second, the enthalpy of hydrophobic interactions is small at 20°C, but greatly increases with an increase in temperature and thus can provide an important contribution to heat effects of protein unfolding (Baldwin, 1986). It is not possible, therefore, to exclude that unfolding of the molten globule state may be accompanied by measurable heat effects, especially at high temperatures.

Unfortunately the available experimental data do not provide much information on this important point. In all proteins that have been studied, the molten globule state can be unfolded by GdmCl or urea in an apparently cooperative transition. (See Section 6, Table 6–2.) In one case —for human α-lactalbumin—this unfolding at 40°C is not accompanied by a large cooperative heat absorption (Pfeil et al., 1986). This suggests that the enthalpy of the molten globule state in this protein and at this temperature does not differ very much from the enthalpy of the unfolded state.

Unfolding of the molten globule state can also be caused by a further increase in electrostatic interactions. For example, the acid molten globule state of carbonic anhydrase can be unfolded by further decrease in the pH (Wong and Hamlin, 1974) and the acid molten state of cytochrome *c* by a decrease in the ionic strength (Ohgushi and Wada, 1983). Interesting effects occur when the ionic strength is varied by an increase of an acid (e.g., HCl) rather than a salt (e.g., KCl). Fink and his collaborators have shown that some proteins (e.g., β-lactamase, cytochrome *c*, and RNase A) first become fully or partly unfolded at acid pH (in the absence of salt), but with further decrease in pH they change into the molten globule state (Goto and Fink, 1989,1990; Goto et al., 1991a,b; Fink et al., 1990,1991). The explanation is simple: after deionization of all (or the majority of) the acid side groups, further decrease of the pH leads, not to an increase, but to a decrease of electrostatic interactions, due to the increase in the concentration of ions.

Data on the temperature melting of the molten globule state are controversial. The acid molten globule states of bovine (Dolgikh et al., 1985) and human (Pfeil et al., 1986) α-lactalbumins do not melt cooperatively upon heating. This can be expected because these proteins have similar molten globule states at low pH and at high temperatures (Dolgikh et al., 1981,1985; Gast et al., 1986). Microcalorimetric data on the acid form of retinol-binding protein (Bychkova et al., 1992) are consistent with a cooperative melting of the molten globule, with a small enthalpy change. Cytochrome *c* is in a state similar to the molten globule at low pH and high ionic strength (see Section 3.5) and melts upon heating in a manner similar to the native protein (Potekhin and Pfeil, 1989). It is unknown to what extent these proteins are unfolded at high temperatures.

4 MODEL OF THE MOLTEN GLOBULE

4.1 Initial Model

As early as 1973, the author predicted the molten globule state as a kinetic intermediate on the protein folding pathway (Ptitsyn, 1973). This state was specified as an "intermediate compact structure which is still different from the unique native protein structure and which formation is determined mainly

by non-specific interactions of amino acid residues with their environment (water and a hydrophobic core of a forming globule)," i.e., by hydrophobic interactions. This hypothesis implied that the intermediate compact structure is formed "by 'the merging' of pre-existing embryos of chain regions with secondary structure and that the mutual arrangement of chain regions in this structure satisfying the crude criterion of the maximal screening of hydrophobic side groups from contacts with water must be close to their arrangement stabilized by specific long-range interactions between space-neighboring residues in the final compact structure" (Ptitsyn, 1973). According to this hypothesis, therefore, the predicted intermediate structure is compact, has native-like secondary structure and a native-like tertiary fold, but differs from the native state by the lack of specific interactions between neighboring residues, i.e., by the lack of a rigid tertiary structure.

Section 3 of this chapter shows that the equilibrium "molten globule" state is compact, has native-like secondary structure, and differs from the native state by the lack of the rigid tertiary structure. There is some preliminary indirect evidence that the molten globule state also has a native-like tertiary fold. The molten globule appears therefore to be consistent with the model that was proposed in 1973.

There are some contradictions, however, between experimental data that need to be explained. (See, e.g., Baldwin, 1990.) One group of experimental data—on specific partial volume and heat capacity (see Section 3.2)—suggests that the molten globule must have a lot of water inside and is something like "an amorphous squeezed coil of fluctuating helices" (Griko et al., 1988). In more rigorous terms, these data are consistent with an unfolded chain below the Θ-temperature[4] (see Flory, 1953); i.e., a chain with a predominance of intramolecular attractions that stabilize its squeezed state by occasional contacts between attracting groups (Finkelstein and Shakhnovich, 1989).

Another group of experiments, however, strongly contradicts this model. Large angle diffuse X-ray scattering shows the existence of a protein core with many van der Waals contacts at distances only ~3% to 4% larger than in the native state. (See Section 3.3.) Moreover, there is evidence that at least some of the contacts of native structure remain also in the molten globule state. (See Section 3.5.)

The most plausible explanation for this contradiction is the assumption on a nonuniform expansion of a protein molecule in the molten globule state. This assumption was proposed several years ago (Ptitsyn et al., 1986; Damaschun et al., 1986) and will be presented in this section.

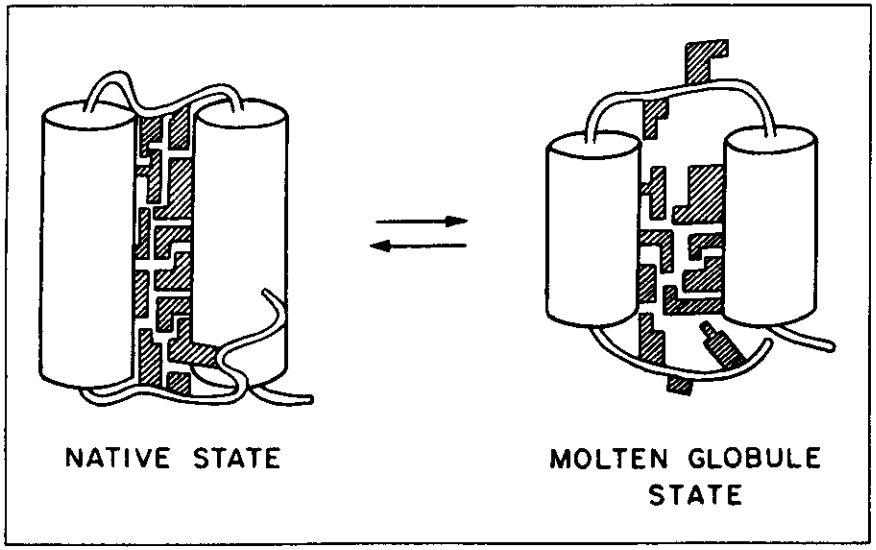
[4]The Θ temperature is that at which interactions between nonadjacent atoms of a polymer compensate for interactions between the polymer and the solvent, so that the polymer has the average dimensions expected for a random coil.

4.2 "Nonuniform" Model

The hypothesis of nonuniform expansion implies that the "frame" of the native protein is essentially preserved in the molten globule state, while other parts of the molecule can be more or less unfolded (Fig. 6–6). The frame includes central parts of the main α-helices and β-sheets, i.e., those parts of these regions that carry nonpolar side chains contributing to the hydrophobic core. The rest of the molecule can be unfolded and includes loops, ends of α- and β-regions, and those α- and β-regions that are not part of the frame of the protein.

According to this hypothesis, a protein core formed mainly by nonpolar groups of α- and β-regions must expand only slightly in the molten globule state, so that water molecules may not penetrate inside this core to any considerable extent. The packing of nonpolar side chains in the core becomes looser (as in usual crystal-liquid transitions), but at least a substantial part of these contacts is preserved. The looser packing of the side chains permits them to increase the amplitudes of their rotational vibrations; moreover, aliphatic side chains probably can even jump from one rotational isomer into another (Shakhnovich and Finkelstein, 1982, 1989). Therefore, this expansion will lead not only to an increase in the van der Waals energy of the core, but also to an increase in its entropy. In this (and only in this!) sense, the molten globule is liquid-like, as the native-like mutual arrangement

FIGURE 6–6. Schematic representation of the native and the molten globule states of protein molecules. Nonpolar side chains are hatched.



of α- and β-regions may be stabilized by liquid-like hydrophobic contacts between their nonpolar side chains. The point is that these side chains belong to rigid α- and β-regions, so their interactions are cooperative (Ptitsyn and Finkelstein, 1980b); this permits preservation of the native-like mutual arrangement of the α- and β-regions. "Liquid crystal" is probably the best way to describe this situation in usual molecular physics terms.

The important point of this hypothesis is that the stability of the molten globule state is due to "liquid-like" (hydrophobic) interactions in a relatively small part of the protein molecule. This possibility is justified by the fact that only a relatively small number of side chains are completely buried in globular proteins; e.g., in globins the number of these side chains is 33 out of approximately 150, or only ~20% (Perutz et al., 1965). As a rule, all, or almost all, of these residues are conserved as nonpolar in all members of the given protein family, e.g., in globins (Perutz et al., 1965). It is natural to assume that these 20% of the side chains (plus probably some nonpolar side chains that are only partly buried in the native state but enter or belong to a hydrophobic core) preserve their contacts in the molten globule state. The energies of these contacts, even when reduced due to the expansion of the core, seem sufficient to stabilize the α- and β-regions and their native-like arrangement.

The hypothesis of a nonuniform expansion probably can help to reconcile the apparently contradictory experimental data mentioned previously. On the one hand, the hypothesis implies preservation of native-like core, which is consistent with the large-angle X-ray scattering and the NMR data. On the other hand, it implies that a large part of the molecule is unfolded and therefore that water penetrates into a thick shell of the molecule, "washing" its nonpolar core, which may be consistent with the data on the partial specific volume and the heat capacity.

The unfolded part of a molecule must have the partial specific volume of an unfolded protein, which is very near to that of a native one (e.g., Bendzko et al., 1988). Therefore, the difference between the partial specific volumes of the molten globule and the native states is reduced to the small expansion of a small part of a molecule that remains impenetrable to the solvent (i.e., to a few percent). In a similar way, the penetration of water inside the molecular shell can lead to a large increase in the heat capacity, even in the case when the small part of a molecule remains screened from water.

Of course, this is no more than a crude draft of a possible explanation for this contradiction. Precise experiments are needed to check the existence of expected differences in partial specific volumes and heat capacities between the molten globule and unfolded proteins. We need quantitative measurements of the partial specific volume, the heat capacity, and the intramolecular contacts in the molten globule state to understand why it preserves many intramolecular contacts, while having a lot of water inside.

Unfolding of the loops of a molecule make them susceptible to rapid hydrogen exchange. This is consistent with the evidence that those NH-protons that are involved in native hydrogen bonds of reverse turns or tertiary structure are not protected from exchange in the molten globule state. (See Section 3.5.) On the other hand, parts of a molecule that are involved in regular secondary structure become much more accessible to water, which facilitates their "breathing" and thus dramatically decreases the protection from exchange of even these fixed parts.

One of the consequences of this "double nature" of the nonuniform expanded molten globule is its high affinity to hydrophobic probes like 8-anilinonapthalene-1-sulfonate (ANS), which is bound by the molten globule state much stronger than by unfolded chains and by the majority of native proteins (Semisotnov et al., 1987,1991a; Rodionova et al., 1989; Ptitsyn et al., 1990). This probe binds to solvent-accessible clusters of nonpolar atoms (Stryer, 1965), which are absent in unfolded chains and relatively rare in native proteins. A molten globule in which the nonpolar core is preserved, but becomes less tightly packed and accessible to the solvent, is the ideal case for this binding.

4.3 Terminology

It is worthwhile now to discuss some terminology questions. As was mentioned in the beginning of this section, the molten globule state was predicted as a compact state with native-like secondary structure and native-like tertiary fold, but without rigid tertiary structure. When this state was detected experimentally (Dolgikh et al., 1981) it was called a "compact globule with native-like secondary structure and with slowly fluctuating tertiary structure." The term *molten globule* was introduced later by Ohgushi and Wada (1983).

When designating the molten globule as a state with a "fluctuating tertiary structure," the term *tertiary structure* was used in its modern sense, i.e., for the *detailed* protein structure at atomic resolution, to distinguish it from the *crude* protein structure (or tertiary fold). It led later, however, to some misunderstandings. For example, Kim and Baldwin (1990) have described "Ptitsyn's model" as "the rapidly fluctuating liquid-like structure" which "is incompatible with a structure determined by fixed tertiary interactions" (p. 642). In fact, my model implied just the opposite picture—a fixed native-like tertiary fold with greatly increased fluctuations of side chains and irregular regions.

Kim and Baldwin (1990) proposed distinguishing between "the term 'molten globule' for Ptitsyn's model and the term 'collapsed form' for the experimentally observed intermediate." It is a good idea to distance the model (which implies the native-like tertiary fold) from the existing

experimental data. The term *molten globule* has been widely used, however, as an operational term for a compact state with pronounced secondary structure but with unknown location of secondary structure elements in the sequence and in space, and it is probably too late to change its meaning.

Therefore, I propose to retain the general term *molten globule* for this state of our knowledge and to introduce the terms *native-like molten globule* and *disordered molten globule* for compact states with and without a native-like tertiary fold. Thus, we shall designate the "molten globule" state of a given protein as either the "native-like" or "disordered" one when we learn whether it does or does not have a native-like tertiary fold. If it happens that the globularization of a protein molecule during folding occurs before formation of the native-like tertiary fold, we can speak of fast formation of a "disordered molten globule" and of its slow transition to a "native-like molten globule." The advantage of this terminology is that it avoids any misunderstanding when reading papers published since the early 1980's on this subject. It is of course too early now to speculate on the possible properties of "disordered" molten globules and particularly on their relation to squeezed coils.

5 TRANSITIONS BETWEEN CONFORMATIONAL STATES

5.1 Denaturation as an All-or-None Transition

Many years ago it was suggested that denaturation of globular proteins is an all-or-none process, i.e., that in the transition region there are only native and fully denatured molecules, without any measurable amount of intermediate states. The first evidence supporting this point of view was the coincidence of denaturation curves measured by different methods (Lumry et al., 1966; Tanford, 1968). The definitive evidence, however, was obtained by Privalov and his collaborators (Privalov and Khechinashvili, 1974; Privalov, 1979; Chapter 3) after they had developed the technique of precise microcalorimetric measurements (Privalov, 1974; Privalov et al., 1975). These measurements have shown that for small, single-domain proteins, the calorimetric enthalpy of denaturation per protein molecule is practically equal to the enthalpy per cooperative unit calculated from the temperature interval of melting. Therefore, the cooperative unit coincides with a protein molecule, i.e., a single-domain protein denatures as a whole without any visible intermediate states. For larger proteins consisting of two or more domains, a cooperative unit usually coincides with a domain (Privalov, 1982). The all-or-none mechanism of protein denaturation has been strictly established only for temperature denaturation. A number of less definitive data, however, strongly suggest that it is true also for other types of denaturation.

The all-or-none mechanism of protein denaturation is very important for the stability of protein function. If proteins could denature non-cooperatively (i.e., in small pieces), thermal motions would destroy a protein structure at all temperatures. As a result, proteins would not have rigid active centers, and this would inhibit their biological activities. Only the high cooperativity of denaturation makes a protein structure resistant to thermal motions unless and until the temperature becomes large enough to destroy the structure as a whole.

Thus, proteins are unique examples of molecules that can possess an *intramolecular all-or-none transition*—a fact that requires a full and convincing explanation.

The first attempt to explain the all-or-none character of protein denaturation (Ptitsyn and Eizner, 1965; Ptitsyn et al., 1968) was based on the "liquid drop" model of a protein molecule and describes protein denaturation as a globule-coil transition. Two approximate versions of the theory have been proposed; the first one used the virial expansion of the free energy of long-range intramolecular interactions (limited by the interactions of pairs and triplets of residues), while the second one used the approximate expression of this free energy similar to the van der Waals expression for real gases. Both of these lines have been followed afterwards—the first one by De Gennes (1975) and others and the second one by Sanchez (1979); for further references, see the review by Chan and Dill (1991). Birshtein and Pryamitsyn (1987,1991) have improved this treatment by replacing distribution functions for the end-to-end distance by that of the radius of gyration, which better represents the linear dimensions of squeezed coils. A more strict and sophisticated, but physically equivalent treatment has been elaborated by Lifshitz et al. (1978).

All of these versions of the theory give the same result—they predict a cooperative but not an all-or-none globule-coil transition for flexible homopolymers. Formally, an all-or-none transition is predicted for rigid chains, but even in this case the barrier between native and denatured states is approximately of magnitude k_BT (A. V. Finkelstein, private communication) and therefore intermediate states can also be populated. The drastic difference between a phase liquid-gas transition and a smooth globule-coil transition is due to the fact that monomers in polymer molecules are connected in a chain; therefore, each monomer has a limited freedom that does not depend upon the dimensions of the entire molecule.

The existence of cooperative coil-globule transitions in a homopolymer has been confirmed experimentally (Anufrieva et al., 1972) by polarized fluorescence, which permits the use of small concentrations of polymer and avoids it aggregation. Neither this nor the subsequent studies (see, e.g., Sun et al., 1980), however, give a clear answer to the character of this transition. See reviews by Chan and Dill (1991), Anufrieva

(1982), and Anufrieva and Krakovyak (1987) for other references. Although there is some evidence against the all-or-none character of such transitions (Anufrieva et al., 1972), the problem still needs careful investigation. Monte Carlo computer experiments on the collapse of a nonintersecting chain on a cubic lattice (Kron et al., 1967) also did not show an all-or-none transition.

The theory of collapse of heteropolymers has also been developed (Dill, 1985; Grosberg and Shakhnovich, 1986a,b; Dill et al., 1989). Grosberg and Shakhnovich (1986a,b) extended the approach of Lifshitz et al. (1978) to heteropolymers and have shown that the coil-globule transition in a very long heteropolymer must be similar to that of a homopolymer.

A liquid drop is, of course, not the best model for a protein molecule, which has rigid α- or β-regions. There are no detailed theories that take this important fact into account. Grosberg (1984) has considered a simple model of a polymer chain in which collapse (coil-globule transition) is coupled with the helix-coil transition. He has shown that this coupling may dramatically increase the cooperativity of the transition, transforming it into an all-or-none transition.

Another approach to this problem (Ptitsyn, 1975) is based on the fact that proteins (or protein domains) consist of only a few "structural blocks" (α-helices and β-strands). If these blocks are unstable in the unfolded state, there are only few states that can be intermediate between the completely unfolded states. In this system, even relatively small cooperativity (e.g., the unfavorable initiation and/or favorable termination of a folded structure) will lead to a small total statistical weight of all the intermediates. This will lead to a transition that practically is all-or-none not because intermediates are forbidden but because there are only a few slightly unstable intermediates. The difference between this case and the real all-or-none transition is the height of the free energy maximum between the folded and unfolded states—for the "quasi all-or-none" transition this maximum must be relatively low.

5.2 Native–Molten Globule Transition

The problem of protein denaturation dramatically changed after it was shown that all-or-none temperature transitions occur in bovine (Dolgikh et al., 1981,1985) and human (Dolgikh et al., 1981; Pfeil et al., 1986) α-lactalbumins. The temperature-denatured states of these proteins are compact (Dolgikh et al., 1981,1985; Gast et al., 1986) and have pronounced secondary structure (Dolgikh et al., 1981,1985), i.e., they are in the molten globule state. Therefore, at least for these proteins, the all-or-none transition is *not* a globule-coil transition, i.e., is not connected with the unfolding of a protein chain or with the destruction of its secondary structure. The single common feature of all denatured proteins—from molten globules to unfolded

chains—is the absence of the tight packing of side chains. This suggests that the common reason for the all-or-none character of protein denaturation is the destruction of the tight packing.

This idea was the basis of the novel theory of protein denaturation proposed by Shakhnovich and Finkelstein (1989). Their theory describes protein denaturation as melting of a crystal rather than as the evaporation of a liquid drop. It explains the all-or-none transition by using the model of tightly packed side chains attached to rigid regions (α-helices and β-strands). Due to this attachment, the movement of side chains can be released only cooperatively, because it is impossible to increase the distance between a few of them, while leaving the other distances unaltered. In the native state, side chain movements typically are limited to their rotational vibrations with relatively small amplitudes. Therefore, the entropy of this state is low. However, its energy also is low (due to large van der Waals attractions, which stabilize this state). If the distances between side chains increase, amplitudes of rotational vibrations increase, but the corresponding gain in entropy is too small to compensate for the increase in energy. At some threshold distance, however, a new type of aliphatic side chain movement is released—these groups begin to jump from one rotational isomer to another. The additional increase in entropy compensates for the increase in energy, and the denatured state becomes stable. As side chain–side chain distances can change only simultaneously, they can be described by a single parameter—the molecular volume. This means that a protein molecule as a whole can have two stable states: an energetically stable native state with a small volume and an entropically stable denatured state with an increased volume. The states with intermediate volumes have high energy and low entropy and therefore are unstable.

According to this theory the "fate" of the denatured protein molecule depends upon the solvent and upon the temperature. Sometimes it can be the molten globule state with a moderately increased volume and release of some new movements. In other conditions (e.g., at high concentrations of strong denaturants), however, it can be the completely unfolded state, with a large increase of the volume and the release of all movements.

Shakhnovich-Finkelstein's theory is a very important step in our understanding of the all-or-none character of protein denaturation. It emphasizes two essential points: the role of the destruction of tight packing of the side chains and the release of new types of movements at some threshold value of the volume. This theory, however, refers to the "uniform" model of the molten globule, i.e., it does not distinguish between the protein core and the protein shell.

According to the nonuniform model of the molten globule state (see Section 4), the main difference between the molten globule and the native protein is the state of the protein shell rather than that of the protein

core; loops forming the protein shell are more or less unfolded, while the protein core still exists, becoming only a little less compact. In native proteins, loops are attached to the hydrophobic core by their nonpolar side chains. The small increase in volume of this core will release these side chains and this process can be cooperative if all side chains need the same threshold value of core volume to be released. This threshold value can be smaller than the value that permits side chains to rotate inside a core, as it is easier to remove a side chain than to rotate it in a dense environment.

The cooperative release of nonpolar side chains belonging to loops may lead to the cooperative unfolding of these loops, because their native structure may become unstable without this support. This unfolding will lead to a large increase in entropy, which can compensate for the increase in the core volume.

The release of a number of nonpolar side chains from the core will permit other side chains to increase the amplitudes of their rotational vibrations and probably even to jump from one rotational isomer to another. Therefore, this interpretation probably supplements the Shakhnovich-Finkelstein theory rather than contradicts it. The main point is that "new" movements that must be switched on at the threshold volume can be not only the rotational isomerization of core side chains, but also, or even predominantly, the unfolding of loops (of course, together with rotational isomerization of their side chains). According to this point of view, the cooperative increase of a core volume may be more the trigger than the reason for the all-or-none transition.

5.3 Globule–Coil Transition

Does the existence of an all-or-none transition between the native and the molten globule states mean that the molten globule state is thermodynamically equivalent to the unfolded state? In other words, is the molten globule just a limiting case of the unfolded state, with maximum compactness and secondary structure, or is it divided from the unfolded state by another all-or-none transition?

GdmCl- and urea-induced molten globule–coil transitions appear to be cooperative, in that they occur within a relatively narrow interval of concentrations of strong denaturants. These intervals are determined by the differences of the number of denaturant molecules "bound" to a cooperative unit in the unfolded and in the molten globule state. We cannot measure the numbers of denaturant molecules bound to one residue in the unfolded and in the molten globule states, so it is practically impossible to evaluate the number of residues in a cooperative unit and to compare it with the number of residues in a protein or in a protein domain. Thus, the cooperativity of the molten globule–coil transition does not give us any information as to whether this transition is all-or-none.

On the other hand, the small enthalpy of the molten globule–coil transition in human α-lactalbumin (Pfeil et al., 1986) does not exclude the existence of an all-or-none transition, as the solvent-induced transition may be connected with a jump in the number of bound denaturant molecules rather than with a jump of enthalpy.

A common, but not definitive, criterion for all-or-none transitions is to compare the transition curves obtained by different methods. If all these curves coincide, it is likely to be an all-or-none transition; if not, this type of transition is excluded. It was this method that had been used in the "precalorimetric" era to suggest that protein denaturation is an all-or-none transition. This method has been used to study the urea-induced molten globule–coil transition in bovine carbonic anhydrase B (Rodionova et al., 1989). The result was that the unfolding of the molten globule state of this protein monitored by the change of its compactness (intrinsic viscosity, polarization, and spectrum of fluorescence) coincided with the unfolding monitored by the change of its secondary structure, i.e., by the ellipticity at 220 nm (Fig. 6–7A). This suggests that in the transition region there are no stable compact intermediates without secondary structure, as well as no noncompact intermediates with secondary structure.

This result is confirmed in Figure 6–7B, which presents the urea-induced unfolding of native carbonic anhydrase. This unfolding process is biphasic. The first phase involves the destruction of the rigid tertiary structure, while the second is the unfolding of the molten globule state. Again the curves monitored by compactness (polarization and spectrum of fluorescence) and by secondary structure coincide, which shows the absence of intermediates between the molten globule and the unfolded states. Similar results have been obtained for GdmCl-induced unfolding of native human carbonic anhydrase B. In addition, the changes of the NMR signals shown in Figure 6–7B demonstrate that the mobility of the aliphatic side chains increases at the first stage of denaturation (at the native–molten globule transition), while the mobility of the aromatic side chains changes only at the second stage (at the molten globule–coil transition).

The results of this study suggest that the transition from the molten globule to the unfolded chain may also be an all-or-none transition. The most direct method to prove this would be to demonstrate a bimodal distribution of protein molecules between the two states in the transition region. Recently, we have used this approach (Uversky et al., 1992) to study the unfolding of the molten globule states of two proteins by FPLC size-exclusion chromatography. The main difficulty with these experiments is that it is possible to separate only slow-exchanging conformational states of protein molecules. We have found that this exchange is slow enough only at a low temperature (4°C) and for GdmCl-induced (not urea-induced!) unfolding. Under these conditions, two peaks in the

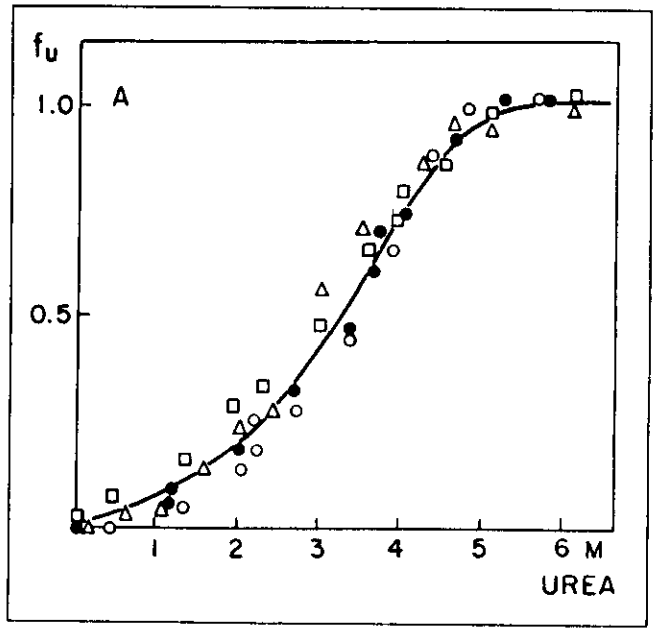


FIGURE 6–7. Urea-induced unfolding of bovine carbonic anhydrase B at pH 3.6 (A) and at pH 7.5 (B). At pH 3.6 in the absence of urea, the protein is in the molten globule state and unfolds by a one-step process. At pH 7.5, the protein is in the native state and unfolds in two steps (native–molten globule and molten globule–unfolded transitions) without additional intermediates. The fraction of unfolding is given by $f_u = (x - x_o)/(x_u - x_o)$, *where* x *is the value of a measured parameter,* x_o *is its value in the absence of urea, and* x_u *is its value at high concentration of urea. (A)* ●, *the increase of intrinsic viscosity* [η], □, *the decrease of* I_{320}/I_{360} (I *is the intensity of tryptophan fluorescence at the given wavelength),* ○, *the increase of* 1/P (P *is the polarization of tryptophan fluorescence),* △, *the decrease of the negative ellipticity at 220 nm. Intrinsic viscosity and the spectrum and polarization of tryptophan fluorescence both reflect the compactness of the molecule, while the ellipticity at 220 nm reflects its secondary structure. (Adapted from Rodionova et al., 1989)*

elution curves have been observed within the equilibrium molten globule–coil transition for bovine carbonic anhydrase B and for β-lactamase from *Staphylococcus aureus*. The first peak is close to that of the native protein (reflecting the almost native compactness of the molten globule state) while the second has a smaller elution volume, demonstrating that this state is more expanded than the first. When the GdmCl concentration increases in the transition zone, the more compact molten globule peak decreases in magnitude, while the unfolded peak increases; this

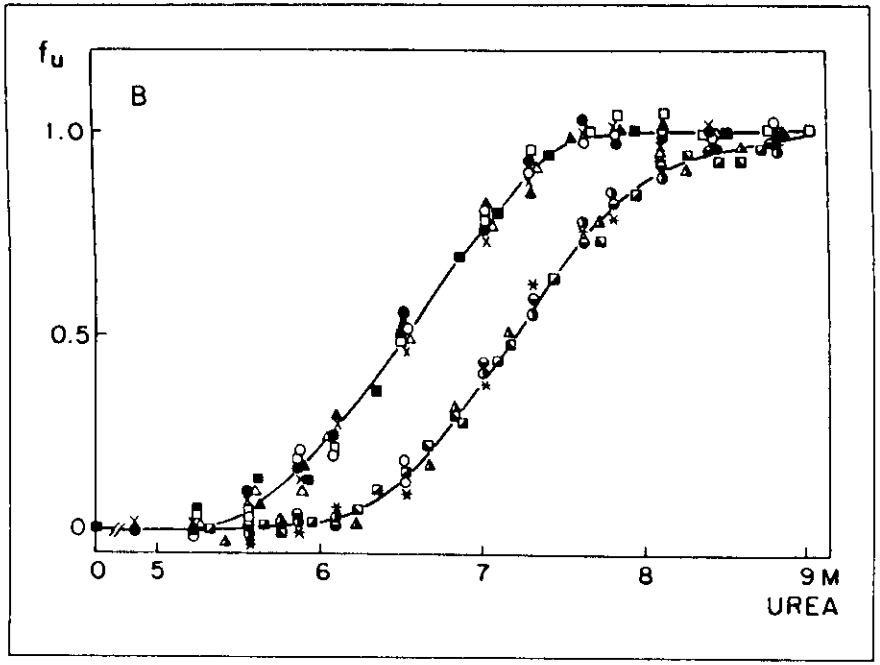


FIGURE 6.7. (B) △, the increase of I_{320}/I_{360}; ◣, the increase of 1/P; ◢ and ■, the decrease of negative ellipticity at 220 nm (◢) and 270 nm (■); ▲, the decrease of the area under high-field NMR resonances; other signs mean the increase of signal intensities of protons of aliphatic side chains at 3.17 (×), 2.97 (○), 2.00 (●), 1.38 (◡), and 0.86 (△) ppm as well as of aromatic side chains at 7.5 (◑), 7.10 (◒), and 6.79 (*) ppm. I_{320}/I_{360} and P reflect the compactness of the molecule, ellipticity at 220 nm reflects its secondary structure, while the ellipticity at 270 mm and the area under the high-field NMR resonances reflect its rigid tertiary structure. (Adapted from Rodionova et al., 1989)

demonstrates the all-or-none transition of proteins from the molten globule to the more expanded state.

This analysis of the molten globule–coil transition leads us back to the general problem of the cooperativity of globule-coil transitions. (See Section 5.1.) As has been already mentioned, the globule-coil transition may be an all-or-none transition of the first order if it is coupled with a change of linear order, e.g., with the helix-coil transition (Grosberg, 1984). Therefore, unfolding of the molten globule, with the simultaneous destruction of its secondary structure, can be similar to the melting of a liquid crystal and may be an all-or-none transition.

It is pertinent to mention that we also know an example of a smooth (not all-or-none) molten globule–coil transition. The first experimental

evidence for the molten globule state was actually obtained for random copolymers of glutamic acid with leucine (Anufrieva et al., 1975; Bychkova et al., 1980; Semisotnov et al., 1981). Upon deionization of the glutamic acid residues, these copolymers first became helical and then collapsed to a compact helical state. The transition of these copolymers from the noncompact helical state to the collapsed helical state clearly is not an all-or-none transition, because the relaxation time of depolarization of fluorescence goes through a maximum. The difference between the pH-induced globule-coil transition in these copolymers and the urea- or GdmCl-induced transitions in proteins is that the pH-induced transition occurs between collapsed helical and noncompact helical polymer, while the influence of urea and GdmCl may transform a protein from a condensed state with secondary structure to an unfolded state without secondary structure. Therefore, this difference suggests that the coupling between condensation and linear structure of a chain may be very important for an all-or-none transition.

On the other hand, the molten globule, unlike globules from random copolymers, may have a unique tertiary fold. (See Section 3.5.) Although a globule-coil transition for a globule with a unique tertiary fold has not been extensively studied theoretically, the possibility cannot be excluded that this transition may be an all-or-none transition (Shakhnovich and Gutin, 1990). Much more experimental and theoretical work is needed to clarify the physical reasons for this unusual phenomenon, the all-or-none globule-coil transitions of denatured proteins.

Depending upon the stability of the molten globule state, it can unfold at different concentrations of strong denaturants (as well as at different pH values and temperatures). Therefore, the unfolded protein at the end of this transition can be expanded to different extents, depending upon the properties of the solvent (Damaschun et al., 1991). One cannot exclude the possibility that very unstable native or molten globule proteins can unfold into "squeezed coils," i.e., coils below the Θ-temperature. (See footnote on p. 261.) These coils can be transformed into Gaussian ones and then into expanded coils by a further increase in the concentration of the strong denaturant. This transition is *not* an all-or-none transition, rather it may be a phase transition of the second order (Lifshitz et al., 1978).

6 OCCURRENCE OF THE MOLTEN GLOBULE

After the molten globule state was first described for bovine and human α-lactalbumins, a number of other proteins have been reported to have similar properties. These are collected and classified in Table 6–2. (See also Bychkova and Ptitsyn, 1992.) Table 6–2A collects proteins for which all three main features of the molten globule state (compactness, the presence of secondary structure, and the absence of rigid tertiary structure) have been

TABLE 6–2. Molten Globule States in Proteins

Protein	Conditions	Compactness[a]	Secondary Structure[b]	Tertiary Structure[c]	Cooperative Unfolding[d]	Binding to Hydrophobic Probe	References
			A. WELL-ESTABLISHED MOLTEN GLOBULE STATES				
Bovine α-lactalbumin	pH 2; 0.05 M KCl	R_D=21.1; N: 18.4; U: 32.3	10,000 N: 9,500	−20 (270 nm) N: -320	—	ANS[e], phospho-lipids	Robbins and Holms, 1970; Kuwajima et al., 1976; Dolgikh et al., 1981,1985; Gilmanshin et al., 1982; Izumi et al., 1983; Kim and Kim, 1986; Bychkova et al., 1990; Semisotnov et al., 1991a
		[η]=3.1; N:3.4; U:6.1	IR	NMR			
		R_g=16.8; 15.7; N: 15.6; 15.5					
		λ_{max}=338					
	pH 7	—	—	—	2.0 and 2.6[f] GdmCl	—	
	pH 7; 50°C; 0.05 M KCl Ca^{2+} -free	[η]=3.0	7,000	−50; NMR	—	—	
	pH 11.3; 0.15 M KCl	R_g=16.8; U: 24.4	$[\Theta]_{208}$=14,500 N: 10,500	−100; N: -300	—	—	
Human α-lactalbumin	pH 2; 0.05 M KCl	[η]=4.2;N:3.1; U:6.6	14,000; N: 9,000	−60 (270 nm) N: −240; U: −50	—	ANS, melittin	Nozaka et al., 1978; Dolgikh et al., 1981; Gilmanshin et al., 1982; Gast et al., 1986; Bychkova et al., 1990; Permyakov et al., 1991
		R_D=19.9; N:17.7; U: 24.7					
		λ_{max}=331					
	pH 7.4; 1.85 M GdmCl	λ_{max}=336	9,000	−60	3.5 GdmCl	—	
	pH 7; 50°C; Ca^{2+} -free	R_D=19.6	Similar to pH 2	Similar to pH 2		—	
Bovine carbonic anhydrase B	pH 3.6	[η]=4.1; N:3.7; U:29	4,600 N: 3,500	+5 (270 nm) N: −100; U: 0	3.5 urea	ANS	Wong and Tanford, 1973; Wong and Hamlin, 1974; Dolgikh et al., 1983; Rodionova et al., 1989; Semisotnov et al., 1991a; Uversky et al., 1992
		P=0.09; N:0.08$_5$; U:0.05	IR				
		λ_{max}=340					
	pH 7	—	—	—	6.5 and 7.2 urea[f]	—	
	pH 7; 4°C; 1.3 M GdmCl	$R_S \equiv (R_S)_N$	—	—	1.5 GdmCl	ANS	
	pH 7; 1.9 M GdmCl	$R_S \equiv (R_S)_N$	Similar to N	Similar to U	2.3 GdmCl	ANS	
Human carbonic anhydrase B	pH 3	—	5,800; N: 600; U: 1,800	−40; N: −180	—	—	Jagannadham and Balasubramanian, 1985; Rodionova et al., 1989
	pH 8; 1.7 M GdmCl	η_{sp}/c similar to N	5,000; N: 970; U: 1,500	10% from N	2.5 GdmCl	ANS	
β-Lactamase *Staphylococcus aureus*	pH 7; 4°C 0.5 M GdmCl	$R_S \equiv (R_S)_N$	$-[m']_{234}$=3,000 N: 3,600 (1 M GdmCl at 20°C)	Nonactive	0.75 GdmCl	—	Robson and Pain, 1976a,b; Uversky et al., 1992
Ribonuclease A	pH 3; 3 M $LiClO_4$	[η] = 5.5 N: 3.4; U: 7.5	8,000 N: 9,000	−20 N: −250	2.5 urea	—	Ahmad and Bigelow, 1979; Denton et al., 1982
	pH <2 (or pH 2 at high KCl)	$R_S \equiv (R_S)_N$	Similar to N	Similar to U	—	—	Fink et al., 1990
T4 Lysozyme	pH 7.4; 2.0 M GdmCl	$\lambda_{max}=(\lambda_{max})_N$	60% from N	—	2.2 GdmCl	—	Desmadril and Yon, 1981
Bovine pancreatic trypsin inhibitor, reduced	0.5 M GdmCl	Specific energy transfer	13,250 (200 nm)[g] N: 19,800 (202.5 nm)	0 (275 nm) N: −520	—	—	Kosen et al., 1981, 1983; Amir and Haas, 1988
		P=0.18; N: 0.18[h]					

continued

Protein	*Conditions*	*Compactness*[a]	*Secondary Structure*[b]	*Tertiary Structure*[c]	*Cooperative Unfolding*[d]	*Binding to Hydrophobic Probe*	*References*
Horse cytochrome *c*	pH 2.2; 0.5 M NaCl	[η]=3.1; N: 2.8; U: 12.7; R_g=20.1 N: 19.8	11,300; N: 10,600	Weak near UV circular dichroism Small protection against H ⇄ D exchange (at 1.5 M NaCl)	—	—	Ohgushi and Wada, 1983; Jeng et al., 1990
Bovine growth hormone, reduced	pH 9.1; 4.5 M urea	$R_S \equiv (R_S)_N$[h]	8,500 N: 15,000	ε_{290}=7,300 N: 9,600	7.8 urea	—	Holzman et al., 1986
Fragment F_c of IgG	pH 4.5; 64°C	s=4.1; N: 4.1[i]	3,500 N: 3,300	-80 (270 nm) N: -120 Fluorescence	—	—	Vonderviszt et al., 1987
Apomyoglobin	pH 4.3; 27°C	[η]=5.0; N: 3.5; U: 20.7	11,500 N[i]: 18,300	0; N: +20	—	ANS, phospho-lipids	Griko et al., 1988; Lee and Kim, 1988
Diphtheria toxin, fragment A	pH 3–5; 0.15 M NaCl	λ_{max}=336 N: 328; U: 350	3,000 N[i]: 3,000 IR	Proteolysis	0.5 urea	Liposomes, TID[j], membranes	Zhao and London, 1988; Dumont and Richards, 1988; Cabiaux et al., 1989
Colicin A (pore-forming domain)	pH 4; 0.1 M KCl	s=3.66; N[i]: 3.66	14,000 N: 14,500	Proteolysis	—	Membranes	Cavard et al., 1988
Colicin E1 (channel peptide)	pH 3.5; 0.1 M NaCl	R_S=20.7; N[i]: 22.7; U: 84.8	Similar to N[l]	Proteolysis	—	Membranes	Merrill et al., 1990
Staphylococcal nuclease, fragment 1-128	pH 7.0; 0.05 M NaCl	R_S=21; N[h]: 21 (0.2 M NaCl)	8,000 N: 13,000	Nonactive (though includes active center)	0.3 GdmCl	—	Shortle and Meeker, 1989
Rhodanese	pH 7.4; 2 M GdmCl	λ_{max}=348; N: 335; U: 356	90% from N	Much less than N; proteolysis	3.3 GuHCl	ANS	Tandon and Horowitz, 1989
	pH 7.4; 4 M urea	λ_{max}=345	80% from N	-SH groups are easily oxidized	5.0 urea	—	Horowitz and Criscimagna, 1990
β-Lactamase *Bacillus cereus*	pH 12; 1.5 M KCl	R_D=31; N: 23; U: 55; λ_{max}=332	10,000 N: 12,000	0; N: -100	—	ANS	Goto and Fink, 1989
	pH 2; 0.5 M KCl	R_S=26.5; N:24; U:51; λ_{max}=332	11,000	+20	—	ANS	Goto et al., 1990a
Aspartate amino transferase (monomer)	0.9–1.1 M GdmCl	s=3.1; N:3.4[k];U:2.2 (6°C) λ_{max}=339 N: 335; U: 355	50% from native dimer	Fluorescence	1.7 GuHCl	—	Herold and Kirschner, 1990
Retinol binding protein (RBP) apo-form	pH 2	η_{sp}/c=9.7[l]; N[i]: 4.4; U: 24.5 λ_{max}=336	7,500 (207 nm) N: 5,000	+5 (285 nm) N: -130; NMR	—	ANS	Bychkova et al., 1992
B. POSSIBLE MOLTEN GLOBULE STATES							
Bovine growth hormone	pH 2; 0.1 M NaCl	η_{sp}/c=10.5; N: 3.0	b_o= -225; N: -225	ε_{290} is changed	—	—	Burger et al., 1966
	pH 8.5; 3.3 M GdmCl	R_S=26; N: 18; U: 37; λ_{max}=342	11,000 N: 15,000	ε_{290}=7,400; N: 9,800; U: 7,400	3.7 GdmCl	—	Brems et al., 1985; Brems and Havel, 1989
	pH 4.0; 4.5 M urea	$\lambda_{max} = (\lambda_{max})_N$	8,000 (50% N)	ε_{290}=7,000 N: 9,000	5.3 urea	—	Holzman et al., 1990
β-lactoglobulin	pH 2; 2.7 M GdmCl	—	6,500 N: 6,000	-15 N: -90	3.0 GdmCl	—	Ananthanarayanan et al., 1977
	pH 2; 90°C 0.1 M KCl	η_{sp}/c=11.5 U: [η]=19.1	11,000 N: 5,500	-10	2 GdmCl	—	Ananthanarayanan and Ahmad, 1977

continued

Protein	Conditions	Compactness[a]	Secondary Structure[b]	Tertiary Structure[c]	Cooperative Unfolding[d]	Binding to Hydrophobic Probe	References
Parvalbumin, carp, apo-form	pH 7; 42°C	—	$-[m']_{233}$=6,500 N: 7,500	NMR	—	—	Cave et al., 1979
	pH 2.5–2.8; 0.1=M NaCl; 4.5°C	—	9,000 N: 14,000	-20 (267 nm) N:-40; U:-15	0.7 GdmCl	—	Kuwajima et al., 1988
Human α-interferon	pH 1.5	—	9,000 N: 15,000	0 (292 nm) N: -150	—	—	Bewley et al., 1982
Human γ-interferon	pH 3.5; 0.05 M NaCl	—	12,000 N: 15,000	0 N: +120	—	—	Arakawa et al., 1987
	pH 2.0; 0.05 M NaCl	—	10,000	+10	—	—	Arakawa et al., 1987

[a] R_s and R_D: Stokes radii (in Å) from FPLC and from diffusion coefficient, respectively; R_g, radius of gyration from diffuse X-ray and neutron scattering data; $[\eta]$ and η_{sp}/c, intrinsic and reduced viscosity (in $cm^3.g^{-1}$), respectively; P, polarization of fluorescence of tryptophan residues; s, sedimentation coefficient in Svedberg units (10^{-13} sec^{-1}); λ_{max}, wavelength of the maximum of tryptophan fluorescence spectrum (in nm). Data for native (N) and unfolded (U) states (if available) are shown for comparison. λ_{max} for N-state is usually equal to 325 to 330 nm and for U-state 350 to 355 nm.

[b] $-[\Theta]_{222}$ or $-[\Theta]_{220}$ (in deg.cm^2.$dmol^{-1}$) unless otherwise stated ($[m']$ in deg.cm^2.$dmol^{-1}$ or b_o in degrees). Data for the N-state are shown for comparison. "IR" means that infrared spectrum shows pronounced secondary structure.

[c] $[\Theta]_{280}$ (in deg.cm^2.$dmol^{-1}$) if other (molar absorption ε_{290} in $1.mol^{-1}.cm^{-1}$) is not stated. Data for N-state are shown for comparison. "NMR" means poor (U-like) NMR spectrum. "Proteolysis" means an increase in susceptibility to proteases as compared to the native state. "Fluorescence" means an increase of the intensity of tryptophan fluorescence as compared to the native state.

[d] Molar concentrations of urea or GdmCl corresponding to the midpoint of the transition to the unfolded state.

[e] ANS: 8-anilino-1-naphthalene sulfonate.

[f] In these cases two overlapping urea- or GdmCl-induced transitions have been observed by $[\Theta]_{222}$ and $[\Theta]_{280}$, which indicates the existence of an intermediate state.

[g] For BPTI, the contribution of aromatic side chains in far- UV CD can be very large (Manning and Woody, 1989).

[h] "Native" values are the values for an intact protein (i.e., for a protein with intact disulfide bridges or to a protein of wild-type) at native conditions.

(continued)

established. Therefore, it is fairly certain that these proteins adopt the molten globule state. Table 6–2B includes proteins for which only the presence of secondary structure and the absence of tertiary structure have been established; the data on compactness are either absent or doubtful. These proteins may be assumed to be in the molten globule state, although additional information is necessary to be certain.

Four conclusions can be drawn from this table:

1. About 20 to 25 different proteins have properties similar to that of the molten globule state. This means that the molten globule state is not a rare exception. Rather, it is probably typical for many proteins under mild denaturation conditions.
2. Proteins can be transformed into the molten globule state by low or high pH, by high temperature, by moderate concentrations of GdmCl or urea, and by the influence of $LiClO_4$ and other salts, i.e., by quite different mild denaturation conditions.
3. Proteins can be transformed into the molten globule state without a change of the environment (i.e., almost under physiological conditions) by small alterations of their chemical structure. Some examples include:
   a. Staphylococcal nuclease with 21 C-terminal residues removed (Shortle and Meeker, 1989);
   b. Point mutants of λ repressor (W. A. Lim and R. T. Sauer, personal communication);
   c. Bovine pancreatic trypsin inhibitor with reduced disulfide bridges (Amir and Haas, 1988);
   d. Bovine and human α-lactalbumins after removal of the Ca^{2+} ion (Kuwajima, 1989).
4. Among proteins that possess a pH-induced molten globule state are a number of membrane-binding proteins (colicins, retinol-binding protein, growth hormone, α- and γ-interferons, diptheria toxin). The question arises as to whether the electrostatic field of the membrane (McLaughlin, 1989; Stegmann et al., 1989) can produce similar transitions in these proteins. This is part of the very important question of the possible role of molten globules in a living cell. (See Section 8.)

Table 6–2 footnotes, continued

[i] "Native" values are the values at pH 7.

[j] TID: 3-(trifluoromethyl)-3-(m-[^{125}I] iodophenyl)-diazirine.

[k] Calculated from data for dimeric protein.

[l] η_{sp}/c at acid pH and low ionic strength may be increased relative to the native state due to electrostatic repulsion of ionized molecules.

All data refer to room temperature, unless otherwise stated.

7 PROTEIN FOLDING

7.1 The Molten Globule and Protein Folding

The problem of the kinetics of protein folding, especially of kinetic intermediates in folding pathways, has been extensively reviewed (e.g., Kim and Baldwin, 1982,1990); Creighton, 1978,1990; Fischer and Schmid, 1990) and is the subject of Chapters 5 and 7 of this book. Therefore, this chapter does not give a full description of the problem, but it concentrates on the role of the molten globule state in protein folding and on some related questions.

The molten globule state was first predicted as a *kinetic* intermediate in protein folding pathway. (See Section 4.) That hypothesis of protein folding (Ptitsyn, 1973) suggested that folding starts with the formation of "fluctuating embryos" of regions with secondary structure (stabilized mainly by hydrogen bonds), followed by the collapse of these regions into an "intermediate compact structure" (stabilized mainly by hydrophobic interactions) and completed by the adjustment of this intermediate structure to the unique native structure (driven by van der Waals and other specific interactions). This hypothesis implies that the positions of the embryos of secondary structure in an unfolded chain must be close to their positions in the folded globular state and that the tertiary fold of the polypeptide chain in a compact intermediate must be similar to that in the final native structure. This scheme has subsequently been designated the "framework model" (Kim and Baldwin, 1982).

The first experimental evidence for the existence of a compact kinetic intermediate during protein folding was obtained using urea-gradient gel electrophoresis near 0°C (Creighton, 1980; Creighton and Pain, 1980). It has also been shown that the folding of β-lactamase (Robson and Pain, 1976b) and carbonic anhydrase (McCoy et al., 1980) is accompanied by the accumulation of kinetic intermediates that have pronounced secondary structure (monitored by far-UV CD) but no rigid tertiary structure (monitored by near-UV CD). Similar evidence has been obtained using amide hydrogen exchange and ^3H labeling for RNase A (Schmid and Baldwin, 1979; Kim and Baldwin, 1980). Finally, Dolgikh et al. (1984) have shown that the early kinetic intermediate of bovine carbonic anhydrase B has viscosity and far-UV ellipticity similar to those of the native state, but has no rigid tertiary structure and no enzyme activity. Therefore, it was concluded that the molten globule state accumulates during protein folding.

All of these studies have not answered, however, the question of the time scale for the formation of the molten globule state and particularly for the time scales of the formation of secondary structure and of the molecular globularization. These questions have been answered only after the development of methods for the rapid monitoring of these processes (Gilmanshin, 1985,1988; Kuwajima et al., 1987; Gilmanshin and Ptitsyn, 1987; Semisotnov et al., 1987,1991a; Ebert et al., 1990).

7.2 Secondary Structure

By combining far-UV CD and the stopped-flow technique, Kuwajima et al. (1987), Gilmanshin and Ptitsyn (1987), Goldberg et al. (1990), and Semisotnov and Kuwajima (unpublished data reviewed by Ptitsyn and Semisotnov, 1991) have shown that the far-UV ellipticity of proteins changes drastically at a *very early stage* of their folding (within the dead time of these experiments, which is approximately 0.01 sec). The negative far-UV ellipticity that is reached after 0.01 sec can be smaller, equal to, or even greater than the native value, but in all cases it is substantially different from the value for the unfolded protein. This suggests the very fast formation of pronounced secondary structure at an early stage of protein folding.

If the far-UV CD ellipticity value reached after 0.01 sec is not equal to the native value, is usually approaches this value relatively slowly, with a rate comparable to the rate of restoration of the rigid tertiary structure monitored by near-UV ellipticity (Kuwajima et al., 1987,1988; Gilmanshin and Ptitsyn, 1987; Goldberg et al., 1990); this is illustrated in Figure 6-8.

The interpretation of these data must take into account that the far-UV CD is very likely influenced by the contribution of aromatic side chains, which can be much greater in the native, rigid state than in nonrigid intermediates. (See Section 3.4.) Therefore, only the change of the far-UV ellipticity that occurs *before* the change in the near-UV can be unambiguously interpreted, while the slow changes can also reflect the increase in the aromatic side chain contribution. Even an absence of slow changes may not indicate that the secondary structure content is constant, as it could be due to changes of secondary structure and the aromatic contribution compensating each other.

7.3 Globularization

To study the next step of protein folding—the collapse of the protein molecule—energy transfer from Trp residues to fluorescent dansyl labels has been used (Semisotnov et al., 1987; Ptitsyn and Semisotnov, 1991). If the Trp residues are close in space to dansyl labels, part of their excitation energy can migrate to the dansyl groups and be emitted as dansyl fluorescence, even if it is the Trp residues that are excited. In proteins with several Trp residues and/or several dansyl labels, energy transfer is averaged over all the Trp-dansyl pairs; it therefore must reflect the overall dimensions of the

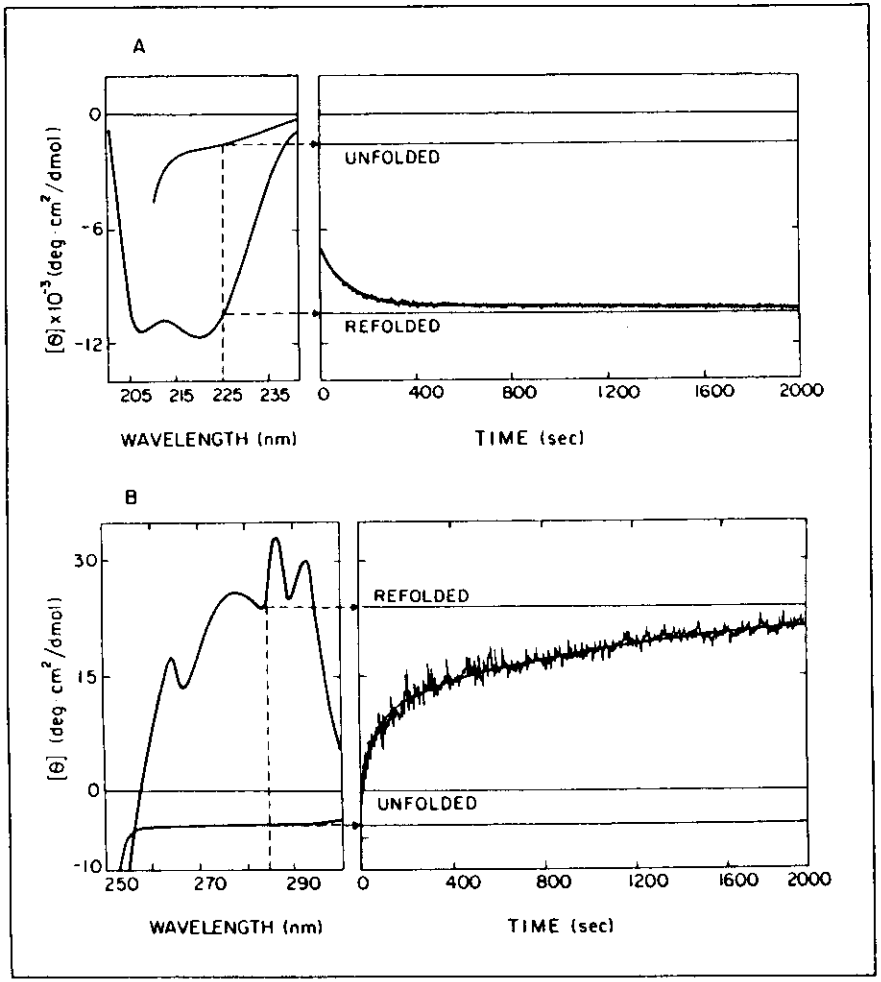


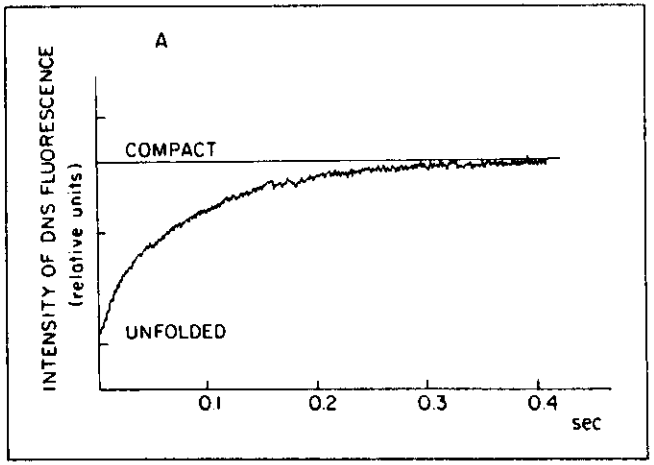
FIGURE 6–8. Far and near UV circular dichroism spectra of β_2 subunit of tryptophan synthase in the unfolded and folded states and time-dependence of molar ellipticities at 225 nm (A) and at 285 nm (B) upon of its refolding from 5.5 to 0.5 M urea at 12°C. (Adapted from Goldberg et al., 1990)

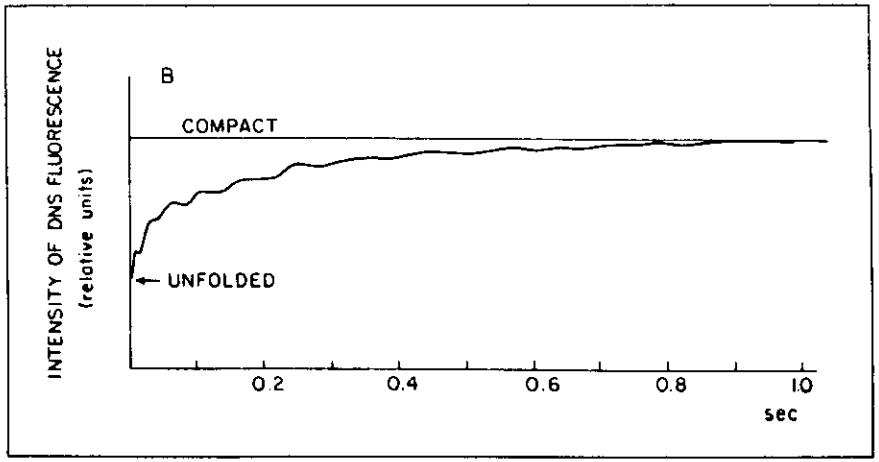
protein molecule. In addition, the intensities of electron spin resonance signals of labels covalently attached to a protein also can be used to monitor the compactness of a chain (Semisotnov et al., 1987; Ebert et al., 1990), because these intensities depend upon the environment of labels. Special experiments on protein equilibrium unfolding (Rodionova, 1990; see also Ptitsyn and Semisotnov, 1991) have shown that the energy transfer and ESR

signals can be used to monitor the kinetics of the molecular collapse of protein folding.

Figure 6–9 illustrates the kinetics of the restoration of the native level of energy transfer, measured by the intensity of dansyl fluorescence at Trp excitation, for bovine carbonic anhydrase B, and for β-lactoglobulin,

FIGURE 6–9. Kinetics of energy transfer increase, monitored by the increase of dansyl fluorescence, during (A) carbonic anhydrase refolding from 8.5 to 4.2 M urea at 23°C and (B) β-lactoglobulin refolding from 4 to 0.4 M GdmCl at 4.5°C. (Adapted from Ptitsyn and Semisotnov, 1991)





which reflects the collapse of the protein molecules. In both proteins, the collapse occurs much more slowly than does the formation of a substantial part of secondary structure; in both proteins, the CD molar ellipticity in the far-UV region changes drastically within the first 0.01 sec.

As has been mentioned already (Section 4), a sensitive test for the molten globule state is the binding of a hydrophobic probe, such as ANS (Semisotnov et al., 1987,1991a; Rodionova et al., 1989). The kinetics of ANS binding in carbonic anhydrase and in β-lactoglobulin, as measured by the increase in ANS fluorescence, consist of two phases (Semisotnov et al., 1991a; Ptitsyn and Semisotnov, 1991). The first phase occurs within the dead time of the experiment (about 0.01 sec), while the second practically coincides with the kinetics of restoration of energy transfer (Semisotnov et al., 1987; Ptitsyn and Semisotnov, 1991). The first phase may reflect ANS binding to the forming secondary structure and/or the first stage of molecular collapse, which is not sufficient to be monitored by energy transfer (Ptitsyn and Semisotnov, 1991). On the other hand, the coincidence of the second stage of ANS binding with the kinetics of restoration of energy transfer suggests that the globularization of a protein molecule leads to the formation of a nonpolar core (Semisotnov et al., 1987,1991a; Ptitsyn and Semisotnov, 1991). Therefore, the increase in ANS fluorescence can be used to monitor the formation of the molten globule state.

The left part of Figure 6–10 illustrates the increase of ANS binding that occurs upon the folding of six proteins (Ptitsyn et al., 1990; Goldberg et al., 1990). These data show that the globularization of a protein molecule, with the formation of a solvent-accessible nonpolar core, occurs in a time that varies from 0.05 sec (for β-lactamase at 23°C) to a few seconds (for the β_2 subunit of tryptophan synthase; Goldberg et al., 1990). As the time required for ANS binding to the molten globule state is less than 0.002 sec (Semisotnov et al., 1991a), these times reflect the rates of formation during protein folding of a kinetic intermediate with strong affinity for ANS.

7.4 Tertiary Structure

The right part of Figure 6–10 shows the decrease in ANS fluorescence that occurs during folding. This release of ANS from the protein suggests that the non- polar core of the protein is being screened from the solvent. At least in one case, for carbonic anhydrase (Semisotnov et al., 1987), it was shown that high field ^{1}H-NMR resonances appear simultaneously with ANS desorption. This suggests that the screening of the nonpolar core occurs simultaneously with its tight packing. The restoration of near-UV CD ellipticity and of protein activity in a number of proteins also occurs simultaneously with ANS desorption (Goldberg et al., 1990; Semisotnov and Kuwajima, unpublished data). In some cases, however, as with carbonic anhydrase (Semisotnov et

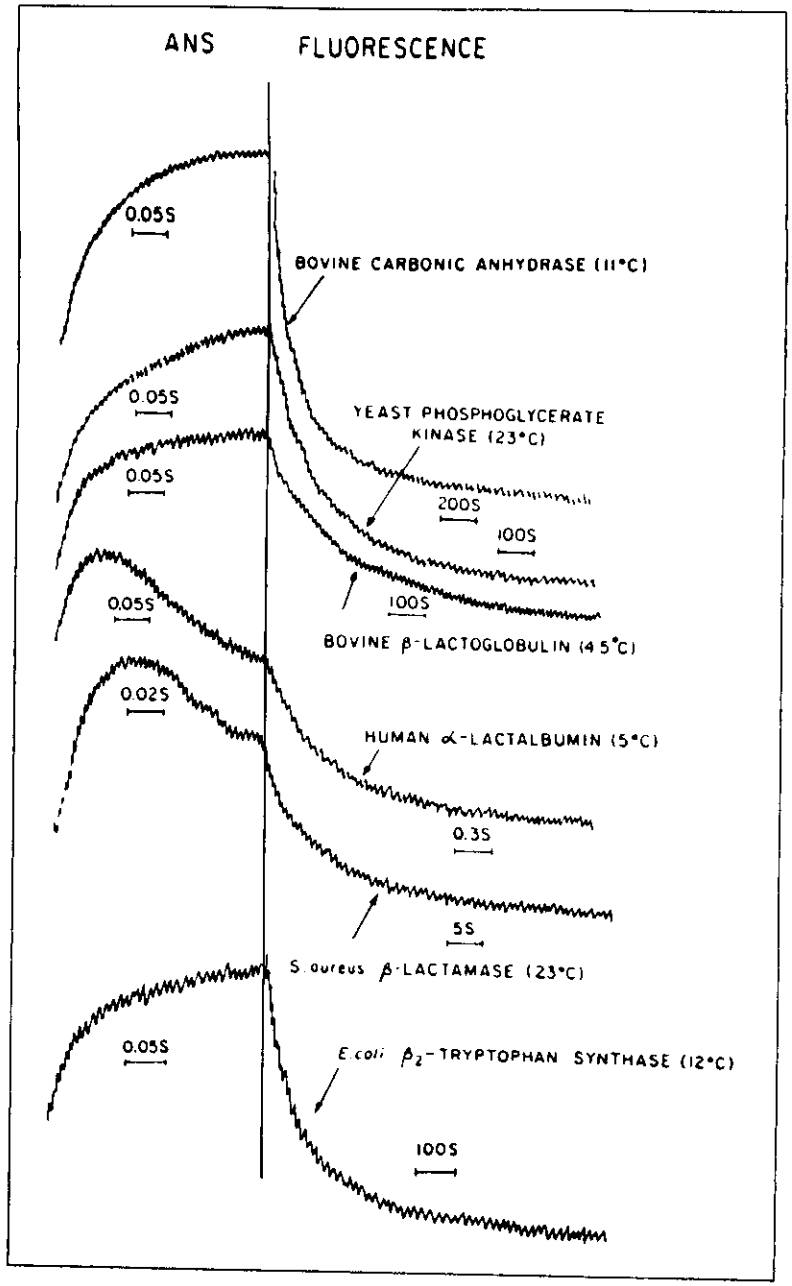


FIGURE 6–10. Refolding of globular proteins from urea or GdmCl monitored by ANS fluorescence. Different time scales are used for the left and the right parts of the figure. The final concentration of denaturant was from 1.95 to 4.0 M urea or from 0.07 to 0.7 M GdmCl.

al., 1987), the restoration of near-UV ellipticity and of enzymatic activity occurs more slowly than does ANS desorption and the appearance of high field ^{1}H-NMR resonance; this suggests that the solidification of a protein core can precede formation of the full native tertiary structure.

Slow phases of protein folding, which can be as slow as 500 to 2,500 sec at 25°C, are often due to *cis,trans* isomerization of peptide bonds preceding Pro residues (Kim and Baldwin, 1982,1990; Chapter 5) as described by the proposal of Brandts et al. (1975) which was extended by Semisotnov et al. (1990). This is not always the case, however; in horse (Betton et al., 1985), pig, and yeast (Semisotnov et al., 1991b) phosphoglycerate kinases, slow refolding stages do not depend upon the time of preincubation of the protein in the unfolded state and therefore are not due to Pro peptide bond isomerization.

7.5 The Molten Globule as a General Kinetic Intermediate

All proteins studies thus far have early kinetic intermediates with strong affinity for ANS, i.e., with a solvent-accessible nonpolar core. The compactness of these intermediates has been confirmed for some proteins by more direct methods: urea gradient electrophoresis, energy transfer, and viscosity measurements. (See above.) All of these kinetic intermediates have pronounced secondary structure, as indicated by their far-UV CD ellipticities (Kuwajima et al., 1987; Goldberg et al., 1990; Ptitsyn and Semisotnov, 1991).

On the other hand, these kinetic intermediates have no rigid tertiary structures. For example, ^{1}H-NMR spectra of kinetic intermediates of carbonic anhydrase (Semisotnov et al., 1987; Ptitsyn and Semisotnov, 1991) and of β-lactamase (Ptitsyn et al., 1990) are much simpler than are the spectra of the native proteins and are similar to the spectra of thermodynamically stable molten globules. Near-UV CD ellipticities of these kinetic intermediates are virtually absent (Robson and Pain, 1976a,b; McCoy et al., 1980; Dolgikh et al., 1984; Kuwajima et al., 1987,1988; Gilmanshin and Ptitsyn, 1987; Goldberg et al., 1990; Ptitsyn and Semisotnov, 1991). In addition, these intermediates are totally inactive enzymatically (Ikai et al., 1978; McCoy et al., 1980; Dolgikh et al., 1984; Mitchinson and Pain, 1985; Semisotnov et al., 1987; Ptitsyn and Semisotnov, 1991).

All of these data taken together leave little doubt that *a protein molecule while folding passes through an early kinetic intermediate that is very similar to the equilibrium molten globule state.*

The list of proteins for which these kinetic intermediates have been observed includes carbonic anhydrase, α-lactalbumin, β-lactoglobulin, β-lactamase, phosphoglycerate kinase (Ptitsyn et al., 1990; Ptitsyn and Semisotnov, 1991), and the β_2 subunit of tryptophan synthase (Goldberg et al., 1990). In addition, strong evidence for a molten globule kinetic intermediate has been obtained by Roder et al. (1988) for cytochrome *c* and by Matouschek et al. (1989,1990) and Bycroft et al. (1990) for barnase (Section 7.6). These proteins belong to different structural types and include both single- and multidomain proteins, as well as proteins with and without disulfide bridges. Therefore, we can conclude that *the molten globule state is a general intermediate in protein folding* (Ptitsyn et al., 1990; Ptitsyn and Semisotnov, 1991).

In addition, there are a number of proteins including RNase A (Schmid and Baldwin, 1979; Kim and Baldwin, 1980), lysozyme (Kuwajima et al., 1985; Ikeguchi et al., 1986), the α subunit of tryptophan synthase (Beasty and Matthews, 1985), growth hormone (Brems et al., 1987) and parvalbumin (Kuwajima et al., 1988) for which it has been shown that their early kinetic intermediates have secondary structure but do not have rigid tertiary structures although the compactness of these intermediates has not been demonstrated.

7.6 Native-like Features of Structural Organization

Now we know (Ptitsyn et al., 1990; Ptitsyn and Semisotnov, 1991) that protein folding includes at least three main stages:

1. Formation of a substantial part of the secondary structure ($<10^{-2}$ sec);
2. Formation of a compact (molten globule) intermediate (≤1 sec);
3. Formation of a rigid tertiary structure (1 to 10^3 sec).

These three stages coincide with the predictions made by Ptitsyn (1973) and occur in the folding of all proteins studied by his group. Sometimes, the last stage can split into two or more stages—the formation of the rigid structure of a protein core can precede the formation of the full tertiary structure described previously. Table 6–3 illustrates these three stages for carbonic anhydrase.

These main stages of protein folding, monitored by methods sensitive to the structure of a molecule as a whole, give a "time frame" for the consideration of other, more specific events that can occur during protein folding and reflect the formation of some native-like features of structural organization.

A powerful method of looking for these specific events is to use hydrogen exchange to label protected NH groups of kinetic intermediates, with subsequent identification of the protected groups in the refolded protein by 2D-NMR (Udgaonkar and Baldwin, 1988; Roder et al., 1988). These studies identify NH groups that have been protected from exchange, and therefore were involved in some structure, at different stages of folding.

TABLE 6–3. Folding Stages of Bovine Carbonic Anhydrase B (4.2 M urea, 23°C)

Formation of	*Techniques*	*Half-time $t_{1/2}$ (sec)*[a]	*References*
Secondary structure	Ellipticity at 222 nm[b]	<<0.01	Semisotnov and Kuwajima, unpublished results
Molten globule (compact state)	Energy transfer	0.04	Semisotnov et al., 1987
	Spin-label immobilization	0.03	Semisotnov et al., 1987; Ebert et al., 1990
Molten globule (solvent-accessible nonpolar core)	ANS binding	0.04	Semisotnov et al., 1987,1991a
Screened and tightly packed nonpolar core	ANS desorption	140	Semisotnov et al., 1987,1991a
	High-field NMR	140	
Native tertiary structure	Ellipticity at 270 nm	580	Semisotnov et al., 1987
	Esterase activity	530	

[a] For monoexponential process with rate constant k, the half-time is given by $t_{1/2} = (\ln 2)/k$.
[b] Ellipticity at 222 nm can also be influenced by the contribution of aromatic side chains.

Interpretation of these data is based on the assumption that "most highly protected amide protons in native proteins are both hydrogen-bonded and inaccessible to solvent" (Udgaonkar and Baldwin, 1988). A good illustration is that in the molten globule state only NH groups involved in α-helices are protected but their protection is much weaker than in the native state. (See Section 3.5.) This suggests that both the presence of secondary structure and its screening from solvent are important for substantial protection from exchange.

Using this method, it has been shown that the NH groups of residues that are part of β-sheets and of the main α-helices in the native conformations of RNase A (Udgaonkar and Baldwin, 1988,1990) and of barnase (Bycroft et al., 1990) are protected at early stages of folding, while NH groups involved in the final tertiary structure are protected later. These data suggest that the secondary structure that is formed at the beginning of the folding process is present in the same positions in the polypeptide chain as in the native conformation, in accordance with the framework model (Udgaonkar and Baldwin, 1988).

Even more interesting results have been obtained for cytochrome *c* (Roder et al., 1988), where the NH groups of the N- and C-terminal α-helices became protected very early in folding, with a half-time of approximately 10^{-2} sec. That these α-helices are protected simultaneously (Roder et al., 1988) and that their protection follows the same pH-dependence

(Roder, private communication) suggests that protection reflects their docking, rather than their independent formation. This is extremely important evidence that the basic feature of the tertiary fold of this simple protein is formed at a very early stage of its folding pathway. It is especially interesting that the rate of this packing is close to the rate of the first stage of fluorescence quenching of the single Trp residue, number 59, as a result of its proximity to the heme group, which suggests that "association of chain termini may be accompanied by a general condensation of the initially unfolded chain" (Roder et al., 1988). In this case also, NH groups involved in the tertiary structure are protected later.

It is worthwhile to mention that the NH groups of some elements of secondary structure, for example, the α-helices of residues 60 to 69 and 71 to 75 in cytochrome *c* (Roder et al., 1988) and the N-terminal α-helix in RNase A (Udgaonkar and Baldwin, 1990), become protected only at the same time as restoration of the tertiary structure, and much later than the other elements of secondary structure.

The protein for which the general time frame of folding has been most carefully compared with specific events is the β_2 subunit of tryptophan synthase, which has been extensively studied using energy transfer (Blond and Goldberg, 1986) and monoclonal antibodies specific for the native protein (Blond and Goldberg, 1987; Murray-Brelier and Goldberg, 1988; Blond-Elguindi and Goldberg, 1990) to monitor specific events of protein folding, such as the approach of two pairs of residues and the formation of an antigenic determinant. These specific events have been compared to the restoration of the overall characteristics of protein structure (Goldberg et al., 1990). The results presented in Table 6–4 show that an antigenic determinant is formed within the molten globule state, before formation of the rigid tertiary structure, while the close interaction of Trp177 with Lys87, which are well separated along the polypeptide chain, occurs simultaneously with the first stage in formation of the rigid tertiary structure.

A new interesting approach to the study of the role of various interactions in protein folding has been proposed by Matouschek et al. (1989,1990). By changing different residues by site-specific mutagenesis, they have determined the contribution of these residues to the free energies of three stages of protein folding: a kinetic intermediate (before the free energy barrier), the transition state (at the barrier), and the fully folded state (after the barrier). Applying this approach to barnase, and combining it with NMR data, they came to the following conclusions:

1. All three α-helices and all β-strands of barnase are already formed in the early kinetic intermediate;
2. The C-terminus of the first α-helix (residues 6 to 18) also is stabilized in this intermediate;

TABLE 6-4. Folding Stages of β_2 Subunit of *E.coli* Tryptophan Synthase (0.5 M urea; 12°C)

		Time Interval											
		<0.01 sec		0–5 sec		0–50 sec		0–200 sec		0–3500 sec			
Formation of:	*Technique*	$t_{1/2}$ (sec)[a]	*A* (%)[b]	$t_{1/2}$ (sec)	*A* (%)	$t_{1/2}$ (sec)	*A* (%)	$t_{1/2}$ (sec)	*A* (%)	$t_{1/2}$ (sec)	*A* (%)	*References*
Secondary structure	Ellipticity at 225 nm[c]	<0.007	57	—	—	5	7	70	29	700	7	Goldberg et al., 1990
Molten globule (solvent-accessible nonpolar core)	ANS binding	—	—	0.03 0.2 1	40 28 32	—	—	—	—	—	—	Goldberg et al., 1990
Native-like antigenic determinant	Monoclonal antibody binding	—	—	—	—	12	100	—	—	—	—	Murry-Brelier and Goldberg, 1988; Blond-Elguindi and Goldberg, 1990
Closing Trp 177 with Lys 87	Energy transfer	—	—	—	—	—	—	35	100	—	—	Blond and Goldberg, 1986
Closing Trp 177 with Cys 170	Energy transfer	—	—	—	—	—	—	90	100	—	—	Blond and Goldberg, 1986
Screened nonpolar core	ANS desorption	—	—	—	—	—	—	35	30	700	70	Goldberg et al., 1990
Native tertiary structure	Ellipticity at 285 nm	—	—	—	—	—	—	35	43	700	57	Goldberg et al., 1990

[a] Half-time of the process.
[b] Amplitude of the process.
[c] Slow change of ellipticity at 225 nm may be determined or at least influenced by contribution of aromatic side chains.

3. The hydrophobic core of barnase is present in the early intermediate, becomes more condensed in the transition state, and reaches its final tight packing in the folded state;
4. The N-termini of the first (residues 6 to 18) and the second (residues 26 to 34) α-helices become stabilized only in the final folded state.

These data are consistent with the kinetic intermediate in barnase being the molten globule state and with the packing of nonpolar groups in the protein core increasing in the sequence molten globule → transition state → folded state, as predicted by the theory of protein denaturation of Shakhnovich and Finkelstein (1989). They contain much additional information, as they present perhaps a detailed picture of what happens at different stages of protein folding.

These data indicate that the native features of protein structure may form at different stages of protein folding. Some features are formed simultaneously with the molten globule, such as the docking of the N- and C-terminal helices in cytochrome *c*, others are formed within the molten globule state such as the antigenic determinant in β_2 subunit of tryptophan synthase, and some (such as the N-termini of α-helices in barnase) occur only in the final native state. Many more experiments are needed to clarify this complicated picture.

8 POSSIBLE ROLES OF THE MOLTEN GLOBULE

Is the molten globule state described in this chapter only a curious feature of proteins *in vitro* or is it involved with proteins in a living cell? Trying to answer this question, it must be remembered that a dilute solution of one protein in a "native" buffer is far from being the ideal model for proteins in a cell. In fact, proteins in a living cell can be subjected to different mild denaturing influences, such as high temperatures, low pH, membranes, and so forth. Therefore, it is not meaningless to consider the possible roles of "slightly denatured" proteins, including the molten globule state, in a living cell.

1. The typical time for biosynthesis of a polypeptide chain is approximately 10 to 100 sec. It is greater than the time for formation of the molten globule state (about 1 sec), but much less than the time for maturation of rigid native structure in some proteins, which can be as long as 10^3 sec. (See Section 7.4.) Therefore, a nascent polypeptide chain during and immediately after its biosynthesis may be in the molten globule state. Cells may even want to keep a nascent protein in this state, as it is sufficiently flexible to adjust to subsequent events,

such as oligomerization, transmembrane transport, and so forth. In fact, it is known that:

a. Heat-shock GroEL protein binds and stabilizes a non-native state of a nascent polypeptide chain (Bochkareva et al., 1988; Rothman, 1989);
b. Some heat-shock proteins bind inactive subunits of oligomeric proteins to prevent them from nonspecific aggregation and/or to transport them to the place of their assembly (Pelham, 1986; Ellis, 1987,1990; Schlesinger, 1990);
c. Some heat-shock proteins bind secretory and imported proteins to prevent them from premature tight folding or aggregation (Deshaies et al., 1988) and are involved in transmembrane protein translocation (Goloubinoff et al., 1989; Scherer et al., 1990; Lazdunski and Benedetti, 1990).

Therefore, it is reasonable to assume that *some heat-shock proteins, particularly GroEL, recognize and bind a nascent protein in the molten globule state* (Bychkova et al., 1988). Recently, this assumption has been confirmed by Semisotnov et al. (1992), who have shown that GroEL retards protein folding, both when it is added at the time when protein is completely unfolded and when it is added 1 to 2 sec later, when the protein is already folded into the molten globule state.

2. There is numerous evidence (Vestweber and Schatz, 1988; Vestweber et al., 1989; Lazdunski and Benedetti, 1990) that proteins cannot be translocated through membranes in the native state, but are competent for translocation when in a non-native state in which they are susceptible to proteolysis. This has given rise to the idea (Bychkova et al., 1988) that the molten globule state might be involved in protein translocation through membranes. It is confirmed by the evidence that proteins can bind to membranes or micelles at low pH (e.g., Zhao and London, 1986; Kim and Kim, 1986), when at least some of them might be in the denatured, molten globule state. It is possible that the interactions of proteins with the membrane surface trigger their transition to the molten globule state, in which they can easily adapt, perhaps by partial unfolding, to various environments, such as water, membrane surface, and the inner part of a membrane. There is evidence (Rassow et al., 1990; Neupert et al., 1990) that the part of the protein chain that is transversing the contact site of mitochondrial membranes is in an extended conformation.
3. It is possible that protein degradation in lysosomes, at acid pH, or by ATP-dependent proteosomes, as in the ubiquitin-dependent system, can be facilitated by prior denaturation, which may be the transition to the molten globule state. It is interesting that a heat-shock protein is involved in lysosomal protein degradation (Chiang et al., 1989).

Although very little is known about non-native protein states in a living cell, the possibility that they might play a role in cell processes is intriguing.

9 CONCLUSION

In summarizing this chapter, it should be emphasized that there is now no doubt that the molten globule state actually exists as a separate equilibrium state of protein molecules, that it occurs frequently in denatured proteins, and that it plays an important role in protein folding. The main question that has to be answered is to what extent the equilibrium molten globule state possesses the native tertiary fold of the corresponding protein and at what stage in the formation or maturation of the molten globule–like kinetic intermediate it achieves the native tertiary fold. A very intriguing question, of course, is the possible role of the molten globule state in protein life in a living cell.

Acknowledgments

Sections 6 and 8 of this chapter were written together with Dr. V. E. Bychkova, who also prepared Table 6–2. Drs. A. V. Finkelstein and G. V. Semisotnov carefully read the manuscript and made many valuable comments. My wife, I. G. Ptitsyna, and M. S. Shelestova have rendered important assistance in preparing the manuscript. I am extremely grateful to all of them for their generous help.

Note

Recently two important papers have been published that directly confirmed the suggestion (Bychkova et al., 1988) that the molten globule may be physiologically important. J. Martin et al. (1991, *Nature 352*, 36–42) have shown that dihydrofolate reductase and rhodanese being in the complexes with GroEL are in the molten globule state. They came to the conclusion that "a molten globule-like pre-folded state is a physiological, early intermediate of chaperonin-mediate folding." F. G. van der Goot et al. (1991, *Nature 354*, 408–410) have shown that pH-dependence of the membrane insertion of the pore-forming domain of colicin A correlates with its transition into the molten globule state. This paper documents a local pH decrease near negatively charged membrane surface (see p. 282), which has been measured and taken into account. Thus the idea of the physiological role of the molten globule state may be meaningful.

REFERENCES

Acharya, K. R., Stuart, D. I., Walker, N. P. C., Lewis, M., and Phillips, D. C. (1989) *J. Mol. Biol. 208*, 99–127.

Ahmad, F., and Bigelow, C. C. (1979) *J. Mol. Biol. 131*, 607–617.

Amir, D., and Haas, E. (1988) *Biochemistry 27*, 8889–8893.

Ananthanarayanan, V. S., and Ahmad, F. (1977) *Can. J. Biochem. 55*, 239–243.

Ananthanarayanan, V. S., Ahmad, F., and Bigelow, C. C. (1977) *Biochim. Biophys. Acta 492*, 194–203.

Anufrieva, E. V. (1982) *Pure Appl. Chem. 54*, 533–584.

Anufrieva, E. V., Bychkova, V. E., Krakovyak, M. G., Pautov, V. D., and Ptitsyn, O. B. (1975) *FEBS Letters 55*, 46–49.

Anufrieva, E. V., and Krakovyak, M. G. (1987) *Vysokomol. Soed. (USSR) A29*, 211–222.

Anufrieva, E. V., Volkenstein, M. V., Gotlib, Yu. Ya., Krakovyak, M. G., Pautov, V. D., Stepanov, V. V., and Skorokhodov, S. S. (1972) *Dokl. Akad. Nauk SSSR 207*, 1379–1382.

Arakawa, T., Hsu, Y. -R., Yphantis, D. A. (1987) *Biochemistry 26*, 5428–5432.

Aune, K. C., Salahuddin, A., Zarlengo, M. H., and Tanford, C. (1967) *J. Biol. Chem. 242*, 4486–4489.

Baldwin, R. L. (1986) *Proc. Natl. Acad. Sci. USA 83*, 8069–8072.

Baldwin, R. L. (1990) *Nature 346*, 409–410.

Baldwin, R. L. (1991) in CIBA Foundation Symposium 161, *"Protein conformation,"* London, January 22–24, 1991.

Baum, J., Dobson, C. M., Evans, P. A., and Hanly, C. (1989) *Biochemistry 28*, 7–13.

Beasty, A. M., and Matthews, C. R. (1985) *Biochemistry 24*, 3547–3553.

Bendzko, P. I., Pfeil, W. A., Privalov, P. L., and Tiktopulo, E. I. (1988) *Biophys. Chem. 29*, 301–307.

Betton, J.-M., Desmadril, M., Mitraki, A., and Yon, J. M. (1985) *Biochemistry 24*, 4570–4577.

Bewley, T. A., Levine, H. L., and Wetzel, R. (1982) *Int. J. Peptide Protein Res. 20*, 93–96.

Birshtein, T. M., and Pryamitsyn, V. A. (1987) *Vysokomol. Soed. (USSR) A29*, 1858–1864.

Birshtein, T. M., and Pryamitsyn, V. A. (1991) *Macromolecules 24*, 1554–1560.

Blond, S., and Goldberg, M. E. (1986) *Proteins: Struct. Funct. Genet. 1*, 247–255.

Blond, S., and Goldberg, M. E. (1987) *Proc. Natl. Acad. Sci. USA 84*, 1147–1151.

Blond-Elguindi, S., and Goldberg, M. E. (1990) *Biochemistry 29*, 2409–2417.

Bochkareva, E. S., Lissin, N. M., and Girshovich, A. S. (1988) *Nature 336*, 254–257.

Bolotina, I. A. (1987) *Mol. Biol. (USSR) 21*, 1625–1635.

Bolotina, I. A., and Lugauskas, V. Yu. (1985) *Mol. Biol. (USSR) 19*, 1409–1421.

Brahms, S., and Brahms, J. (1980) *J. Mol. Biol. 138*, 149–178.

Brandts, J. F., Halvorson, H. R., and Brennan, M. (1975) *Biochemistry 14*, 4953–4963.

Brandts, J. F., and Hunt, L. (1967) *J. Am. Chem. Soc. 89*, 4826–4838.

Brazhnikov, E. V., Chirgadze, Yu. N., Dolgikh, D. A., and Ptitsyn, O. B. (1985) *Biopolymers 24*, 1899–1907.

Brems, D. N., and Havel, H. A. (1989) *Proteins: Struct. Funct. Genet. 5*, 93–95.

Brems, D. N., Plaisted, S. M., Dougherty, J. J. Jr., and Holzman, T. F. (1987) *J. Biol. Chem. 262*, 2590–2596.

Brems, D. N., Plaisted, S. M., Havel, H. A., Kauffman, E. W., Stodola, J. D., Eaton, L. C., and White, R. D. (1985) *Biochemistry 24*, 7662–7668.

Burger, H. G., Edelhoch, H., and Condliffe, P. G. (1966) *J. Biol. Chem. 241*, 449–457.

Bushnell, G. W., Louie, G. V., and Brayer, G. D. (1990) *J. Mol. Biol. 214*, 585–595.

Bychkova, V. E., Bartoshevich, S. F., and Klenin, S. I. (1990) *Biofizika (USSR) 35*, 242–248.

Bychkova, V. E., Berni, R., Rossi, G. L., Kutyshenko, V. P., and Ptitsyn, O. B. (1992) *Biochemistry*, in press.

Bychkova, V. E., Pain, R. H., and Ptitsyn, O. B. (1988) *FEBS Letters 238*, 231–234.

Bychkova, V. E., and Ptitsyn, O. B. (1992) *Biofizika (USSR)*, in press.

Bychkova, V. E., Semisotnov, G. V., Ptitsyn, O. B., Gudkova, O. V., Mitin, Yu. V., and Anufrieva, E. V. (1980) *Mol. Biol. (USSR) 14*, 278–286.

Bycroft, M., Matouschek, A., Kellis, J. T. Jr., Serrano, L., and Fersht, A. R. (1990) *Nature 346*, 488–490.

Cabiaux, V., Brasseur, R., Wattiez, R., Falmagne, P., Ruysschaert, J. -M., and Goormaghtigh, E. (1989) *J. Biol. Chem. 264*, 4928–4938.

Cavard, D., Sauve, P., Heitz, F., Pattus, F., Martinez, C., Dijkman, R., and Lazdunski, C. (1988) *Eur. J. Biochem. 172*, 507–512.

Cave, A., Pages, M., Morin, P., and Dobson, C. M. (1979) *Biochimie 61*, 607–613.

Chan, H. S., and Dill, K. A. (1991) *Ann. Rev. Biophys. Biophys. Chem. 20*, 447–490.

Chiang, H.-L., Terlecky, S. R., Plant, C. P., and Dice, J. F. (1989) *Science 246*, 382–385.

Chirgadze, Yu. N., Fedorov, O. V., and Trushina, N. P. (1975) *Biopolymers 14*, 679–694.

Creighton, T. E. (1978) *Progr. Biophys. Mol. Biol. 33*, 231–297.

Creighton, T. E. (1980) *J. Mol. Biol. 137*, 61–80.

Creighton, T. E. (1990) *Biochem. J. 270*, 1–16.

Creighton, T. E., and Pain, R. H. (1980) *J. Mol. Biol. 137*, 431–436.

Damaschun, G., Damaschun, H., Gast, K., Zizwer, D., and Bychkova, V. E. (1991) *Int. J. Biol. Macromol. 13*, 217–221.

Damaschun, G., Gernat, C., Damaschun, H., Bychkova, V. E., and Ptitsyn, O. B. (1986) *Int. J. Biol. Macromol. 8*, 226–230.

De Gennes, P.-G. (1975) *J. Phys. Letters 36*, 55–57.

Denton, J. B., Konishi, Y., and Scheraga, H. A. (1982) *Biochemistry 21*, 5155–5163.

Deshaies, R. J., Koch, B. D., Werner-Washburne, M., Craig, E. A., and Schekman, R. (1988) *Nature 332*, 800–805.

Desmadril, M., and Yon, J. M. (1981) *Biochem. Biophys. Res. Commun. 101*, 563–569.

Dill, K. A. (1985) *Biochemistry 24*, 1501–1509.

Dill, K. A., Alonso, D. O. U., and Hutchinson, K. (1989) *Biochemistry 28*, 5439–5449.

Dill, K. A., and Shortle, D. (1991) *Ann. Rev. Biochem. 60*, 795–825.

Dobson, C. M., Hanley, C., Radford, S. E., Baum, J. A., and Evans, P. A. (1991) in *Conformations and Forces in Protein Folding* (B. T. Nall and K. A. Dill, eds.), AAAS, Washington, D.C., pp. 175–181.

Dolgikh, D. A., Abaturov, L. V., Bolotina, I. A., Brazhnikov, E. V., Bychkova, V. E., Bushuev, V. N., Gilmanshin, R. I., Lebedev, Yu. O., Semisotnov, G. V., Tiktopulo, E. I., and Ptitsyn, O. B. (1985) *Eur. Biophys. J. 13*, 109–121.

Dolgikh, D. A., Abaturov, L. V., Brazhnikov, E. V., Lebedev, Yu. O., Chirgadze, Yu. N., and Ptitsyn, O. B. (1983) *Dokl. Akad. Nauk SSSR 272*, 1481–1484.

Dolgikh, D. A., Gilmanshin, R. I., Brazhnikov, E. V., Bychkova, V. E., Semisotnov, G. V., Venyaminov, S. Yu., and Ptitsyn, O. B. (1981) *FEBS Letters 136*, 311–315.

Dolgikh, D. A., Kolomiets, A. P., Bolotina, I. A., and Ptitsyn, O. B. (1984) *FEBS Letters 165*, 88–92.
Dumont, M. E., and Richards, F. M. (1988) *J. Biol. Chem. 263*, 2087–2097.
Ebert, B., Semisotnov, G. V., and Rodionova, N. A. (1990) *Studia Biophys. 137*, 125–131.
Echols, G. H., and Anderegg, J. W. (1960) *J. Am. Chem. Soc. 82*, 5085–5093.
Ellis, R. J. (1987) *Nature 328*, 378–379.
Ellis, R. J. (1990) *Science 250*, 954–959.
Fedorov, B. A., and Ptitsyn, O. B. (1977) *Dokl. Akad. Nauk SSSR 233*, 716–718.
Fink, A. L., Calciano, L. J., Goto, Y., and Palleros, D. R. (1990) in *Current Research in Protein Chemistry* (J. Villafranca, ed.), Academic Press, New York, pp. 417–424.
Fink, A. L., Calciano, L. J., Goto, Y., and Palleros, D. R. (1991) in *Conformations and Forces in Protein Folding* (B. T. Nall and K. A. Dill, eds.), AAAS, Washington, D. C., pp. 169–174.
Finkelstein, A. V., Badretdinov, A. Ya., and Ptitsyn, O. B. (1991) *Proteins: Struct. Funct. Genet. 10*, 287–299.
Finkelstein, A. V., and Ptitsyn, O. B. (1987) *Progr. Biophys. Mol. Biol. 50*, 171–190.
Finkelstein, A. V., and Shakhnovich, E. I. (1989) *Biopolymers 28*, 1681–1694.
Fischer, G., and Schmid, F. X. (1990) *Biochemistry 29*, 2205–2212.
Flory, P. J. (1953) *Principles of Polymer Chemistry*, Cornell University Press, Ithaca, New York.
Foster, J. F. (1960) in *Plasma Proteins*, vol. 1, Academic Press, New York, pp. 179–239.
Gast, K., Zirwer, D., Welfle, H., Bychkova, V. E., and Ptitsyn, O. B. (1986) *Int. J. Biol. Macromol. 8*, 231–236.
Gernat, C., Damaschun, G., Kröber, R., Bychkova, V. E., and Ptitsyn, O. B. (1986) *Studia Biophys. 112*, 213–219.
Gilmanshin, R. I. (1985) *Biofizika (USSR) 30*, 581–587.
Gilmanshin, R. I. (1988) *Biofizika (USSR) 33*, 27–30.
Gilmanshin, R. I., Dolgikh, D. A., Ptitsyn, O. B., Finkelstein, A. V., and Shakhnovich, E. I. (1982) *Biofizika (USSR) 27*, 1005–1016.
Gilmanshin, R. I., and Ptitsyn, O. B. (1987) *FEBS Letters 223*, 327–329.
Goldberg, M. E., Semisotnov, G. V., Friguet, B., Kuwajima, K., Ptitsyn, O. B., and Sugai, S. (1990) *FEBS Letters 263*, 51–56.
Goloubinoff, P., Christeller, J. T., Gatenby, A. A., and Lorimer, G. H. (1989) *Nature 342*, 884–889.
Goto, Y., Calciano, L. J., and Fink, A. L. (1990a) *Proc. Natl. Acad. Sci. USA 87*, 573–577.
Goto, Y., and Fink, A. L. (1989) *Biochemistry 28*, 945–952.
Goto, Y., and Fink, A. L. (1990) *J. Mol. Biol. 214*, 803–805.
Goto, Y., Takahashi, N., and Fink, A. L. (1990b) *Biochemistry 29*, 3480–3488.
Grigoryev, A. I., Volkova, L. A., and Ptitsyn, O. B. (1971) *FEBS Letters 15*, 217–219.
Griko, Yu. V., Privalov, P. L., Venyaminov, S. Yu., and Kutyshenko, V. P. (1988) *J. Mol. Biol. 202*, 127–138.
Grosberg, A. Yu. (1984) *Biofizika (USSR) 29*, 569–573.
Grosberg, A. Yu., and Shakhnovich, E. I. (1986a) *Zh. Exp. Theor. Fiz. (USSR) 91*, 2159–2170.
Grosberg, A. Yu., and Shakhnovich, E. I. (1986b) *Biofizika (USSR) 31*, 1054–1057.
Harding, M. M., Williams, D. H., and Woolfson, D. N. (1991) *Biochemistry 30*, 3120–3128.
Herold, M., and Kirschner, K. (1990) *Biochemistry 29*, 1907–1913.
Holladay, A., Hammonds, R. G., Jr., and Puett, D. (1974) *Biochemistry 13*, 1653–1661.
Holzman, T. F., Brems, D. N., and Dougherty, J. J., Jr. (1986) *Biochemistry 25*, 6907–6917.
Holzman, T. F., Dougherty, J. J., Jr., Brems, D. N., and MacKenzie, N. E. (1990) *Biochemistry 29*, 1255–1261.
Horowitz, P. M., and Criscimagna, N. L. (1990) *J. Biol. Chem. 265*, 2576–2583.
Hughson, F. M., and Baldwin, R. L. (1989) *Biochemistry 28*, 4415–4422.
Hughson, F. M., Barrik, D., and Baldwin, R. L. (1991) *Biochemistry 30*, 4113–4118.
Hughson, F. M., Wright, P. E., and Baldwin, R. L. (1990) *Science 249*, 1544–1548.
Ikai, A., Tanaka, S., and Noda, H. (1978) *Arch. Biochem. Biophys. 190*, 39–45.
Ikeguchi, M., Kuwajima, K., Mitani, M., and Sugai, S. (1986) *Biochemistry 25*, 6965–6972.
Izumi, Y., Miyake, Y., Kuwajima, K., Sugai, S., Inoue, K., Iizumi, M., and Katano, S. (1983) *Physica 120 B*, 444–448.
Jagannadham, M. V., and Balasubramanian, D. (1985) *FEBS Letters 188*, 326–330.
Jeng, M. F., Englander, S. W., Elöve, G. A., Wang, A. J., and Roder, H. (1990) *Biochemistry 29*, 10433–10437.
Kim, J., and Kim, H. (1986) *Biochemistry 25*, 7867–7874.
Kim, P. S., and Baldwin, R. L. (1980) *Biochemistry 19*, 6124–6129.
Kim, P. S., and Baldwin, R. L. (1982) *Ann. Rev. Biochem. 51*, 459–489.
Kim, P. S., and Baldwin, R. L. (1990) *Ann. Rev. Biochem. 59*, 631–660.
Kosen, P. A., Creighton, T. E., and Blout, E. R. (1981) *Biochemistry 20*, 5744–5754.
Kosen, P. A., Creighton, T. E., and Blout, E. R. (1983) *Biochemistry 22*, 2433–2440.
Kron, A. K., Ptitsyn, O. B., Skvortsov, A. M., and Fedorov, A. K. (1967) *Mol. Biol. (USSR) 1*, 576–582.
Kronman, M. J., Blum, R., and Holmes, L. G. (1966) *Biochemistry 5*, 1970–1978.
Kuwajima, K. (1977) *J. Mol. Biol. 114*, 241–258.
Kuwajima, K. (1989) *Proteins: Struct. Funct. Genet. 6*, 87–103.
Kuwajima, K., Harushima, Y., and Sugai, S. (1986) *Int. J. Peptide Protein Res. 27*, 18–27.
Kuwajima, K., Hiraoka, Y., Ikeguchi, M., and Sugai, S. (1985) *Biochemistry 24*, 874–881.
Kuwajima, K., Nitta, K., Yoneyama, M., and Sugai, S. (1976) *J. Mol. Biol. 106*, 359–373.
Kuwajima, K., Sakuraoka, A., Fueki, S., Yoneyama, M., and Sugai, S. (1988) *Biochemistry 27*, 7419–7428.
Kuwajima, K., Yamaya, H., Miwa, S., Sugai, S., and Nagamura, T. (1987) *FEBS Letters 221*, 115–118.
Lazdunski, C. J., and Benedetti, H. (1990) *FEBS Letters 268*, 408–414.
Lee, J. W., and Kim, H. (1988) *FEBS Letters 241*, 181–184.
Lifshitz, I. M., Grosberg, A. Yu., and Khokhlov, A. R. (1978) *Rev. Modern Phys. 50*, 683–713.
Lumry, R. T., Biltonen, R., and Brandts, J. F. (1966) *Biopolymers 4*, 917–944.
Manning, M. C., and Woody, R. W. (1989) *Biochemistry 28*, 8609–8613.
Matouschek, A., Kellis, J. T., Jr., Serrano, L., Bycroft, M., and Fersht, A. R. (1990) *Nature 346*, 440–445.
Matouschek, A., Kellis, J. T., Jr., Serrano, L., and Fersht, A. R. (1989) *Nature 340*, 122–126.
McCoy, L. F., Rowe, E. S., and Wong, K. -P. (1980) *Biochemistry 19*, 4738–4743.
McLaughlin, S. (1989) *Ann. Rev. Biophys. Biophys. Chem. 18*, 113–136.

Merrill, A. R., Cohen, F. S., and Cramer, W. A. (1990) *Biochemistry 29*, 5829–5836.
Mitchinson, C., and Pain, R. H. (1985) *J. Mol. Biol. 184*, 331–342.
Murry-Brelier, A., and Goldberg, M. E. (1988) *Biochemistry 27*, 7633–7640.
Neupert, W., Harte, F.-U., Craig, E. A., and Pfanner, N. (1990) *Cell 63*, 447–450.
Nozaka, M., Kuwajima, K., Nitta, K., and Sugai, S. (1978) *Biochemistry 17*, 3753–3758.
Ohgushi, M., and Wada, A. (1983) *FEBS Letters 164*, 21–24.
Pelham, H. R. B. (1986) *Cell 46*, 959–961.
Permyakov, E. A., Grishchenko, V. M., Kalinichenko, L. P., Orlov, N. Y., Kuwajima, K., and Sugai, S. (1991) *Biophys. Chem. 39*, 111–117.
Perutz, M. F., Kendrew, J. C., and Watson, H. C. (1965) *J. Mol. Biol. 13*, 669–678.
Pfeil, W., Bychkova, V. E., and Ptitsyn, O. B. (1986) *FEBS Letters 198*, 287–291.
Potekhin, S. A., and Pfeil, W. (1989) *Biophys. Chem. 34*, 55–62.
Privalov, P. L. (1974) *FEBS Letters 40*, S140–S153.
Privalov, P. L. (1979) *Adv. Protein Chem. 33*, 167–241.
Privalov, P. L. (1982) *Adv. Protein Chem. 35*, 1–104.
Privalov, P. L., and Khechinashvili, N. N. (1974) *J. Mol. Biol. 86*, 665–684.
Privalov, P. L., and Makhatadze, G. I. (1990) *J. Mol. Biol. 213*, 385–391.
Privalov, P. L., Plotnikov, V. V., and Filimonov, V. V. (1975) *J. Chem. Thermodyn. 7*, 41–47.
Ptitsyn, O. B. (1973) *Dokl. Akad. Nauk SSSR 210*, 1213–1215.
Ptitsyn, O. B. (1975) *Dokl. Akad. Nauk SSSR 223*, 1253–1255.
Ptitsyn, O. B. (1985) *Suppl. J. Biosci.* (Proc. Int. Symp. Biomol. Struct. Interactions) *8*, 1–13.
Ptitsyn, O. B. (1987) *J. Protein Chem. 6*, 273–293.
Ptitsyn, O. B., Damaschun, G., Gernat, C., Damaschun, H., and Bychkova, V. E. (1986) *Studia Biophys. 112*, 207–211.
Ptitsyn, O. B., Dolgikh, D. A., Gilmanshin, R. I., Shakhnovich, E. I., and Finkelstein, A. V. (1983) *Mol. Biol. (USSR) 17*, 569–576.
Ptitsyn, O. B., and Eizner, Yu. Ye. (1965) *Biofizika (USSR) 10*, 3–6.
Ptitsyn, O. B., and Finkelstein, A. V. (1980a) *Quart. Rev. Biophys. 13*, 339–386.
Ptitsyn, O. B., and Finkelstein, A. V. (1980b) in *Protein Folding* (R. Jaenicke, ed.), Elsevier/North Holland Biomedical Press, Amsterdam-New York, pp. 101–115.
Ptitsyn, O. B., Kron, A. K., and Eizner, Yu. Ye. (1968) *J. Polymer Sci. Pt C 16*, 3509–3517.
Ptitsyn, O. B., Pain, R. H., Semisotnov, G. V., Zerovnik, E., and Razgulyaev, O. I. (1990) *FEBS Letters 262*, 20–24.
Ptitsyn, O. B., Reva, B. A., and Finkelstein, A. V. (1989) *Highlights of Modern Biochemistry* (Proc. 14th Intern. Congress of Biochemistry, Prague, Czechoslovakia, July 10–15, 1988), vol. 1 (A. Kotyk, J. Škoda, V. Paĉes, and V. Kostka, eds.), VSP, Utrecht-Tokyo, pp. 11–17.
Ptitsyn, O. B., and Semisotnov, G. V. (1991) in *Conformations and Forces in Protein Folding* (B. T. Nall and K. A. Dill, eds.), AAAS, Washington, D. C., pp. 155–168.
Rassow, J., Hartl, F.-U., Guiard, B., Pfanner, N., and Neupert, W. (1990) *FEBS Letters 275*, 190–194.
Robbins, F. M., and Holmes, L. G. (1970) *Biochim. Biophys. Acta 221*, 234–240.
Robson, B., and Pain, R. H. (1976a) *Biochem. J. 155*, 325–330.

Roder, H., Elöve, G. A., and Englander, S. W. (1988) *Nature 335*, 700–704.
Rodionova, N. A. (1990) The study of compactization of globular proteins. Ph.D. thesis, Moscow Physico-Technical Institute.
Rodionova, N. A., Semisotnov, G. V., Kutyshenko, V. P., Uversky, V. N., Bolotina, I. A., Bychkova, V. E., and Ptitsyn, O. B. (1989) *Mol. Biol. (USSR) 23*, 683–692.
Rothman, J. E. (1989) *Cell 59*, 591–601.
Sanchez, I. C. (1979) *Macromolecules 12*, 980–988.
Scherer, P. E., Krieg, U. C., Hwang, S. T., Vestweber, D., and Schatz, G. (1990) *EMBO J. 9*, 4315–4322.
Schlesinger, M. J. (1990) *J. Biol. Chem. 265*, 12111–12114.
Schmid, F. X., and Baldwin, R. L. (1979) *J. Mol. Biol. 135*, 199–215.
Sears, D. W., and Beychok, S. (1973) in *Circular Dichroism in Physical Properties and Techniques of Protein Chemistry*, part C (S. J. Leach, ed.), Academic Press, New York, pp. 445–593.
Semisotnov, G. V., Kutyshenko, V. P., and Ptitsyn, O. B. (1989) *Mol. Biol. (USSR) 23*, 808–815.
Semisotnov, G. V., Rodionova, N. A., Kutyshenko, V. P., Ebert, B., Blank, J., and Ptitsyn, O. B. (1987) *FEBS Letters 224*, 9–13.
Semisotnov, G. V., Rodionova, N. A., Razgulyaev, O. I., Uversky, V. N., Gripas, A. F., and Gilmanshin, R. I. (1991a) *Biopolymers 31*, 119–128.
Semisotnov, G. V., Sokolovsky, I. V., Bochkareva, E. S., and Girshovich, A. S. (1992), in preparation.
Semisotnov, G. V., Uversky, V. N., Sokolovsky, I. V., Gutin, A. M., Razgulyaev, O. I., and Rodionova, N. A. (1990) *J. Mol. Biol. 213*, 561–568.
Semisotnov, G. V., Vas, M., Chemeris, V. V., Kashparova, N. J., Kotova, N. V., Razgulyaev, O. I., and Sinev, M. A. (1991b) *Eur. J. Biochem.*, in press.
Semisotnov, G. V., Zikherman, K. Kh., Kasatkin, S. B., Ptitsyn, O. B., and Anufrieva, E. V. (1981) *Biopolymers 20*, 2287–2309.
Shakhnovich, E. I., and Finkelstein, A. V. (1982) *Dokl. Akad. Nauk SSSR 267*, 1247–1250.
Shakhnovich, E. I., and Finkelstein, A. V. (1989) *Biopolymers 28*, 1667–1680.
Shakhnovich, E. I., and Gutin, A. M. (1990) *Nature 346*, 773–775.
Shortle, D., and Meeker, A. K. (1989) *Biochemistry 28*, 936–944.
Spolar, R. S., Ha, J.-H., and Record, M. T., Jr., (1989) *Proc. Natl. Acad. Sci. USA 86*, 8382–8385.
Stegmann, T., Doms, R. W., and Helenius, A. (1989) *Ann. Rev. Biophys. Biophys. Chem. 18*, 187–211.
Stryer, L. (1965) *J. Mol. Biol. 13*, 482–495.
Sun, S. T., Nishio, I., Swislow, G., and Tanaka, T. (1980) *J. Chem. Phys. 73*, 5971–5975.
Tandon, S., and Horowitz, P. M. (1989) *J. Biol. Chem. 264*, 9859–9866.
Tanford, C. (1968) *Adv. Protein Chem. 23*, 121–282.
Timchenko, A. A., Dolgikh, D. A., Damaschun, H., and Damaschun, G. (1986) *Studia Biophys. 112*, 201–206.
Udgaonkar, J. B., and Baldwin, R. L. (1988) *Nature 335*, 694–699.
Udgaonkar, J. B., and Baldwin, R. L. (1990) *Proc. Natl. Acad. Sci. USA 87*, 8197–8201.
Uversky, V. N., Semisotnov, G. V., Pain, R. H., and Ptitsyn, O. B. (1992), in preparation.
Venyaminov, S. Yu., and Kalnin, N. N. (1991) *Biopolymers*, in press.
Vestweber, D., Brunner, J., Baker, A., and Schatz, G. (1989) *Nature 341*, 205–209.

References for Lectures on Optical Triggering in Protein Kinetics and Dynamics

William A. Eaton, Laboratory of Chemical Physics, National Institutes of Health, Bethesda

I. "*Photophysics, protein physics, and myoglobin - the hydrogen atom of biology*"

Theoretical interpretation of the optical spectra of hemes in myoglobin and hemoglobin
Eaton, Hanson, Stephens, Sutherland and Dunn, *J. Am. Chem. Soc. 100*, 4491 (1978).

Sub-picosecond observation of the photodissociation of heme complexes
Martin, Migus, Poyart, Lecarpentier, Astier, and Antonetti, *Proc. Natl. Acad. Sci. USA 80*, 173 (1983)
Anfinrud, Han, and Hochstrasser, *Proc. Natl. Acad. Sci. USA 86*, 8387 (1989).

Simulation of photodissociation, geminate rebinding, and laser heating by molecular dynamics
Henry, Levitt, and Eaton, *Proc. Natl. Acad. Sci. USA 82*, 2034 (1985)
Henry, Eaton, and Hochstrasser, *Proc. Natl. Acad. Sci. USA 83*, 8982 (1986).
Petrich, Lambry, Kuzcera, Karplus, Poyart, and Martin, *Biochemistry 30*, 3975 (1991)
Schaad, Zhou, Szabo, Eaton, and Henry, *Proc. Natl. Acad. Sci. USA 90*, 9547 (1993).
Li, Elber, and Straub, *J. Biol. Chem. 268*, 17908 (1993).

Discovery of conformational substates and multiple barriers
Austin, Beeson, Eisenstein, Frauenfelder, and Gunsalus, *Biochemistry 14*, 5355 (1975).
Frauenfelder, Sligar, and Wolynes, *Science* 254, 1598 (1991).

Discovery of geminate recombination of CO and O_2 in liquid solution
Greene, Hochstrasser, Weisman, and Eaton, *Proc. Natl. Acad. Sci. USA 75*, 5255 (1978)
Duddell, Morris, and Richards, *J. C. S. Chem. Comm.* 75 (1979).
Chernoff, Hochstrasser, and Steele, *Proc. Natl. Acad. Sci. USA 77*, 5606 (1980)
Henry, Sommer, Hofrichter, and Eaton, *J. Mol. Biol. 166*, 443 (1983)

Effect of viscosity on protein kinetics
Beece, Eisenstein, Frauenfelder, Good, Marden, Reinisch, Reynolds, Sorensen, and Yue, *Biochemistry 19*, 5147 (1980).
Ansari, Jones, Henry, Hofrichter, and Eaton, *Science 256*, 1796 (1992).
Zwanzig, *J. Chem. Phys. 97*, 3587 (1992).

Role of protein relaxation in kinetics of geminate rebinding
Agmon and Hopfield, *J. Chem. Phys. 79*, 2042 (1983).
Srajer, Reinisch, and Champion, *J. Am. Chem. Soc. 110*, 6656 (1988).
Steinbach, Ansari, Berendzen, Braunstein, Chu, Cowen, Ehrenstein, Frauenfelder, Johnson, Lamb, Luck, Mourant, Nienhaus, Ormos, Philipp, Xie, and Young, *Biochemistry 30*, 3988 (1991)
Ansari, Jones, Henry, Hofrichter, and Eaton, *Biochemistry 33*, 5128 (1994).
Jackson, Lim, and Anfinrud, *Chem. Phys. 180*, 131 (1994).

II. "*Cooperativity in hemoglobin - the hydrogen molecule of biology*"

Structure, stereochemical mechanism, and allostery
Perutz, *Nature 228*, 726 (1970).
Szabo and Karplus, *J. Mol. Biol. 72*, 163 (1972).
Perutz, Fermi, Luisi, Shaanan, and Liddington, *Acc. Chem. Res. 20*, 309 (1987).
Lee, Karplus, Poyart, and Bursaux, *Biochemistry 27*, 1285 (1988).
Rivetti, Mozzarelli, Rossi, Henry, and Eaton, *Biochemistry 32*, 2888 (1993).

Photoselection in polarized photolysis experiments
Ansari and Szabo, *Biophys. J. 64*, 838 (1992).
Ansari, Jones, Henry, Hofrichter, and Eaton, *Biophys. J. 64*, 852 (1993).

Application of linear free energy relations to protein conformational changes and determination of a reaction path
Eaton, Henry, and Hofrichter, *Proc. Natl. Acad. Sci. USA 88*, 4472 (1991).
Janin and Wodak, *Biopolymers 24*, 509 (1985).

Kinetics of geminate recombination, the quaternary conformational change, and the allosteric mechanism
Sawicki and Gibson, *J. Biol. Chem. 251*, 1533 (1976).
Hofrichter, Sommer, Henry, and Eaton, *Proc. Natl. Acad. Sci. USA 80*, 2235 (1983).
Murray, Hofrichter, Henry, Ikeda-Saito, Kitagishi, Yonetani, and Eaton, *Proc. Natl. Acad. Sci. USA 85*, 2151 (1988).
Jones, Ansari, Henry, Christoph, Hofrichter, and Eaton, *Biochemistry 31*, 6692 (1992).

III. "*Nucleation of hemoglobin S polymers and sickle cell disease*"

Reviews
Eaton and Hofrichter, *Blood 70*, 1245 (1987).
Eaton and Hofrichter, *Ad. Prot. Chem. 40*, 63-279 (1990)

IV. "*Early events in the folding of cytochrome c*"

The experiment

Jones, Henry, Hu, Chan, Luck, Bhuyan, Roder, Hofrichter, and Eaton, *Proc. Natl. Acad. Sci. USA 90*, 11860 (1993).

Recent relevant reviews on protein folding

Ptitsyn in *Protein Folding*, ed. Creighton (Freeman, NY), p. 243 (1992).

Karplus and Shakhnovich, *ibid.*, p. 127. (1992).

Bryngelson, Onuchic, Socci, and Wolynes, *Proteins* (in press).