

## RESEARCH ARTICLES

## Structural Energetics of the Molten Globule State

Donald T. Haynie and Ernesto Freire

*Departments of Biology and Biophysics and the Biocalorimetry Center, The Johns Hopkins University, Baltimore, Maryland 21218*

**ABSTRACT** Certain partly ordered protein conformations, commonly called “molten globule states,” are widely believed to represent protein folding intermediates. Recent structural studies of molten globule states of different proteins have revealed features which appear to be general in scope. The emerging consensus is that these partly ordered forms exhibit a high content of secondary structure, considerable compactness, nonspecific tertiary structure, and significant structural flexibility. These characteristics may be used to define a general state of protein folding called “the molten globule state,” which is structurally and

thermodynamically distinct from both the native state and the denatured state. Despite extensive knowledge of structural features of a few molten globule states, a cogent thermodynamic argument for their stability has not yet been advanced. The prevailing opinion of the last decade was that there is little or no enthalpy difference or heat capacity difference between the molten globule state and the unfolded state. This view, however, appears to be at variance with the existing database of protein structural energetics and with recent estimates of the energetics of denaturation of  $\alpha$ -lactalbumin, cytochrome *c*, apomyoglobin, and T4

Received September 17, 1992; revision accepted December 1, 1992.

Address reprint requests to Dr. Ernesto Freire, Department of Biology and Biocalorimetry Center, Johns Hopkins University, Charles and 34th Streets, Baltimore, MD 21218.

Abbreviations used: *a*, ligand activity;  $a_{\text{H}}$ , hydrogen ion activity; apo- $\alpha$ -LA, apo form of  $\alpha$ -lactalbumin; apo-Mb, apo form of myoglobin; A, the acid pH-induced intermediate state of  $\alpha$ -lactalbumin;  $\Delta A_{\text{ap}}$ , change in apolar surface area accessible to solvent;  $\Delta A_{\text{pol}}$ , change in polar surface area accessible to solvent; CD, circular dichroism;  $\Delta C_p$ , heat capacity change at constant pressure;  $\Delta C_{p,\text{ap}}$ , constant pressure heat capacity change per  $\text{\AA}^2$  of apolar surface;  $\Delta C_{p,\text{D}}$ , heat capacity change at constant pressure for total unfolding;  $\Delta C_{p,\text{D,calc}}$ , calculated heat capacity change at constant pressure for total unfolding;  $\Delta C_{p,\text{D,exp}}$ , measured heat capacity change at constant pressure for total unfolding;  $\Delta C_{p,\text{A} \rightarrow \text{D}}$ , heat capacity change at constant pressure for unfolding from the acid-pH stabilized; state of  $\alpha$ -lactalbumin to the denatured state;  $\Delta C_{p,\text{J}}$ , heat capacity change at constant pressure for the native state to intermediate state; transition of the acid pH-stabilized state of apomyoglobin;  $\Delta C_{p,\text{J} \rightarrow \text{D}}$ , heat capacity change at constant pressure for unfolding from the acid pH-stabilized; state of apomyoglobin to the denatured state;  $\Delta C_{p,\text{J}}$ , heat capacity change at constant pressure for the native state to intermediate state transition;  $\Delta C_{p,\text{P}}$ , heat capacity change at constant pressure for unfolding from the native state to the extended intermediate;  $\Delta C_{p,\text{P,calc}}$ , calculated heat capacity change at constant pressure for unfolding from the native; state to the extended intermediate;  $\Delta C_{p,\text{P} \rightarrow \text{D,calc}}$ , calculated heat capacity change at constant pressure for unfolding from the extended; intermediate to the denatured state;  $\Delta C_{p,\text{pol}}$ , constant pressure heat capacity change per  $\text{\AA}^2$  of polar surface; C-terminal, carboxyl-terminus of polypeptide; cyt *c*, cytochrome *c*; D, denatured (unfolded) state;  $\Delta G$ , Gibbs free energy change;  $\Delta G_{\text{J}}$ , Gibbs free energy change for the native state to intermediate state transition;  $\Delta G_{\text{D}}$ , Gibbs free energy change for the native state to denatured state transition; GuHCl, guanidine hydrochloride;  $\Delta H$ , enthalpy change;  $\Delta H_{\text{b}}$ , enthalpy of ligand binding;  $\Delta H_{\text{cal}}$ , calorimetric enthalpy change;  $\Delta H_{\text{D,calc}}$ , calculated enthalpy change for transition from the native state to the denatured state;  $\Delta H_{\text{D,exp}}$ , measured

enthalpy change for transition from the native state to the denatured state;  $\Delta H_{\text{D}}$ , enthalpy change for native state to unfolded state transition;  $\Delta H_{\text{I} \rightarrow \text{D}}$ , enthalpy change for a transition from the acid pH-induced intermediate state of apomyoglobin to the denatured state;  $\Delta H_{\text{J}}$ , enthalpy change for the native state to intermediate state transition;  $\Delta H_{\text{P} \rightarrow \text{D}}$ , enthalpy change for a transition from the extended intermediate to the unfolded state;  $\Delta H_{\text{vH}}$ , van't Hoff enthalpy change; I, the acid pH-induced intermediate state of apomyoglobin;  $k_{\text{b}}$ , microscopic ligand binding constant;  $K_{\text{D},i}$ , protonation constant of the *i*th ionizable group in the unfolded state;  $K_{\text{J},i}$ , protonation constant of the *i*th ionizable group in the intermediate state J of a three state equilibrium;  $K_{\text{N},i}$ , protonation constant of the *i*th ionizable group in the native state;  $\alpha$ -LA,  $\alpha$ -lactalbumin; Mb, myoglobin;  $\Delta n_{\text{J}}$ , increase in number of ligand binding sites upon a transition from the native state to the intermediate state;  $\Delta n_{\text{D}}$ , increase in number of ligand binding sites upon a transition from the native state to the unfolded state; N, native state; NMR, nuclear magnetic resonance;  $N_{\text{res,helix}}$ , number of residues in a helical conformation in the native state; N-terminal, amino-terminus of polypeptide;  $\omega_i$ , degeneracy of state *i*; P, the acid pH-induced extended intermediate state of cytochrome *c*;  $P_{\text{D}}$ , population of the denatured state;  $P_{\text{J}}$ , population of the intermediate state J of a three state equilibrium;  $pK_{\text{a}}$ , pH value at which half of the residues of a given species are protonated;  $pK_{\text{D}}$ , negative log of protonation constant of a given ionizable group in the unfolded state;  $pK_{\text{A}}$ , negative log of protonation constant of a given ionizable group in state A of  $\alpha$ -LA;  $pK_{\text{N}}$ , negative log of protonation constant of a given ionizable group in the native state; *Q*, partition function; *R*, molar gas constant;  $\Delta S$ , entropy change;  $\Delta S_{\text{D}}$ , entropy change for the native state to denatured state transition;  $\Delta S_{\text{J}}$ , entropy change for the native state to intermediate state transition; *T*, absolute temperature;  $T_{\text{m}}$ , melting temperature;  $T_{\text{m,D}}$ , melting temperature for native state to unfolded state transition ( $\Delta H_{\text{D}}/\Delta S_{\text{D}}$ );  $T_{\text{m,J}}$ , melting temperature for native to intermediate state transition ( $\Delta H_{\text{J}}/\Delta S_{\text{J}}$ );  $T_{\text{R}}$ , reference temperature; T4 lys, T4 lysozyme; UV, ultraviolet; 2D, two-dimensional; Ib, an acid pH-induced compact intermediate state of cytochrome *c*; Ic, an acid pH-induced compact intermediate state of cytochrome *c*.

lysozyme. We discuss these four proteins at length. The results of structural studies, together with the existing thermodynamic values for fundamental interactions in proteins, provide the foundation for a structural thermodynamic framework which can account for the observed behavior of molten globule states. Within this framework, we analyze the physical basis for both the high stability of several molten globule states and the low probability of other potential folding intermediates. Additionally, we consider, in terms of reduced enthalpy changes and disrupted cooperative interactions, the thermodynamic basis for the apparent absence of a thermally induced, cooperative unfolding transition for some molten globule states. © 1993 Wiley-Liss, Inc.

**Key words:** molten globule state, protein folding intermediates, secondary structure, cytochrome *c*

## INTRODUCTION

A major goal in biological research is to understand how the specific amino acid sequence of a typical small globular protein encodes all the information required for it to fold spontaneously into the unique, three-dimensional native conformation.<sup>1,2</sup> Although innumerable conformations are potentially accessible to a given polypeptide, structural transitions between a disordered coil and the native structure cannot involve a random search of all possible conformations.<sup>2</sup> Indeed, kinetic refolding experiments indicate that submolecular structural units can be obtained from the disordered state within a second or less, a time interval *many* orders of magnitude shorter than the one required for a random search of the entirety of conformation space.<sup>3</sup> During folding the disordered polypeptide chain collapses, the elements of secondary structure form, and the specific noncovalent tertiary structural cross-links interlock. The study of partly folded structures existing along the folding pathway is of fundamental importance for comprehending the mechanisms by which a polypeptide attains the native conformation.<sup>4-9</sup>

The protein folding process is, however, highly cooperative. From a statistical mechanical point of view, this means that under "normal" equilibrium conditions the vast majority of partly folded states are energetically so unfavorable that their respective statistical weights are effectively negligible.<sup>10</sup> Consequently, their aggregate population is essentially zero. Experimental results indicate that at equilibrium partly folded intermediates usually constitute no more than 5% of the total population of most globular proteins.<sup>11</sup> Under some experimental conditions, however, the equilibrium population of

intermediate states can be substantial, making possible investigations of their structural and energetic features.<sup>13-22,73,94</sup> Studies of such states have shown that most intermediates share some common characteristics and collectively define a distinct protein state, often called the molten globule state. Many investigators believe that the molten globule state may represent a "universal" intermediate in the folding pathway; as a result, the topic has attracted significant attention in recent years.<sup>27-37</sup> We begin this review by summarizing some of the principal features of the native state, the denatured state, and the molten globule state.

Crystal structures of water soluble proteins indicate that a considerable fraction of the folded polypeptide chain follows a regular course, forming the regions of higher structural order known as  $\alpha$ -helices,  $\beta$ -sheets, and reverse turns (see ref. 12). Specific noncovalent polar and apolar contacts mediate intramolecular interactions between these structural units. Along with the effect of solvent exclusion, these contacts more than compensate the negative entropy change associated with organizing the polypeptide into the native conformation. Each residue occupies a unique environment. The folded structure is compact and rigid; therefore, conformational fluctuations are small. Disulfide bridges, when present, further reduce the amplitude of intramolecular fluctuations. A nonpolar core excludes water almost completely, and buried water molecules are typically hydrogen bonded to the protein. In some cases, both the complete folding and thermal stability of the native state depend on the presence of specific ligands, often prosthetic groups or metal ions. Because few apolar residues are exposed to solvent, the specific heat capacity of the native state is lower than that of the unfolded polypeptide chain.<sup>11</sup>

In the flexible unfolded form, on the other hand, most, if not all, of the noncovalent intramolecular interactions of the native state are absent. Effectively all residues are exposed to solvent.<sup>70</sup> Structurally, the disordered polypeptide chain can and does undergo large conformational fluctuations, but the state may be approximated by a "random" coil only to the extent that excluded volume effects and intramolecular noncovalent linkages such as disulfide bridges are duly accounted for.

In 1981 Ptitsyn's group introduced a general model of partly organized protein structure, the main features of which were based on experimental studies of the acid form of  $\alpha$ -lactalbumin: "a compact globule with native-like secondary structure and with slowly fluctuating tertiary structure."<sup>14</sup> Shortly thereafter, reporting on the predominant state of cytochrome *c* at low pH, high chloride concentration, and room temperature, Ohgushi and Wada coined the term "molten globule" to describe this "new type of structural state of the globular protein."<sup>16</sup> The label refers to two of the state's prin-

cipal characteristics: "molten" to its nonspecific, "liquid-like" tertiary structure interactions; "globule" to its native-like compactness.

Many structural studies of different proteins have shown that these partly folded states are not unique to  $\alpha$ -lactalbumin or cytochrome *c* (see refs. 26 and 99 for recent reviews). In general, these intermediates display considerable structural flexibility and non-native tertiary structure; hence, most tertiary structure hydrogen bonds of the native state are broken. Yet they are highly compact and possess a significant percentage of the secondary structure content of the native state. The elements of secondary structure of these states are, however, less rigid than their native state counterparts and are stabilized by nonspecific intramolecular hydrophobic contacts.

Though discussions pertaining to intermediate states presently enjoy considerable popularity in the scientific literature, the detailed structure and especially the energetics of these states are far from being understood (see refs. 4, 6, 9, 23–26, and 99 for reviews). Characterizations of molten globule states are requisite for understanding the principles of structure stabilization and the mechanisms of protein folding, since they model the otherwise elusive, transiently populated kinetic intermediates on the folding pathway.<sup>4,24,25</sup> The structural energetics of these states constitute the focus of this review.

## STRUCTURE OF THE MOLTEN GLOBULE STATE

In this section we summarize the results of structural studies of several partly ordered states. The most studied molten globule state is the low pH form of  $\alpha$ -lactalbumin even though no high resolution structural information is yet available. Extensive structural studies of molten globule states of cytochrome *c* and apomyoglobin have also appeared in the literature. Therefore, we center our attention on these three globular proteins. We conclude with a discussion of the consensus structure of the molten globule state, which we define in accordance with the common structural features of the molten globule states described below.

### $\alpha$ -Lactalbumin

$\alpha$ -Lactalbumin ( $\alpha$ -LA) is produced in the lactating mammary gland, where it promotes synthesis of milk lactose by modifying the substrate specificity of galactosyltransferase (reviewed in ref. 42). It is one of the major "whey" proteins of bovine milk. At 123 residues in length,  $\alpha$ -LA is morphologically similar to *c*-type lysozymes, of which the hen egg white variety is best known. A comparison between the two gene structures suggests that a duplication event produced  $\alpha$ -LA. All four disulfide bridges are conserved, as are about one third of the residues. Both proteins have two nonsequential structural lobes,

the larger one being formed from about two-thirds of the linear chain, including the amino- and carboxyl-termini. The four principal helices are in the large subdomain; the other lobe is partly comprised by one  $\beta$ -sheet and an irregular loop region (see Fig. 1).

Despite the similarities between  $\alpha$ -LA and lysozyme, certain structural, chemical, and thermal properties are quite different. Only the former is able to bind calcium with high affinity<sup>43</sup> at a location near the lobe interface.<sup>44</sup> Removal of bound calcium dramatically attenuates thermal stability, some of which can be restored through specific binding of other cations.<sup>45–47</sup> The saccharide binding sites of lysozyme are blocked in  $\alpha$ -LA.<sup>42</sup> The two-state character of lysozyme thermal denaturation is firmly established under a broad range of experimental conditions; whereas for  $\alpha$ -LA, a molten globule state is stable at extremes of pH, intermediate guanidine hydrochloride (GuHCl) concentrations, or low ionic strengths.<sup>25</sup> Even though all three intermediate forms have been described as molten globule states, they might not be structurally (or energetically) equivalent. In accordance with the notation of Kuwajima, the native state, low pH molten globule state, and denatured state of  $\alpha$ -LA are here designated "N," "A," and "D," respectively.<sup>48</sup> At pH 2, apparently all molecules are in the acid pH molten globule state.<sup>48</sup> Unless noted otherwise, all results summarized below pertain to bovine  $\alpha$ -LA; in most cases the forms from different species have similar properties.<sup>24,25,49</sup>

Far-UV circular dichroism (CD) spectroscopy suggests that the secondary structure content of A is N-like. Early estimates of the helical content of A were about 10% greater than values for N, since the ellipticity of the former at 222 nm is markedly larger.<sup>50</sup> The shape of the spectrum, however, is considerably different.<sup>51</sup> Ptitsyn and coworkers found an N-like, overall  $\alpha$ -helix content of 30–40%. The quantity of  $\beta$ -sheet, too, was estimated to be about the same for both states. Using infrared spectroscopy, the same group measured a total content of  $\alpha$ -helix differing from that of N by no more than about 10%, a  $\beta$ -sheet content no different than about 5%.<sup>52</sup>

The secondary structural elements of A are significantly more stable than helical segments in isolated small peptides. Whereas small peptides generally undergo rapid interconversion with a large population of disordered conformations, A displays specific structural regions which fluctuate little at room temperature.<sup>49,53</sup> Tryptophan residues lie buried in inflexible regions.<sup>52,54</sup> In Dobson and co-worker's solvent exchange experiments on guinea pig  $\alpha$ -LA, some of the amide hydrogens most highly protected from exchange correspond to residues in an  $\alpha$ -helix of N.<sup>49</sup> This helix, which together with another small one forms the calcium binding site, is intact and rigid in A; however, all amide hydrogens

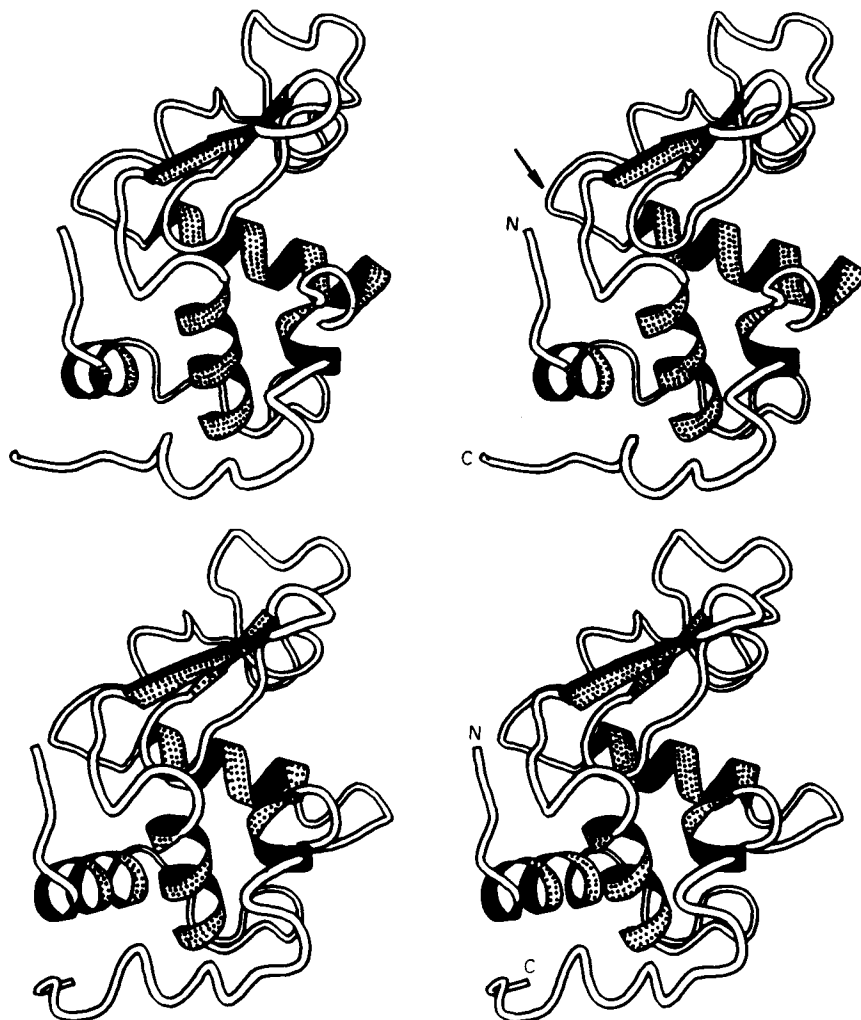


Fig. 1. Stereo ribbon drawings of **(top)** baboon  $\alpha$ -lactalbumin and **(bottom)** hen egg-white lysozyme. In top panel the arrow points to the calcium binding loop, which is located at the interface of two structural lobes. The small lobe, consisting of about one-third of all residues in the middle of the polypeptide chain, contains

a  $\beta$ -sheet, a disordered loop, and a small amount of helix. The polypeptide chain loops back on itself, so that both the N- and C-termini are found in the large lobe, which contains the hydrophobic box and is mainly helical. Adapted from Acharya, K.R., et al.<sup>44</sup>

are not uniformly protected in the partly unfolded form:  $\beta$ -sheet and disordered loop amide protons, for example, exchange rapidly in A.<sup>49</sup> These data, when compared to the results of similar experiments on homologous lysozyme, suggest that the lobe of  $\alpha$ -LA containing the  $\beta$ -sheet is disordered in A.<sup>55</sup>

Hydrodynamic studies indicate that A is highly compact. An early quantitative estimate of the molecular size of A corresponds to a radius about 20% larger than that of N.<sup>56</sup> Later, Ptitsyn's group found that the intrinsic viscosity of A exceeds that of N by no more than 30% and that the gyration radii of the two forms are effectively commensurate.<sup>14,52</sup> Separate measurements have shown that the sedimentation constants of the two states differ by less than 15%.<sup>51,52</sup>

The stable helix mentioned above is amphipathic

and borders on the hydrophobic box region of N.<sup>44</sup> Although dynamic averaging of side-chain interactions in this region is enhanced in A, the aromatic residues located there have some of the most significantly perturbed chemical shifts.<sup>49</sup> This region of the protein is, therefore, "a highly structured part of the molecule."<sup>49</sup> The stability of the helix probably depends on hydrophobic interactions with the aromatic residues of the hydrophobic box.<sup>49</sup> A is more like a collapsed conformation than an expanded one.

Tertiary structure packing in A resembles that in N, according to diffuse X-ray scattering data for the human protein.<sup>57</sup> These experiments show that in A interatomic distances are less than 5% greater and the volume of the protein interior is about 10% larger than in N. Many apolar residues are buried.

Therefore, a hydrophobic core is present; little water penetrates into the interior of A.<sup>57</sup>

Notwithstanding the N-like character of the helices and the persistent hydrophobic core, it is clear that the higher level order of A is distinct. There are large scale fluctuations in the tertiary structure of A, attributable more to greater freedom of rotation of side chains rather than to "melting out" of helical regions. This conclusion follows from several experimental results. One, the near-UV CD spectrum of A closely resembles that of D.<sup>51</sup> Two, the proton nuclear magnetic resonance (<sup>1</sup>H NMR) spectrum is similar to the typical spectra of free amino acids or of unfolded proteins, though some resonance lines remain.<sup>52</sup> The extent of averaging is not uniform, however, and some resonances show large deviations from random coil shifts.<sup>49</sup> Three, solvent exchange data show that some regions of A have slowly exchanging amide hydrogens and are stable, while other regions are completely unfolded (ref. 49 and personal comm. cited in ref. 58). And four, results of polarization of luminescence experiments indicate that the intramolecular mobility of tryptophan residues is, during a period of tens of nanoseconds, about as restricted in both forms.<sup>52</sup>

In GuHCl denaturation experiments, native tertiary structure is destroyed first, secondary structure second.<sup>48</sup> Ewbank and Creighton report that the stability of A is independent of a single disulfide bond pairing scheme.<sup>53</sup> Moreover, in Kuwajima's folding model of 1977, neither disulfide linkages nor electrostatic interactions stabilize A.<sup>13</sup> It is said, therefore, that hydrophobic interactions are necessary for stabilization of A.<sup>13</sup> That is, A appears to be stabilized by time-averaged, mostly nonspecific hydrophobic interresidue contacts between secondary structure elements and increased configurational entropy.

### Cytochrome *c*

Horse heart cytochrome *c* (cyt *c*) is an electron-transferring protein involved in oxidative phosphorylation in the inner membrane of mitochondria (see refs. 60 and 61 for reviews). Five helices, three major ( $\alpha$ ) and two minor ( $3_{10}$ ), are distributed along the chain, and no  $\beta$ -sheet is present<sup>60,61</sup> (see Fig. 2). The protein is small and globular. A single polypeptide chain, cyt *c* contains 104 amino acids, 24 of which are basic residues. There are no free sulfhydryl groups or disulfide bonds; a sole tryptophan residue lies in the native structure about 10 Å from a heme moiety. In the native holoprotein, the heme is coordinately bound to the polypeptide by two thioester linkages and two strong field ligands, a histidine and a methionine. Embedded in a crevice, the prosthetic group forms numerous noncovalent interactions with the polypeptide chain. These intramolecular contacts are essential for stabilization of the native globular protein conformation. Acidification

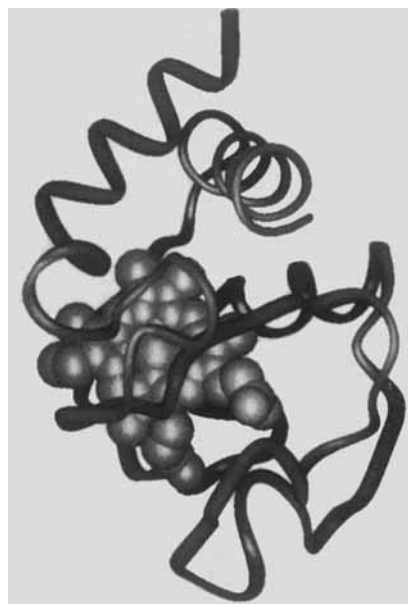


Fig. 2. Diagram of cytochrome *c* generated using Insight II (Biosym Technologies, San Diego, CA). Though a heme group separates the polypeptide chain into two distinct regions, the N- and C-terminal helices are spatially proximal and interact strongly.

of the protein solution results in displacement of the two strong field ligands and significantly reduces thermal stability. Under such conditions, the acidic side chains are neutralized, yielding a highly charged molecule; the histidine ligand is protonated. An increase in anion concentration decreases the electrostatic repulsion by specific binding, thereby enabling the protein to partly refold.<sup>62,63</sup> Various conformations of cyt *c* have been described. These states are maximally populated under the following conditions: "N" at neutral pH and low temperatures, "D" at high temperatures, at acid pH (~2) and high chloride concentrations: "IIb" at low temperatures and "IIc" at intermediate temperatures,<sup>64</sup> and "P" at acid pH and low chloride concentrations.

At acid pH and low chloride concentration, segments of the native structure are held apart by strong electrostatic repulsions.<sup>62,63,65</sup> Jeng and Englander have characterized this state (P) using a variety of methods.<sup>65</sup> Their fluorescence experiments show that under these conditions the indole ring of tryptophan is much farther from the heme than in N and is exposed to water. The intrinsic viscosity of P is about four times greater than that of N, reflecting a lack of interlocked atomic close packing interactions; the polypeptide is greatly extended.<sup>65</sup> Nonetheless, ellipticity measurements at 222 nm suggest that about 85% of the helical content of N is intact at 20°C.<sup>65</sup> In order to check whether the CD signals represent the native helices, the same group conducted hydrogen exchange experiments. They found that some of the secondary structural hydrogen

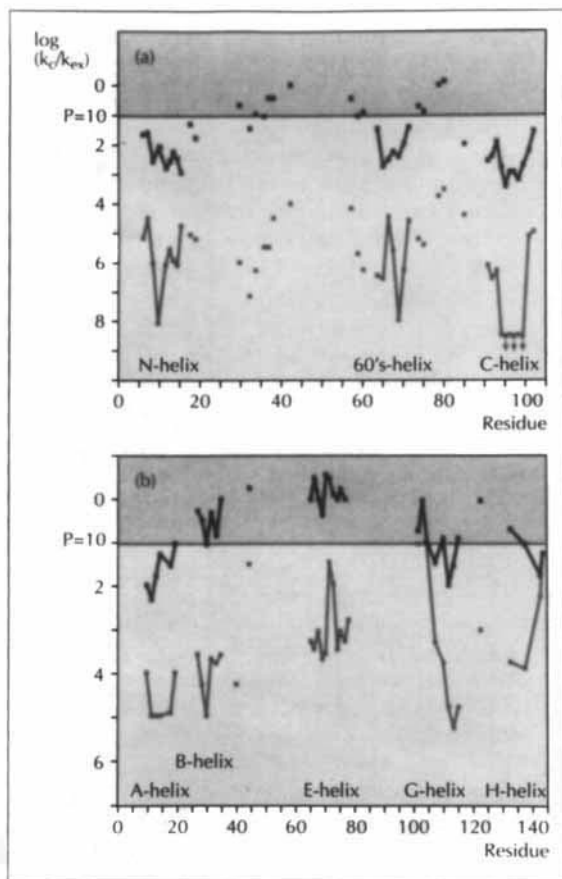


Fig. 3. Profiles of hydrogen exchange protection factors for (a) oxidized horse cyt *c* in the IIb (dark) and N (light) states, and (b) sperm whale apomyoglobin in the I (dark) and N (light) states. The logarithm of the protection factor for amide proton exchange is plotted against residue number. Lines connecting data points show which residues form helices in the native structures. Arrows designate very slowly exchanging amide protons. Reproduced from Roder and Baldwin<sup>7</sup> with permission of Current Science.

bonding is "moderately stable," the pattern of protection of N being preserved in P<sup>65</sup> (see Fig. 3). However, these data, collected at one temperature, do not imply that N is never significantly populated: CD studies by Wada and colleagues show that N can be stable in the same buffer system and that its maximum population is about 40% near 0°C.<sup>64</sup>

The N-terminal segment of cyt *c* is helical in N. When isolated, however, it is a disordered coil, according to Baldwin and co-workers (personal communication cited in ref. 65). As for the C-terminal segment, it is about 25% helical at 0°C.<sup>65</sup> Therefore, the stability of P is said to derive from intrinsic helix propensities of certain sequences and intersegment "soft-packing" interactions, probably hydrophobic in nature.<sup>65</sup> This state, called a "premolten globule state," does not undergo cooperative unfolding.<sup>65</sup>

At low pH but high chloride concentration, on the other hand, a partly folded, compact globular con-

formation (IIb) is predominantly stable.<sup>16</sup> Intrinsic viscosity data show that, compared with N, IIb is only slightly expanded.<sup>16</sup> Even though IIb is compact, the main-chain and side-chain protons show motional averaging, suggesting that the conformation is fluid-like.<sup>16</sup> The aromatic region of the <sup>1</sup>H NMR spectrum is similar to that of D.<sup>16</sup> Addition of chloride does not fully restore the tertiary structure interactions of N, as demonstrated by differences in aromatic CD signals.<sup>16</sup> Still, some higher order contacts are formed; there is a hydrophobic core, and IIb does unfold cooperatively.<sup>64,100</sup>

Hydrogen exchange experiments show that the slowly exchanging amide protons in the three major helices of N are strongly protected from exchange at 20°C, acid pH, and high salt.<sup>66</sup> These protons, however, are much less protected than in N.<sup>66</sup> Clusters of protected helices in IIb correspond to helices in N; the boundaries of the helices are nearly the same in both states.<sup>66</sup> In contrast, many amide protons involved in tertiary structure hydrogen bonds in N are only marginally protected in IIb, suggesting a drastic destabilization of those tertiary interactions.<sup>66</sup> These results clearly indicate that IIb contains N-like elements of helical secondary structure and that nonhelix hydrogen bonds are either unstable or absent.<sup>66</sup> Therefore, some regions of the tertiary structure must be highly flexible, and the increases in structural flexibility are apparently local rather than global.<sup>66</sup>

There is some evidence for a second molten globule state of cyt *c*.<sup>64</sup> At acid pH and high chloride concentrations, the form stable at low temperatures is the partly folded state IIb, described above. At high temperatures, D is most stable.<sup>64</sup> Since thermal denaturation data collected under these solution conditions cannot be fit by a two-state model, a second intermediate state must become significantly populated.<sup>64</sup> This state, called "IIc," has been observed only under these conditions and is maximally populated at about 40°C.<sup>64</sup> Experiments by Wada's group on the quenching of tryptophan fluorescence by heme suggest that IIc also is a compact intermediate.<sup>64</sup>

### Apomyoglobin

Myoglobin (Mb), an oxygen-carrying protein in vertebrates, is located in muscle, where it facilitates the movement of molecular oxygen vital to cellular respiration.<sup>67</sup> Skeletal muscle Mb from sperm whale was the first protein to be seen at atomic resolution.<sup>110</sup> The native structure is extremely compact and stabilized by a liganded heme. There are eight  $\alpha$ -helices, designated "A" through "H," constituting about 75% of the main chain.

A globular state is stable in the absence of heme.<sup>68</sup> Griko et al., have shown that when the pH is near neutral, apo-Mb is compact (sedimentation); also, it has unique spatial structure (1-D NMR) and

an extensive hydrophobic core (differential scanning calorimetry).<sup>68</sup> Upon removal of heme at pH > 4.8 and 30°C, the intrinsic viscosity increases by about 20%, and the overall helicity decreases by 20–25%.<sup>68</sup> This state Baldwin and co-workers call “N.”<sup>69</sup>

pH titration of apo-Mb reveals a two stage unfolding process.<sup>68,69</sup> The pH value (pH 4.8–4.5) at which the first stage of unfolding is complete depends on temperature, whereas the second stage is nearly temperature independent.<sup>69</sup> The percentage increase in viscosity is smaller for the high pH transition than for the low pH transition, suggesting that the intermediate state has a molecular volume similar to that of N.<sup>68</sup> On the basis of amide proton exchange and CD spectroscopy data, it is known that little of the helical structure of N is stable when the pH is extremely low.<sup>58</sup> The physical state following the first stage of acid unfolding is called “I.”<sup>58</sup> Only at very low pH or at high temperatures does the denatured state (D) become significantly populated.

At pH 4.2 and room temperature, I is the predominant state.<sup>69</sup> Thermograms from differential scanning calorimetry experiments under conditions which stabilize I show no heat absorption peak.<sup>68</sup> Also, the residues of I are considerably less solvated than those of acid denatured metMb, which is taken to be completely unfolded.<sup>68,70</sup> Therefore, some of the hydrophobic side chains must be largely exposed to water, even though a reduced hydrophobic core must still exist.<sup>68</sup> Though the near-UV CD spectrum of I resembles that of D, indicating the disruption of native tertiary structure, the far-UV signal is large, as in N.<sup>58,68</sup> The  $\alpha$ -helix content is about 35%.<sup>58</sup> Helices A, G, and H protect many of their amide protons from solvent exchange<sup>58</sup> (see Fig. 3). In I, the pattern of amide protection is preserved in about 40 sites along the chain; however, the protection factors are up to four orders of magnitude lower than those of N, indicating that I is globally significantly less rigid than N.<sup>58</sup>

To assess the role of hydrophobic interactions in stabilizing I, Baldwin's group introduced mutations in the A–H and G–H helix packing sites.<sup>69</sup> They found that the mutations tended to destabilize N relative to D and that the N to I transition involves disruption of tight packing interactions in the vicinity of the mutations.<sup>69</sup> Mutations which increased the nonpolar surface area in the A–H and G–H helix packing sites yielded small increases in the stability of I; those which decreased the nonpolar surface area decreased the stability of I.<sup>69</sup>

Using the above data, the same group built a structural model of I,<sup>58,69</sup> a schematic of which is shown in Figure 4. In this model, a compact subdomain of the polypeptide has N-like structure, while the remainder of the protein is essentially unfolded.<sup>58,69</sup> Helices B through E, which constitute the other compact subdomain in N, are unfolded.



Fig. 4. Schematic drawing of sperm whale myoglobin.  $\alpha$ -helices A, F, G, and H form a subdomain of the tertiary structure; the other four helices constitute a second, smaller subdomain. In the model of state I proposed by Baldwin and co-workers, helices B–E (no shading) are unfolded; the F helix (light shading) is disordered; and helices A, G, and H (dark shading) are native-like. The overall fold is such that the N- and C-termini form contacts in the large subdomain. Reproduced from Hughson et al.,<sup>58</sup> with the permission of AAAS. Copyright 1990 by the AAAS.

Though I lacks well-defined tertiary structure and N-like helix–helix interactions, it does have “loose,” nonspecific hydrophobic interactions.<sup>58,69</sup> These nonnative interactions stabilize the intact helices. The intact hydrophobic core is also believed to stabilize the compact structure. The tertiary interactions of I probably do not completely exclude water.<sup>58,69</sup>

### Consensus Structure of the Molten Globule State

The foregoing survey of specific observations of partly ordered states of three different globular proteins clearly shows that some of the described states exhibit similar features. These shared characteristics can be used to provide a minimal definition of the protein conformation commonly known as the molten globule state. This state has three predominant features: (1) a relatively high content of secondary structure, (2) secondary structural elements which do not establish tertiary structure contacts satisfying the rigid geometric constraints of the native state, and (3) a sizable hydrophobic core. The hydrophobic core is believed to be responsible for maintaining compactness, and the overall structure exhibits a high degree of flexibility and undergoes relatively large fluctuations.

Thus, the conformations A ( $\alpha$ -LA), IIB and IIC (cyt c), and I (Mb) all may be classified as molten globule states. State P of cyt c, on the other hand, is not a molten globule state: for even though it has intact elements of secondary structure, it is not compact; it has no hydrophobic core. Therefore, P is better called a “pre-molten globule state,” as proposed by Jeng and Englander.<sup>65</sup> The structural features of

the molten globule state appear to be common not only to the molten globule states considered above, but also to those of other proteins described in the literature (see refs. 4, 6, 9, 24–26, and 99 for reviews).

Over the past few years the term “molten globule” has been employed in ever broader contexts, giving rise to a considerable degree of confusion. Experimentally detected intermediates, models of protein structure, and a folding model have all been labeled “molten globule.” In order that the term not become meaningless, we suggest that only those states that satisfy the structural criteria described above be classified as molten globule states.

In certain cases, a partly folded state might be described more comprehensively as a collection of isoenergetic conformations in equilibrium. In other words, the state would be degenerate. In this situation, all conformers meeting the specific structural requirements of the definition would constitute a molten globule state. Also, according to the prescribed usage, one may say that a kinetic intermediate or transition state of a given protein is a molten globule state, if the intermediate satisfies the definition of the state. Therefore, it is possible for a particular protein to have more than one molten globule state (e.g., states IIb and IIc of cytochrome c). In such a situation the different molten globule states would not simply be different isoenergetic conformations.

### EXPECTED ENERGETICS OF THE MOLTEN GLOBULE STATE

Thermodynamics provides a quantitative description of the forces that drive protein folding and stabilize the native structure under equilibrium conditions. An important goal of thermodynamic studies is to determine the relative molar Gibbs free energy—and its dependence on temperature, pH, ionic strength, etc.—of all conformations pertinent to the folding process. Such information permits calculation of the populations of accessible states under various conditions. The difference between the molar Gibbs free energy function of any state and a reference state (typically the native state),  $\Delta G(T)$ , can be computed from the state's relative enthalpy ( $\Delta H$ ), entropy [ $\Delta S(T_m) = \Delta H(T_m)/T_m$ , for a two-state transition, where  $T_m$  is the melting temperature], and heat capacity at constant pressure ( $\Delta C_p = \partial \Delta H / \partial T$ ), according to the standard relationship:

$$\Delta G(T) = \Delta H(T) - T \cdot \Delta S(T) \quad (1a)$$

$$= \Delta H(T_R) + \Delta C_p \cdot (T - T_R) - T \cdot [\Delta S(T_R) + \Delta C_p \cdot \ln(T/T_R)] \quad (1b)$$

where  $T_R$  is an appropriate reference temperature. Throughout the discussion  $\Delta C_p$  will be assumed to be temperature independent. Since the early days of physical biochemistry it has been assumed that

there is a direct correspondence between the structure of a protein and its energetics. Only recently, however, has this relation been brought to a quantitative and predictive level.<sup>72–75</sup> It is our object to relate changes in structural transitions involving partly folded states to changes in the thermodynamic functions just mentioned.

Because the molten globule state has a relatively high content of secondary structure, its unfolding must be accompanied by a significant enthalpy contribution, owing to the rupture of hydrogen bonds and the hydration of groups initially shielded from the solvent. Similarly, the conformational entropy of the molten globule state must be higher than that of the native state, since the overall structure exhibits a relatively high degree of flexibility; but the regularity of some of its structure means that its entropy must be considerably lower than that of the denatured state. As to heat capacity, if a particular conformation exhibits a hydrophobic core, some apolar amino acid side chains must be inaccessible to solvent. The heat capacity of the denatured state must be higher than that of the native state, and the molten globule state should reasonably be expected to have a lower heat capacity than the denatured state. As we discuss below, the heat capacity difference between a given state and the reference state is central to a thorough treatment of protein thermodynamics, since it determines the temperature dependence of both the enthalpy and the entropy. In fact, it is the increased heat capacity of the denatured state which gives rise to the phenomenon of “cold denaturation.”<sup>109</sup>

We begin this section by reviewing the concepts of heat capacity change, enthalpy change, and entropy change in relation to protein energetics studies. Then we apply these ideas to the intermediate structures of  $\alpha$ -LA, cyt c, and apo-Mb described above and to one of T4 lysozyme (T4 lys). After discussing the cooperativity of folding/unfolding transitions in the context of thermal denaturation, we conclude with a description of the expected energetics of the molten globule state.

### Heat Capacity Changes

#### Heat capacity changes of total unfolding

It is now evident that heat capacity changes<sup>72–75,106,107</sup> and enthalpy changes<sup>72–75</sup> between the native and denatured states of a protein can be predicted accurately using data obtained from crystallographic structures. One of the main conclusions of those studies is that the heat capacity difference between two different protein conformations is directly proportional to the change in apolar and polar surface areas accessible to the solvent:

$$\Delta C_p = \Delta C_{p,ap}^{\circ} \cdot \Delta A_{ap} + \Delta C_{p,pol}^{\circ} \cdot \Delta A_{pol} \quad (2)$$

where  $\Delta C_{p,ap}^{\circ}$  ( $0.45 \pm 0.02 \text{ cal K}^{-1} [\text{mol} \cdot \text{\AA}^2]^{-1}$ ) and  $\Delta C_{p,pol}^{\circ}$  ( $-0.26 \pm 0.03 \text{ cal K}^{-1} [\text{mol} \cdot \text{\AA}^2]^{-1}$ ) are the



TABLE I. Heat Capacity Changes for Complete Unfolding\*

Protein	$\Delta A_{\text{ap,D}}$ ( $\text{\AA}^2$ )	$\Delta A_{\text{pol,D}}$ ( $\text{\AA}^2$ )	$\Delta C_{\text{p,calc}}$ ( $\text{kcal K}^{-1} \text{mol}^{-1}$ )	$\Delta C_{\text{p,exp}}$ ( $\text{kcal K}^{-1} \text{mol}^{-1}$ )	Ref.
$\alpha$ -Lactalbumin	6780	4750	$1.8 \pm 0.3$	$1.5 \pm 0.3$	84
				$1.821 \pm 0.001$	85
Cytochrome <i>c</i>	5420	3850	$1.4 \pm 0.2$	$1.6 \pm 0.1$	92
Myoglobin	8850	5860	$2.5 \pm 0.4$	$2.6 \pm 0.1$	92
T4 lysozyme	9700	6580	$2.7 \pm 0.4$	$2.18 \pm 0.01$	93

\*Values were calculated as previously described in ref. 75. Native state accessible surface areas were calculated from the crystal structures (Protein Data Bank atomic coordinate identity codes 1ALC [ $\alpha$ -LA], 5CYT [cyt *c*], 4MBN [Mb], and 3LZM [T4 lys]) by Dr. Scott Presnell's (University of California, San Francisco, CA) implementation of the Lee and Richards algorithm. The probe radius was 1.4  $\text{\AA}$ , the slice width 0.25  $\text{\AA}$ . Accounting for additional exposure at the polypeptide chain ends, we modeled each unfolded state accessible surface area as the sum of the accessible surface areas of each residue in an extended Ala-X-Ala tripeptide. The tripeptides were generated using the program Quanta (Polygen, Waltham, MA).

elementary apolar and polar contributions to the total heat capacity increment; and  $\Delta A_{\text{ap}}$  and  $\Delta A_{\text{pol}}$  are the differences between the solvent accessible apolar and polar surface areas of the two states.<sup>74</sup> These parameter values have been shown to account not only for the heat capacity increment associated with the dissolution of solid dipeptides but also for the positive heat capacity change upon complete globular protein unfolding. Here "complete" means that all residues of the polypeptide are exposed to the solvent in the denatured form.<sup>74,75</sup>

Table I shows the increases in solvent accessible surface area and the predicted ( $\Delta C_{\text{p,D,calc}}$ ) and experimental ( $\Delta C_{\text{p,D,exp}}$ ) heat capacity changes for complete unfolding of  $\alpha$ -LA, cyt *c*, Mb and T4 lys. It should be noted that the experimental values, which were obtained from thermal denaturation measurements, are in good agreement with the calculated values. This suggests strongly that heat denaturation of these proteins involves complete or near complete (>85%) exposure to solvent of all side chains.

The validity of heat capacity measurements for assessing the completeness of unfolding was evaluated recently by Privalov and co-workers, who used microcalorimetry, in conjunction with other methods, to probe the character of the heat denatured state of several proteins.<sup>70</sup> They concluded that the heat capacity of a protein is a sensitive measure of the degree of unfolding, since this parameter is directly proportional to the extent of solvation of the polypeptide chain. From a thermodynamic point of view, this means that the heat denatured state of a typical globular protein is effectively devoid of secondary structure, if its heat capacity corresponds to that of a fully solvated peptide and the enthalpy change is that expected for complete unfolding. This is an important issue because several thermally denatured proteins display a large signal in the far-UV region of the CD spectrum,<sup>14,25,70,76,77</sup> which some investigators have interpreted as residual secondary structure.<sup>14,76</sup>

Ribonuclease A (RNase A) is a case in point. The heat denatured state of this protein shows a significant CD spectrum in the far-UV region. This spectrum can be reduced to a low level by adding GuHCl, which suggests that substantial secondary structure persists after the cooperative thermal transition.<sup>78</sup> However, using quenched-flow methods coupled with 2D  $^1\text{H}$  NMR spectroscopy, Robertson and Baldwin have found that this form of RNase A lacks secondary structure, as judged by the low protection against exchange of 36 amide protons.<sup>77</sup> The NMR data agree with calorimetric results, which indicate that the heat capacity of the thermally denatured form of RNase A corresponds to that of a fully solvated peptide.<sup>70</sup> Hence, independent experiments suggest that the thermally denatured form contains little or no stable hydrogen bonded structure.

Now, although CD spectra of native states of soluble globular proteins are readily interpretable, those of heat denatured forms appear to present a different situation. Since a large signal in the peptide region does not necessarily correspond to the presence of secondary structure, considerable care must be exercised in interpreting CD data. Within the limitations of the techniques, heat capacity measurements and amide hydrogen exchange rate determinations together provide a means of circumventing some of the ambiguities associated with estimating the secondary structure content of partly folded or unfolded states by optical methods.

#### Heat capacity changes of extended intermediates

The first type of intermediate we consider has an overall extended conformation (Fig. 5). The secondary structural elements of the native state are intact or largely intact while tertiary structural contacts are mostly or completely disrupted. This model represents states like the P state of cytochrome *c* described above and the extended intermediate forms of ribonuclease A stable at sub-zero temperatures.<sup>79,80</sup> These states might also represent the proposed con-

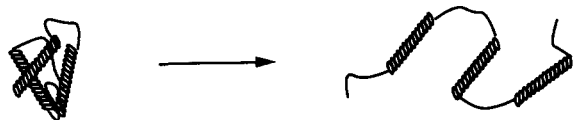


Fig. 5. Schematic illustration of a hypothetical N to P state transition. Although elements of native secondary structure are intact in P, this state is not a molten globule state, since it does not meet the structural criteria of the definition of molten globule state offered in this work: P is not compact, and it contains no hydrophobic core.

formations associated with the translocation of proteins across membranes. According to the definition given above, such states cannot be classified as molten globule states. Nonetheless, we briefly discuss extended intermediates to contrast their energetic properties with those of molten globule states.

The heat capacity difference between the extended intermediate and the native state ( $\Delta C_{p,P}$ ) can be calculated from the heat capacity change of  $\alpha$ -helix unfolding, which for an isolated peptide has been estimated at  $-7.6 \pm 1 \text{ cal K}^{-1} \text{ mol-res}^{-1}$ .<sup>72</sup> This value is the sum of the unfolding heat capacity changes attributable to increased  $\alpha$ -carbon solvent exposure ( $6.7 \pm 0.2 \text{ cal K}^{-1} \text{ mol}^{-1}$ ) and to hydrogen bond disruption ( $-14.3 \pm 1.4 \text{ cal K}^{-1} \text{ mol}^{-1}$ ).<sup>81,82</sup> The results of calculations for the extended intermediate forms of the four proteins under consideration are listed in Table II. Each value was obtained using the approximation that all  $\alpha$ -helices of the native state are intact in the extended intermediate. Note that the calculated values of  $\Delta C_{p,P}$  are somewhat larger than those of complete unfolding. The reason for this is that intrahelix interactions are predominantly polar; solvent exposure of polar surface results in a heat capacity decrease. Therefore, the formation of an  $\alpha$ -helix from a disordered chain is accompanied by a small *increase* in heat capacity.

Because a significant amount of polar surface remains buried in such partly folded extended states, the values in Table II may be considered upper limits on  $\Delta C_{p,D}$ . The data imply that, in the case of cold denaturation at a temperature low enough to intrinsically stabilize helices, the heat capacity increase of cold denaturation should be slightly larger than that of heat denaturation. The main conclusion of this subsection is that a transition from the native state to an extended conformation containing residual helical structure is expected to have a heat capacity increase larger than that expected for complete unfolding of the molecule.

#### Heat capacity changes of compact intermediates

The heat capacity difference between a compact intermediate state and any other conformation depends on the associated changes in solvent accessible polar and apolar surface areas. If the structure of

the intermediate is known or if a model can be built from structural information, then these quantities can be estimated using Eq. (2). For instance, surface area calculations based on Baldwin and coworkers' model of state I of apo-Mb permit estimation of  $\Delta C_{p,I \rightarrow D}$ . Under the assumptions that the folded regions of I (helices A, G, and H) are in a structural arrangement identical to that of the native state, that the helices are tight-packed, and that the remainder of the molecule is unfolded,  $\Delta C_{p,I \rightarrow D}$  is  $870 \pm 130 \text{ cal K}^{-1} \text{ mol}^{-1}$ . More realistically, because the interactions of this subdomain are much less specific in I than in N, a somewhat smaller  $\Delta C_{p,I \rightarrow D}$  should be expected experimentally. Indeed, examination of the heat capacity measurements by Privalov and coworkers indicate that  $\Delta C_{p,I \rightarrow D}$  is on the order of  $520 \text{ cal K}^{-1} \text{ mol}^{-1}$  (at 30°C states I and D are the principal ones at pH 4.2 and pH 2.5, respectively).<sup>68</sup> The experimental value suggests that approximately 33% of the hydrophobic core is intact in I, even though roughly 65% of the secondary structure is disrupted.

Cyt c presents another example. Wada and coworkers have measured the energetics of various intermediate states using spectroscopic methods.<sup>64</sup> They estimate a heat capacity increase for total unfolding on the order of  $1.4 \text{ kcal K}^{-1} \text{ mol}^{-1}$ , a figure close to both the (calorimetric) measured value and the calculated one (see Table I). The heat capacity differences  $\Delta C_{p,Ib \rightarrow D}$  and  $\Delta C_{p,Ic \rightarrow D}$  were estimated to be  $1.10 \pm 0.14 \text{ kcal K}^{-1} \text{ mol}^{-1}$  and  $0.40 \pm 0.01 \text{ kcal K}^{-1} \text{ mol}^{-1}$ . These values suggest that a significant hydrophobic core is intact for both Ib (~75%) and Ic (~30%). A large hydrophobic core in Ib is consistent with the considerable compactness exhibited by the state (see section on Structure).

A "transition state" of T4 lysozyme represents a similar situation. Schellman and colleagues recently estimated the relative heat capacities of this intermediate and the denatured state to be  $0.5 \pm 0.1 \text{ kcal K}^{-1} \text{ mol}^{-1}$  and  $2.18 \pm 0.01 \text{ kcal K}^{-1} \text{ mol}^{-1}$ .<sup>86,93</sup> Although no structural data have been published on the transition state, the thermodynamic values are nevertheless consistent with the energetics expected of a molten globule state. It appears that about 80% of the hydrophobic core of the folded form is intact in the transition state. In all likelihood structural studies will show that the transition state of T4 lys is in fact a molten globule state.

A recent thermodynamic study of the thermal unfolding transition of  $\alpha$ -LA, conducted at neutral pH and moderate GuHCl concentrations, suggested a  $\Delta C_{p,A}$  of about  $330 \text{ cal K}^{-1} \text{ mol}^{-1}$  and a  $\Delta C_{p,D}$  of about  $1800 \text{ cal K}^{-1} \text{ mol}^{-1}$ .<sup>85</sup> These figures correspond to a hydrophobic core that is 80% intact. Murphy et al. noticed that the experimental energetics are consistent with a local unfolding of the protein, probably involving only the disordering of the  $\text{Ca}^{2+}$  binding loop or a region of similar size.<sup>74</sup> According

TABLE II. Heat Capacity Changes for Extended Intermediates\*

Protein	$N_{\text{res, helix}}$	$\Delta C_{p,P,\text{calc}}$ (kcal K <sup>-1</sup> mol <sup>-1</sup> )	$\Delta C_{p,P \rightarrow D,\text{calc}}$ (kcal K <sup>-1</sup> mol <sup>-1</sup> )
$\alpha$ -Lactalbumin	64	2.3 $\pm$ 0.3	-0.49 $\pm$ 0.06
Cytochrome <i>c</i>	52	1.8 $\pm$ 0.2	-0.40 $\pm$ 0.05
Myoglobin	121	3.4 $\pm$ 0.4	-0.92 $\pm$ 0.12
T4 lysozyme	107	3.5 $\pm$ 0.4	-0.81 $\pm$ 0.11

\*Values were calculated as previously described in ref. 75. Data on the number of residues in a helical conformation were obtained from the Protein Data Bank files mentioned in the footnote to Table I.

to these data, the heat capacity difference between the unfolded state and the putative molten globule intermediate is on the order of 1400 cal K<sup>-1</sup> mol<sup>-1</sup>.

As regards A, the acid form of  $\alpha$ -LA, its compactness is believed to arise from the existence of a significant hydrophobic core. One of the first estimates of  $\Delta C_{p,A \rightarrow D}$  from denaturation experiments was 420  $\pm$  50 cal K<sup>-1</sup> mol<sup>-1</sup>, made by Kuwajima in 1977.<sup>13</sup> Since  $\Delta C_{p,D}$  is 1.5–1.8 kcal K<sup>-1</sup> mol<sup>-1</sup>,<sup>83,84</sup> the early  $\Delta C_{p,A \rightarrow D}$  value implies that about 25% of the hydrophobic core remains intact in the acid state. Remarkably, this percentage is similar to the one calculated for state I of apo-Mb.

In further consideration of  $\alpha$ -LA, the hydrogen exchange data of Dobson and co-workers on the guinea pig protein<sup>27</sup> and hen egg white lysozyme<sup>55</sup> are consistent with the following structural model of A. As described above, the experimental results suggest that the lobe of  $\alpha$ -LA containing the  $\beta$ -sheet is disordered in A and that all helical structures of the native state are intact in A, excluding the one found in the smaller of the two lobes (see Fig. 1). Accordingly, in the model residues 34–73 are completely unfolded, the remaining residues interacting as in the native state. Calculations based on this structure indicate that approximately 60% of the hydrophobic core is intact (the expected  $\Delta C_{p,A \rightarrow D}$  being 1100  $\pm$  160 cal K<sup>-1</sup> mol<sup>-1</sup>). However, because the interhelix polar and apolar contacts of this model are necessarily more specific than those of the protein in the acid state, this calculated value must be an overestimate, possibly by as much as a few hundred cal K<sup>-1</sup> mol<sup>-1</sup>, as in the case of apo-Mb discussed above.

Recently, Mark and van Gunsteren<sup>113</sup> performed molecular dynamics simulations aimed at trapping a molten globule state of lysozyme. Using this computational approach, these authors concluded that lysozyme unfolds in a two-stage process, the intermediate state being a quasistable one and having an arrangement of secondary structure different from that of the native form. Structural thermodynamic calculations based on the structure of the intermediate, which exhibits all the features of the consensus structure of the molten globule state described above, indicate that about 64% of the hydrophobic core is intact. The calculated heat capacity differ-

ence between the molten globule and denatured states is about 800 cal K<sup>-1</sup> mol<sup>-1</sup>.

It is interesting to compare the relative heat capacities discussed above with the predominant view of the last several years. During that time many researchers believed that the unfolding of the molten globule was accompanied by a small, if not zero, enthalpy change and by a negligible heat capacity change.<sup>25</sup> These ideas appear to have originated from interpretations of a calorimetric study of the acid state of human  $\alpha$ -LA.<sup>87</sup> According to the data obtained by Pfeil et al.,<sup>87</sup> the heat effect of the molten globule state to denatured state transition appears to be either small or absent altogether. Also, the heat capacity function of A exhibits no thermal transition peak and has a value close to that of the heat denatured *holoprotein*. However, it should be noted that it might not be appropriate to compare these two forms, since the latter is known to precipitate upon thermal denaturation. Precipitation lowers the apparent partial heat capacity in the region after the transition and reduces  $\Delta C_{p,D}$ .

Examination of the low temperature region of the data presented by Pfeil et al. shows that  $\Delta C_{p,A}$  for human  $\alpha$ -LA is about 850 cal K<sup>-1</sup> mol<sup>-1</sup>. Since the total heat capacity difference between the denatured and native forms is 1500–1800 cal K<sup>-1</sup> mol<sup>-1</sup>,  $\Delta C_{p,A \rightarrow D}$  must lie between 650 and 950 cal K<sup>-1</sup> mol<sup>-1</sup>. The high end of this range of values for  $\Delta C_{p,A \rightarrow D}$  does fall within the uncertainty of the heat capacity increase calculated from the model of the acid state based on hydrogen exchange data (1100  $\pm$  160 cal K<sup>-1</sup> mol<sup>-1</sup>). Moreover, this range of values encompasses that calculated for the molten globule state of lysozyme defined by Mark and van Gunsteren<sup>113</sup> (see above). On the other hand, the value of  $\Delta C_{p,A \rightarrow D}$  predicted by the model of Murphy et al.<sup>74</sup> is on the order of 500 cal K<sup>-1</sup> mol<sup>-1</sup> higher. This may mean that, though the acid state and the intermediate induced by GuHCl at neutral pH have similar properties,<sup>25,48</sup> the two forms differ energetically. Of course, any detailed analysis of these two forms must account for the enthalpies of protonation of all titratable groups and the heat of denaturant binding (see below). In any case, the acid state of  $\alpha$ -LA is believed to have intact elements of secondary structure and a sizable hydrophobic core; and

TABLE III. Enthalpy Changes at 60°C for Complete Unfolding\*

Protein	$\Delta H_{D,calc}$ (kcal mol <sup>-1</sup> )	$\Delta H_{D,exp}$ (kcal mol <sup>-1</sup> )	Ref.
$\alpha$ -Lactalbumin	94 $\pm$ 20	93 $\pm$ 6	84
		95.5 $\pm$ 0.1	85
Cytochrome c	77 $\pm$ 17	78 $\pm$ 5	92
Myoglobin	110 $\pm$ 30	101 $\pm$ 5	92
T4 lysozyme	120 $\pm$ 30	104 $\pm$ 0.5	93

\*Calculated values were obtained as previously described in ref. 75. The experimental figure for T4 lys was extrapolated from 28°C under the assumption that  $\Delta C_p$  is temperature independent.

regardless of its precise value,  $\Delta C_{p,A \rightarrow D}$  must be substantially greater than zero. Consequently, even if  $\Delta H_{A \rightarrow D}$  is zero at some temperature, it cannot be zero at all temperatures. Accurate enthalpy and heat capacity determinations under conditions in which the protein does not aggregate or precipitate are required for a precise assessment of the differences between these two forms.

In summary, increasing experimental evidence indicates that the heat capacity of the molten globule state must be substantially lower than that of the denatured form. This is primarily due to the existence of a sizable hydrophobic core in the molten globule state. Consequently, the molten globule and the unfolded state cannot be enthalpically equivalent.

## Enthalpy Changes

### Enthalpy changes of total unfolding

A conformational change is usually accompanied by a heat capacity change. The enthalpy change associated with each transition is, therefore, temperature dependent and obeys the standard thermodynamic equation

$$\Delta H(T) = \Delta H(T_R) + \Delta C_p \cdot (T - T_R) \quad (3a)$$

$$= \Delta H(T_R) + (\Delta C_{p,ap}^o \cdot \Delta A_{ap} + \Delta C_{p,pol}^o \cdot \Delta A_{pol}) \cdot (T - T_R). \quad (3b)$$

Evaluation of  $\Delta H(T)$  at any temperature requires knowledge of  $\Delta H$  at some reference temperature  $T_R$ , denoted " $\Delta H(T_R)$ ." A convenient reference for enthalpy calculations is the "enthalpy convergence temperature for proteins," 100  $\pm$  6°C, at which apolar contributions to  $\Delta H$  are assumed to be zero but polar ones are substantial.<sup>75,82</sup> Experimental evidence indicates that  $\Delta H(100^\circ\text{C})$  scales with the amount of polar surface buried in the *native* state and equals 35  $\pm$  1 cal mol<sup>-1</sup> per Å<sup>2</sup> of buried polar surface.<sup>75</sup> In conjunction with the above equations, these parameter values have been used to predict transition enthalpies which, on the average, deviate from the calorimetrically determined values after extrapolation to 60°C by  $\sim$ 7%.<sup>74</sup> Table III shows the experimental and calculated values of  $\Delta H_D(60^\circ\text{C})$  for  $\alpha$ -LA, cyt c, Mb, and T4 lys. Clearly, the experimen-

tal and calculated enthalpy and heat capacity (Table I) values are in good agreement, which suggests that all side chains of the temperature denatured states of these molecules are solvated and that these forms contain little, if any, intact secondary structure.

### Enthalpy changes of extended intermediates

Most of the heat effect of unfolding an extended intermediate comes from heat absorption during  $\alpha$ -helix unfolding. Recent calorimetric measurements indicate that this enthalpy increase is close to 1.3–1.5 kcal mol<sup>-1</sup> of amino acid residue.<sup>88</sup> This value can be used to calculate the expected enthalpy change of the hypothetical extended intermediates of the four proteins under discussion.

According to Jeng and Englander, about 85% of the helical content of N is intact in P, the extended state of cyt c.<sup>65</sup> This value corresponds to about 44 residues in helical conformation and suggests a  $\Delta H_{P \rightarrow D}$  of about 50–60 kcal mol<sup>-1</sup>. Although this enthalpy increase is relatively large, P does not undergo a cooperative thermal transition. The most likely reason for this behavior is that the helices of P do not interact strongly, as described above; consequently, each helix probably unfolds independently, giving rise to a very broad transition profile (see below). Extended intermediates have been described for RNase A also, but they are stable at sub-zero temperatures only.<sup>79,80</sup> In contrast, to date no intermediate of this type has been described for  $\alpha$ -LA, Mb, or T4 lys.

### Enthalpy changes of compact intermediates

The enthalpy difference between the molten globule and unfolded states can be estimated using the procedures employed above. The molten globule state contains a hydrophobic core and exhibits a heat capacity increase upon unfolding; therefore, its denaturation enthalpy is temperature dependent. Also, since most molten globule intermediates are stable only at extremely low pH values, in the presence of mild concentrations of denaturant, or at extreme salt concentrations, it is necessary to consider explicitly the contributions of protonation, denaturant binding, or ion binding to the apparent enthalpy

differences between states. This is important, because a complete characterization of the enthalpy of denaturation consists not simply of the enthalpy of conformational change but of the sum of all contributions to the enthalpy difference between states.

For example, the conformational contribution to the unfolding enthalpy of the low pH intermediate state I of apo-Mb ( $\Delta H_{I \rightarrow D}$ ) can be estimated using Eq. (3b), if the buried polar and apolar surface areas are known. In Baldwin and co-workers' structural model of I, about 3,150 Å<sup>2</sup> of apolar surface and 2,090 Å<sup>2</sup> of polar surface are inaccessible to solvent. These values translate into a  $\Delta H_{I \rightarrow D}(25^\circ\text{C})$  of  $8 \pm 11$  kcal mol<sup>-1</sup>. Since the buried imidazole rings of buried histidine residues become protonated when the protein unfolds at pH values below 6, the approximate  $pK_a$  of histidine in aqueous solution, and since the enthalpy of histidine protonation is relatively large (about -7 kcal mol<sup>-1</sup>), it is essential to consider these contributions to the measured enthalpy change. When the heat of histidine protonation is taken into account, the effective  $\Delta H_{I \rightarrow D}(25^\circ\text{C})$  is close to zero. The significance of this result will be discussed below.

In the structural model of the acid pH molten globule state of  $\alpha$ -LA described above, about 4,010 Å<sup>2</sup> of apolar surface and 2,600 Å<sup>2</sup> of polar surface are inaccessible to solvent. By Eq. (3b),  $\Delta H_{A \rightarrow D}$  is close to zero ( $6 \pm 14$  kcal mol<sup>-1</sup>) at 25°C. However, it must be remembered that  $\Delta C_{p,A \rightarrow D}$  is about 1,100 cal K<sup>-1</sup> mol<sup>-1</sup>. Thus at 43°C, the approximate transition temperature of apo- $\alpha$ -LA in the presence of 200 mM NaCl at pH 8.0,  $\Delta H_{A \rightarrow D}$  is about 25 kcal mol<sup>-1</sup>, a value considerably greater than zero. This example illustrates the typical range of variation in the magnitude of the unfolding enthalpy of a molten globule state within reasonable folding temperature limits.

The above discussion shows that the conformational portion of the denaturation enthalpies of two molten globule states approaches zero at 25°C. This should be not surprising, since 25°C is close to the inversion temperature for many globular proteins.<sup>89</sup> At room temperature both the molten globule to unfolded state transition and the native to unfolded state transition have enthalpy changes close to zero.

### Thermally Induced Cooperative Folding/Unfolding Transitions

An important characteristic of some molten globule states, the acid form of  $\alpha$ -LA in particular, is the absence of a measurable thermally induced cooperative transition.<sup>87</sup> There are two possible and compatible reasons for such behavior: (1) the enthalpy difference between the intermediate state and the unfolded state is near zero in the neighborhood of the transition temperature; and (2) cooperative interactions are drastically disrupted.

As long as there is a heat capacity difference between two conformational states, the associated

enthalpy difference cannot be temperature independent, even if it reduces to zero at some experimentally accessible temperature. This temperature dependence of  $\Delta H$  has a characteristic effect on thermal transitions. For instance, consider a perturbation of the melting temperature ( $T_m$ ) of a two-state transition by a change in pH, salt concentration, or denaturant concentration. Typically, the change in  $T_m$  so elicited does not significantly affect the temperature dependence of the enthalpy change.<sup>11</sup> That is, the intrinsic heat capacity change is equal to the one deduced from the temperature dependence of the enthalpy change, if all relevant binding enthalpies have been accounted for. Figure 6 illustrates this situation, presenting a representative family of calculated calorimetric transition curves. It is clear from this heat capacity surface that the thermal transition is sharp and is accompanied by a large transition enthalpy (area below the peak), when  $T_m$  is high. On the other hand, if  $T_m$  is low, the transition enthalpy is small and the curve is relatively broad; the calorimetric signal is weak. At some point the transition can no longer be detected using differential scanning calorimetry.

The foregoing example illustrates a possible source of incorrect interpretations of data pertaining to potential molten globule intermediates. As shown in Figure 6, *the absence of a measurable transition implies neither the existence of a molten globule intermediate nor a temperature independent transition enthalpy of zero.* If the native state can be destabilized by changing certain solvent conditions, then when  $T_m$  is low the transition will be broad and beyond the limits of detection, even if a molten globule state never becomes stable under the specific physical conditions of the folding/unfolding experiment. At any given temperature the sample would simply be a mixture of molecules in the native and unfolded states, although some of its spectroscopic properties would probably resemble those of a partly folded intermediate. Therefore, for experimental studies of partly folded states, it is advisable to investigate the dependence of several physical properties of the system on temperature and an additional independent variable, i.e., to take a multidimensional approach, globally analyzing the data to estimate system parameters.<sup>90</sup> Such an approach should help to preclude the situation encountered in Yutani et al.<sup>114</sup> On the basis of the absence of a measurable transition in a calorimetric scan of apo- $\alpha$ -LA at very low ionic strength and neutral pH, those authors concluded that the molten globule state is enthalpically equivalent to the unfolded state. As we have shown here, however, the absence of a thermally induced transition is not sufficient proof that two states of a protein are enthalpically equivalent.

Furthermore, because the enthalpy differences between the native and molten globule states, and be-

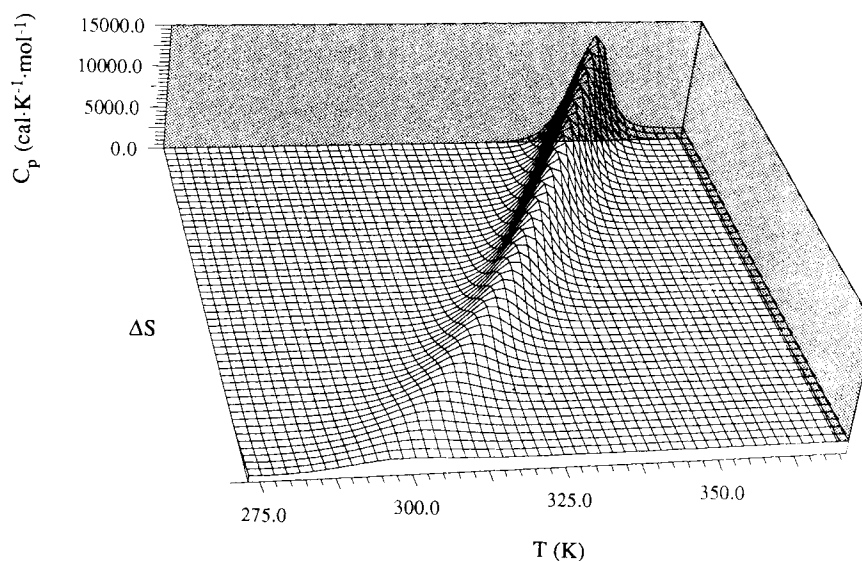


Fig. 6. Simulated excess heat capacity surface based on thermodynamic parameters for unfolding of  $\alpha$ -LA at neutral pH and in the absence of denaturants. The family of curves were calculated assuming a two-state transition mechanism. The heat capacity function ( $C_p$ ) is plotted against transition entropy ( $\Delta S$ ) and tem-

perature ( $T$ ). The transition enthalpy decreases with transition temperature, which itself declines as a result of changes in the temperature independent portion of the transition entropy (see text).

tween the molten globule and denatured states, are nearly zero at 25°C, it is better to choose a different temperature at which to evaluate the relative energetics of the relevant conformational states. This is one of the reasons why experimental determinations of the energetics of molten globule states have been so difficult. Additionally, under most experimental conditions the molten globule state is not the only species present in solution: it is in equilibrium with the native and denatured states. Consequently, in most cases experimental data must be corrected for the typically sizable populations of molecules in the other energetic states.

As regards extended intermediates, their unfolding is expected to be accompanied by a significant conformational enthalpy increase. This enthalpy difference is however only slightly temperature dependent, since the associated heat capacity change is small. The absence of cooperative thermal transitions for such states most likely follows from a breakdown in cooperative interactions between the individual helices along the polypeptide chain, i.e., from reduced tertiary structure interactions. Figure 7 shows a computer simulated heat capacity surface for a hypothetical P to D transition. The second independent variable is degree of cooperativity, which we have expressed (in the standard way) as the ratio of the van't Hoff enthalpy change to the calorimetric enthalpy change ( $\Delta H_{vH}/\Delta H_{cal}$ ).<sup>91</sup> A  $\Delta H_{vH}/\Delta H_{cal}$  of one is characteristic of a two-state transition and indicates maximal cooperativity for the unfolding of a monomeric protein; a ratio less than one implies a reduced cooperativity and a multistate process.

To summarize, Figure 7 illustrates two important points regarding the cooperativity of a transition. First, as the cooperativity of the transition decreases the transition curve broadens, so that at some point the latter diminishes beyond the limits of detectability. This occurs even when the temperature independent portion of the transition enthalpy is held constant. Second, as the transition becomes less cooperative and spans a wider temperature range, structural elements remain intact over a broader temperature interval.

The above examples show that the absence of a transition does not necessarily imply a zero enthalpy difference between the molten globule state and the denatured state. Also, contrary to some proposals, if a homogeneous protein sample exhibits a thermal unfolding transition the initial conformation may be a molten globule state: for it has been shown experimentally that molten globule states of cyt c<sup>62,100</sup> and of retinol binding protein,<sup>101</sup> for example, do in fact undergo thermal denaturation transitions. We expect that in the future more thermal transitions between molten globule and unfolded states will be reported.

### Stability of the Molten Globule State

#### General statistical thermodynamic formalism

In the preceding sections we have illustrated direct relationships between the thermodynamic parameters  $\Delta C_p$  and  $\Delta H$  and structural parameters. As to estimation of the configurational entropy, however, the situation is different. Currently it is not

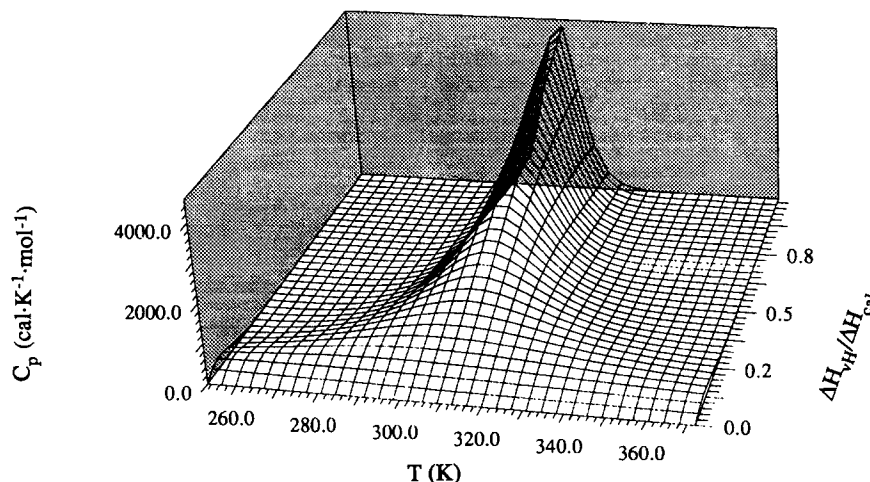


Fig. 7. The effect of cooperativity on the excess heat capacity ( $C_p$ ) associated with an hypothetical P to D transition. The enthalpy change corresponds to the isolated helical structures of cyt *c* (see text for details). The heat capacity function has been plotted against temperature ( $T$ ) and cooperativity ( $\Delta H_{vH}/\Delta H_{cal}$ ).  $\Delta H_{vH}/\Delta H_{cal} = 1$  corresponds to maximally cooperative thermal denaturation;  $\Delta H_{vH}/\Delta H_{cal} = 0$  to noncooperative unfolding. Though the

temperature dependence of the calorimetric transition enthalpy is the same for all values of cooperativity, the transition peak broadens as the cooperativity decreases. When structural domains of the protein unfold independently, the excess heat capacity is above zero over a broad temperature range. This implies that weakly cooperative systems possess elements of structure over a wider temperature range than highly cooperative systems.

possible to calculate accurate entropy changes from structural parameters. Nevertheless, it can be safely assumed that the conformational entropy of a partly folded state lies between those of the native and denatured states. We now illustrate a means of estimating the magnitude of entropy changes. With values for  $\Delta C_p$ ,  $\Delta H$ , and  $\Delta S$  in hand, we will be in a position to analyze the thermodynamic stability of an intermediate state.

The most fundamental quantity for describing a system in thermal equilibrium is the partition function. The general form of the canonical partition function,  $Q(T)$ , the sum of the statistical weights of all  $n + 1$  accessible conformational states, is

$$Q(T) = \sum_{i=0}^n \omega_i \cdot \exp(-\Delta G_i/RT) \quad (4a)$$

$$= 1 + \sum_{i=1}^{n-1} \omega_i \cdot \exp(-\Delta G_i/RT) + \omega_D \cdot \exp(-\Delta G_D/RT) \quad (4b)$$

$$= 1 + \sum_{i=1}^{n-1} \exp(-\Delta G_i/RT) + \exp(-\Delta G_D/RT) \quad (4c)$$

where  $\omega_i$  is the degeneracy of state  $i$ ;  $\Delta G_i$  and  $\Delta G_i$  are abbreviations of  $\Delta G_i(T)$  and  $\Delta G_i(T)$ ;  $\Delta G_i$  is the molar Gibbs free energy of the  $i$ th state relative to the free energy of a reference state [see Eqs. (1)];  $\Delta G_i$  is the form of the molar free energy when the degeneracy of the  $i$ th state has been expressed implicitly in the relative entropy function,  $\Delta S_i$ . Subscript D refers to the denatured state. Note that because the native state (N) has been selected as the reference state, its statistical weight is unity. A statisti-

cal weight which refers solely to a conformational change is sometimes called an intrinsic statistical weight. The partition function contains all the thermodynamic information of the protein system and is therefore sufficient to describe the thermal unfolding of globular proteins.<sup>10,105</sup> This treatment requires no a priori assumptions regarding the character of the transition or the magnitude of the thermodynamic parameters associated with the folding/unfolding equilibrium.

In the special case where only two states can be substantially populated (e.g., a typical globular protein in the absence of denaturant<sup>11</sup>), Eq. (4c) reduces to

$$Q \approx 1 + \exp(-\Delta G_D/RT) \quad (5)$$

where  $Q = Q(T)$ . This is the partition function for the two-state model, and it holds in the typical case because the molar free energies of the vast majority of intermediates are so high that the associated statistical weights effectively vanish, leaving just the two terms shown.<sup>10</sup>

The probability of a molecule being in a given state—the population of that state—is given by the statistical weight of the state divided by the sum of all statistical weights<sup>105</sup>:

$$P_N = 1/Q \quad (6a)$$

and

$$P_D = \exp(-\Delta G_D/RT)/Q \quad (6b)$$

where  $Q$  is given by Eq. (5) in the approximation that the relation is an identity. These two populations depend on temperature and are identical at




INDEX	STATE	FREE ENERGY	STAT. WEIGHT
N		0	1
J		$\Delta G_J = G_J - G_N$	$e^{-\Delta G_J/RT}$
D		$\Delta G_D = G_D - G_N$	$e^{-\Delta G_D/RT}$

Fig. 8. Matrix summarizing the thermodynamic formalism for three state unfolding. The first column contains the state labels. Schematic diagrams of the native (N) state, the molten globule state (J), and the denatured state (D) appear in column two. The third and fourth columns show the relative molar Gibbs free energy and the statistical weight for each state. The partition function equals the sum of the statistical weights.

$T_m$ . At this temperature the free energy difference between the states is zero.

Now, if a stable intermediate state can be significantly populated, it will not be possible to account for the folding/unfolding data with the two-state model. Let us suppose that only one additional state, say a molten globule state, is relevant to a thermodynamic description of the folding process. N, D, and the compact intermediate state are depicted schematically in Figure 8 along with their associated free energies and intrinsic statistical weights. To avoid confusion with the state I of apo-Mb and to underscore the generality of the present treatment, the subscript J (instead of I) will be used to represent the intermediate. For this case the partition function is

$$Q = 1 + \exp(-\Delta G_J/RT) + \exp(-\Delta G_D/RT) \quad (7)$$

and the populations of states are

$$P_N = 1/Q \quad (8a)$$

$$P_J = \exp(-\Delta G_J/RT)/Q \quad (8b)$$

and

$$P_D = \exp(-\Delta G_D/RT)/Q. \quad (8c)$$

Applying the above mathematical apparatus for a three state equilibrium process, we can probe the conditions necessary to stabilize molten globule states. The enthalpy and heat capacity changes can be calculated from the structure as described above or, for the case of the denatured state, they can be measured experimentally. The only remaining parameters are the entropy changes. One of these relative entropies ( $\Delta S_D$ ) can be obtained experimentally using differential scanning calorimetry. The other parameter ( $\Delta S_J$ ) cannot be calculated accurately; therefore, we must estimate its magnitude,

setting lower and upper limits using experimentally derived constraints. One method for doing this is to simulate the expected behavior of the system as a function of the appropriate independent variable (e.g., temperature) and the unknown entropy term. This procedure is illustrated below.

In the first example, we analyze the behavior of Eq. (8b) under different conditions by substituting in energetic parameter values for the complete unfolding of apo- $\alpha$ -LA. The effective overall energetic parameters ( $\Delta H_D$ ,  $\Delta C_{p,D}$ ,  $\Delta S_D$ ) for this case have been estimated experimentally by several authors and computationally as well (see Tables I and III). Figure 9A shows the temperature dependence of the population of J as a function of  $\Delta S_J$ ; the enthalpy and heat capacity values [ $\Delta H_J(25)$ ,  $\Delta C_{p,J}$ ] were taken from Xie et al.<sup>85</sup> With the energetic parameters fixed at the values given in the figure legend,  $\Delta S_J(25^\circ\text{C})$  values on the order of  $17 \text{ cal K}^{-1} \text{ mol}^{-1}$  or less yield negligible populations of the intermediate at all temperatures. In contrast,  $\Delta S_J(25^\circ\text{C})$  values of about  $30 \text{ cal K}^{-1} \text{ mol}^{-1}$  give intermediate populations of approximately 90%, for temperatures below the transition region. Of course, higher entropies would result in higher intermediate populations over wider temperature ranges. Thus, for the energetic parameters in Figure 9,  $\Delta S_J(25^\circ\text{C})$  must be between 17 and  $30 \text{ cal K}^{-1} \text{ mol}^{-1}$  in order for the population of the intermediate to be between 10 and 90%.

For comparison, Figure 9B displays the expected behavior when the enthalpy and heat capacity differences between N and J and between N and D are assumed to be the same. This situation corresponds to the view of Yutani et al.,<sup>114</sup> i.e., that the molten globule state and the unfolded state are enthalpically equivalent. Under this constraint, J is not significantly populated for  $\Delta S_J(25^\circ\text{C}) < 90 \text{ cal K}^{-1} \text{ mol}^{-1}$ . Moreover, when  $\Delta S_J = \Delta S_D$  states J and D are equally populated at temperatures above the transition temperature, whereas when  $\Delta S_J > 92 \text{ cal K}^{-1} \text{ mol}^{-1}$ , J is the predominant conformation in the same temperature range. Since we have set  $\Delta C_{p,J} = \Delta C_{p,D}$ , the solvation contributions to the entropy change must be identical for both states, and the entropy differences must be of a configurational origin. Thus, if it is assumed that the molten globule to denatured state transition involves no heat capacity change and no enthalpy change, then the intrinsic stability of the molten globule must be of conformationally entropic origin. It is difficult, however, to rationalize a higher configurational entropy for the molten globule than for the denatured state, especially since it is known that the molten globule is partly ordered and compact.

In general, there are two main contributions to  $\Delta S_J$ . The positive one is attributable to the increased number of degrees of freedom; the negative one arises from the increased exposure of hydrophobic



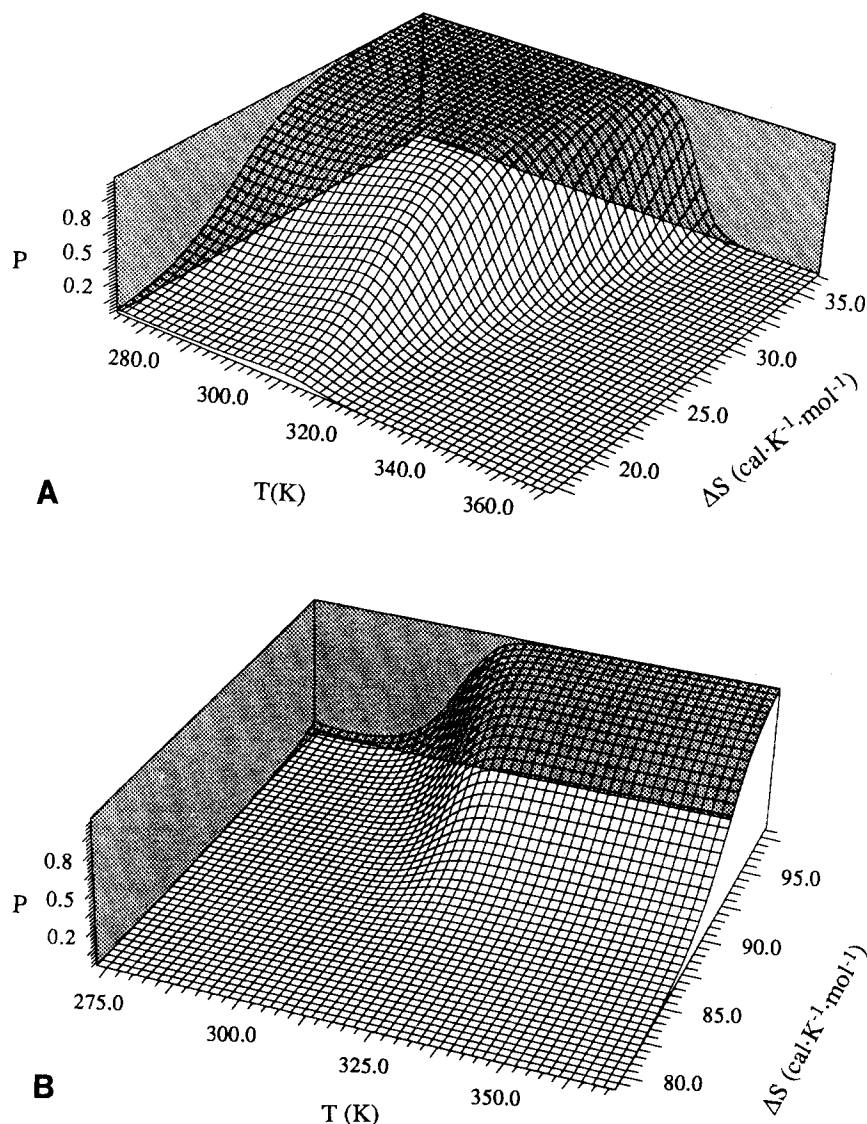


Fig. 9. Two surfaces showing the dependence of the population of the intermediate state on temperature and transition entropy,  $\Delta S_{\ddagger}$ . The surfaces were calculated using  $\alpha$ -LA energetic parameter values, which are based on data from the crystal structure of baboon  $\alpha$ -lactalbumin and on experiments performed at pH 8.0 in the presence of 200 mM NaCl. (a) From Tables I and III,  $\Delta C_{p,D} = 1,800 \text{ cal K}^{-1} \text{ mol}^{-1}$  and  $\Delta H_D(25^\circ\text{C}) = 30 \text{ kcal mol}^{-1}$ . Under the stated solution conditions, the maximum in the heat capacity function ( $T_m$ ) occurs at about  $45^\circ\text{C}$ . Knowledge of  $T_m$

permits calculation of the transition entropy at  $T_m$ , which can be extrapolated to any other temperature under the assumption of a constant  $\Delta C_p$ :  $\Delta S_D(25^\circ\text{C}) = 92 \text{ cal K}^{-1} \text{ mol}^{-1}$  [Eq. (1b)]. Experimental data (ref. 85) suggest that for the GuHCl-induced intermediate  $\Delta C_{p,\ddagger}$  and  $\Delta H_{\ddagger}(25^\circ\text{C})$  may be on the order of 400 and 7 kcal  $\text{mol}^{-1}$ , respectively. (b) In this case, the enthalpy and heat capacity of the molten globule and denatured states have been assumed to be identical;  $\Delta H_D(25^\circ\text{C}) = \Delta H_{\ddagger}(25^\circ\text{C}) = 30 \text{ kcal mol}^{-1}$  and  $\Delta C_{p,D} = \Delta C_{p,\ddagger} = 1800 \text{ cal K}^{-1} \text{ mol}^{-1}$ .

groups. The latter is reflected in a  $\Delta C_{p,\ddagger}$  on the order of  $400 \text{ cal K}^{-1} \text{ mol}^{-1}$  for the neutral pH molten globule state. The contribution to the entropy change due to solvation can be calculated easily<sup>75</sup> and is close to  $-100 \text{ cal K}^{-1} \text{ mol}^{-1}$  at  $25^\circ\text{C}$ , or about  $-80 \text{ cal K}^{-1} \text{ mol}^{-1}$  at  $43^\circ\text{C}$ , the midpoint temperature of the N  $\rightarrow$  D transition of bovine  $\alpha$ -LA at pH 8.0 and in the presence of 200 mM NaCl. Upon passing from the native state to the unfolded state, an amino acid gains an average of  $4.3 \pm 0.1 \text{ cal K}^{-1} \text{ mol}^{-1}$  of configurational entropy.<sup>89</sup> This value includes

the increase in number of degrees of freedom of both the side chains and the backbone. If we assume that about 20% of the total number of residues have lost their tertiary structural contacts and are as disordered as in the denatured state, then the corresponding configurational entropy gain is about  $105 \text{ cal K}^{-1} \text{ mol}^{-1}$ . Recently, the average gain in entropy upon sidechain melting was estimated to be on the order of  $2.8 \text{ cal K}^{-1} \text{ mol}^{-1}$ .<sup>96</sup> So if only the side chains were to gain configurational freedom, the entropy increase would be about  $70 \text{ cal K}^{-1} \text{ mol}^{-1}$ . At

25°C, the combined contributions give the total  $\Delta S_J$ , which is about  $5 \text{ cal K}^{-1} \text{ mol}^{-1}$ . This value falls short of the  $\sim 25 \text{ cal K}^{-1} \text{ mol}^{-1}$  required to stabilize the molten globule (see above). The situation is similar for our structural model of the acid molten globule state of  $\alpha$ -LA: the expected solvation entropy is  $-180 \text{ cal K}^{-1} \text{ mol}^{-1}$  and the calculated configurational entropy gain is  $172 \text{ cal K}^{-1} \text{ mol}^{-1}$ .

The absence of specific tertiary structural interactions is expected to further stabilize the molten globule state. All the configurational entropic factors combined bring the intrinsic statistical weight of this state into a range in which the additional free energy provided by "molten globule stabilizing agents" is sufficient for stabilization. Thus, the range of expected entropy values suggests that partly folded states having structural characteristics of the molten globule state are "quasistable": molten globules are not intrinsically stable, but because their relative free energy is close enough to zero over some temperature range, they can be stable in the presence of some additional factors, including GuHCl, protons, salt, etc. This situation is not generally true for purely local intermediates, in which a small region of the protein is unfolded but the rest is native-like.<sup>74</sup>

### Stabilization of the molten globule state by GuHCl

The thermal stability of the molten globule state depends significantly on interactions arising from perturbed solvent conditions. For example, moderate GuHCl concentrations have been shown to stabilize the molten globule state of apo- $\alpha$ -LA at neutral pH.<sup>48</sup> In such cases, the partition function in Eq. (7) must be modified suitably by a term dependent on the concentration or activity of GuHCl. For the independent and identical site-binding model of GuHCl-protein interaction<sup>112</sup> the form of the partition function is

$$Q = 1 + \exp(-\Delta G_J/RT) \cdot (1 + k_b \cdot a)^{\Delta n_J} + \exp(-\Delta G_D/RT) \cdot (1 + k_b \cdot a)^{\Delta n_D} \quad (9)$$

where  $k_b = k_b(T)$ , the average binding constant of GuHCl, which is assumed to be the same in all three states;  $a = a_{\text{GuHCl}}$ , the activity of GuHCl;  $Q$  has been used even though (9) represents a class of partition function which is usually denoted by a different symbol; and  $\Delta n_J$  and  $\Delta n_D$  are the difference in the number of GuHCl binding sites between J and N and between D and N, respectively. Note that Eq. (9) reduces to Eq. (7) in the case of  $\Delta n_J = \Delta n_D = 0$ .

Clearly,

$$(1 + k_b \cdot a)^{\Delta n_i} \geq 1, \quad i = J, D \quad (10)$$

since  $k_b$ ,  $a$ ,  $\Delta n_J$ , and  $\Delta n_D$  all are greater than or equal to zero. Moreover, this factor must be greater for D than for J. How, then, does GuHCl stabilize the

molten globule state and enable it to become the predominant species? For physically reasonable conditions (for example, a population of intermediate no greater than 10% in the absence of denaturant), such a situation is possible for temperatures at which the native state is predominantly populated and the statistical weights of J and D are of similar magnitude (Haynie and Freire, in preparation). Under conditions in which the denatured state is already the most significantly populated state, however, GuHCl cannot stabilize the molten globule due to the constraint that  $\Delta n_J < \Delta n_D$ .

The population of any state can be computed from a suitably modified version of the relevant Eq. (8). For the intermediate state we have

$$P_J = [\exp(-\Delta G_J/RT) \cdot (1 + k_b \cdot a)^{\Delta n_J}] / Q. \quad (11)$$

This relation has some interesting features which allow us to address important questions related to the stability of the molten globule state. For instance, What GuHCl activity should be used to maximize the population of the intermediate and to minimize contributions from the native state and denatured state?

In order to find the GuHCl activity at which the population of the intermediate state is a maximum, we differentiate equation (11) with respect to denaturant activity and solve the equation ( $\partial P_J / \partial a = 0$ ) for  $a$ . The resulting expression (Haynie and Freire, in preparation) is

$$a|_{P_J = \max} = \{[\Delta n_J / (\Delta n_D - \Delta n_J)] / \exp(-\Delta G_D/RT)\}^{1/\Delta n_D} - 1 / k_b. \quad (12)$$

The form of Eq. (12) leads to the conclusion that, for the 3-state folding/unfolding model represented by Eq. (9), the denaturant activity at which the population of J is maximal is *independent* of the energetics of J, the partly folded state. Instead, this function depends weakly on  $\Delta n_J$  and strongly on the (temperature dependent) intrinsic statistical weight of D. This result resembles an established ligand-binding property: the mass-law description of the binding of a ligand requires that the population of the  $i$ th ligated species be maximal when the average amount of ligand specifically bound per mole of macromolecule equals  $i$ , not when the intrinsic energetics of the state meet some special criteria.<sup>95</sup> Additionally, Eq. (12) indicates that the location of the maximum of  $P_J$  with respect to ligand activity varies with the value of the binding constant, another temperature dependent function. Figure 10 shows a pair of curves calculated with experimental parameters for apo- $\alpha$ -LA. These effective parameters have been employed here mainly to illustrate the effect of denaturant binding. In this case, the only unknown parameter was  $\Delta n_J$ , which as Figure 10 shows is expected to be around 20, since experiments have shown that the

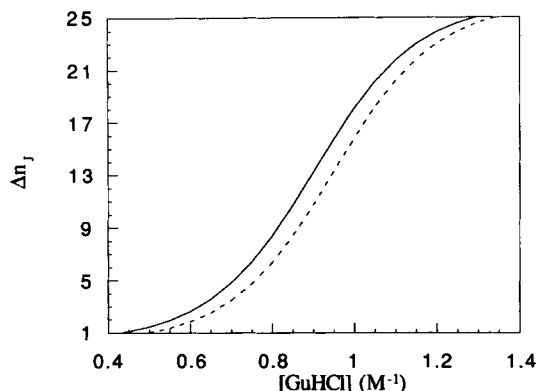


Fig. 10. Number of binding sites in the intermediate state of apo  $\alpha$ -LA versus concentration of GuHCl. At 298 K, the concentration of denaturant at which the population of J is a maximum depends on the binding parameters and on the molar free energy increase of *total* unfolding; it is independent of the intrinsic energetics of the intermediate state. In this case, Eq. (12) has been solved for  $\Delta n_J$ . The curves represent the required increase in number of binding sites upon unfolding from the native state to the intermediate state, for a given concentration of GuHCl, in order that the population of the intermediate be a maximum.  $\Delta G_D(25^\circ\text{C})$  was calculated using the parameters given in (a) of the legend of Figure 9 (solid line) and those given in ref. 84 (dashed line).

intermediate population is maximal above 1 M GuHCl.

Although the denaturant activity promoting the maximum population of intermediate state is independent of the relative free energy of the intermediate state, Eq. (11) indicates that the population of J certainly does depend on its energetics. Perhaps more importantly, it depends on the magnitude of  $\Delta G_J$  relative to  $\Delta G_D$ . This equation can be used to ascertain whether such a model can describe the observed experimental behavior by substituting in values for  $\alpha$ -LA, for example.

To discover the physical conditions consistent with the model of Eq. (9) which permit GuHCl stabilization of the molten globule intermediate, we have performed a series of computer simulations. For a given binding model and a given set of energetic parameters for the N to D transition, the algorithm exhaustively samples all possible sets of N to J transition parameters, GuHCl activities, and temperatures; it determines which sets of parameters that are consistent with a stable intermediate state (Haynie and Freire, in preparation).

For example, for the case of apo  $\alpha$ -LA at pH 8.0 and in the presence of 200 mM NaCl, the energetics of the N to D transition is defined by the following experimental parameters (see Tables I, III, and references therein):

$$\begin{aligned}\Delta H_D(25^\circ\text{C}) &= 30 \text{ kcal mol}^{-1} \\ T_{m,D} &= 316 \text{ K} \\ \Delta C_{p,D} &= 1800 \text{ cal K}^{-1} \text{ mol}^{-1} \\ \Delta n_D &= 26.\end{aligned}$$

The denaturant binding parameters are:

$$\begin{aligned}k_b(25^\circ\text{C}) &= 0.76 \text{ M}^{-1} \\ \Delta H_b &= -2600 \text{ cal mol}^{-1}\end{aligned}$$

according to Makhatadze and Privalov.<sup>102</sup> The intrinsic parameter values should be considered effective ones, since they include the ion binding energetics. Under the specified conditions the “neutral pH” intermediate can be observed at moderate GuHCl activities. The folding/unfolding parameters above have been estimated independently by several authors and are in substantial agreement with each other.<sup>84,85</sup> As to  $\Delta n_D$ , the value for  $\alpha$ -LA is close to one derived from a non-linear least squares fitting of differential scanning calorimetry data<sup>85</sup> and to that obtained from isothermal titration calorimetry experiments by Okuda and Sugai.<sup>97</sup> Furthermore, it is nearly identical to the one resolved by calorimetry for homologous hen egg white lysozyme.<sup>102</sup>

The remaining parameter values, which define the parameter space to be “sampled” and characterize the intermediate state transition, were permitted to assume values at discrete points within the following domains of the function  $P_J$ :

$$\begin{aligned}-30 \text{ kcal mol}^{-1} &\leq \Delta H_J(25^\circ\text{C}) \leq 30 \text{ kcal mol}^{-1} \\ 200 \text{ cal mol}^{-1} &\leq \Delta C_{p,J} \leq 1800 \text{ cal mol}^{-1} \\ 316 \text{ K} &\leq T_{m,J} \leq 396 \text{ K} \\ 0 &\leq \Delta n_J \leq 25.\end{aligned}$$

Each set of parameters defines a “parameter vector,” the “magnitude” and “direction” of which depend on the four intrinsic properties of the protein state indicated above. Each vector “points” to a two-dimensional matrix in parameter space. The elements of this matrix constitute a “sample” of the population contour as a function of  $a_{\text{GuHCl}}$  and temperature, two experimentally adjustable parameters which were allowed to vary from 0 to 6 M and from 0 to 100°C, respectively.

This type of analysis shows that it is possible to shrink the otherwise vast parameter space down to the relatively small subdomain in which the molten globule state is stabilized. The shape of the curve in Figure 11 depends on a set of parameter values that is consistent with the observed properties of  $\alpha$ -LA under the solution conditions mentioned above. Figure 11A represents the limiting situation in which there is no enthalpy difference and no heat capacity difference between the intermediate and denatured states [i.e.,  $\Delta H_J(25^\circ\text{C}) = \Delta H_D(25^\circ\text{C})$ ;  $\Delta C_{p,J} = \Delta C_{p,D}$ ]. Apparently, these conditions do not lead to a significant stabilization of J: for even if  $\Delta n_J$  is allowed to increase to  $\Delta n_D - 1$  (certainly an overestimate of  $\Delta n_J$ ), the population of J never exceeds 30% of the total for all points in the experimentally accessible domain. Significant stabilization of the intermediate is obtained only for  $\Delta H_J(25^\circ\text{C}) < 20 \text{ kcal mol}^{-1}$ . As  $\Delta H_J(25^\circ\text{C})$  increases,  $\Delta n_D$  also must in-

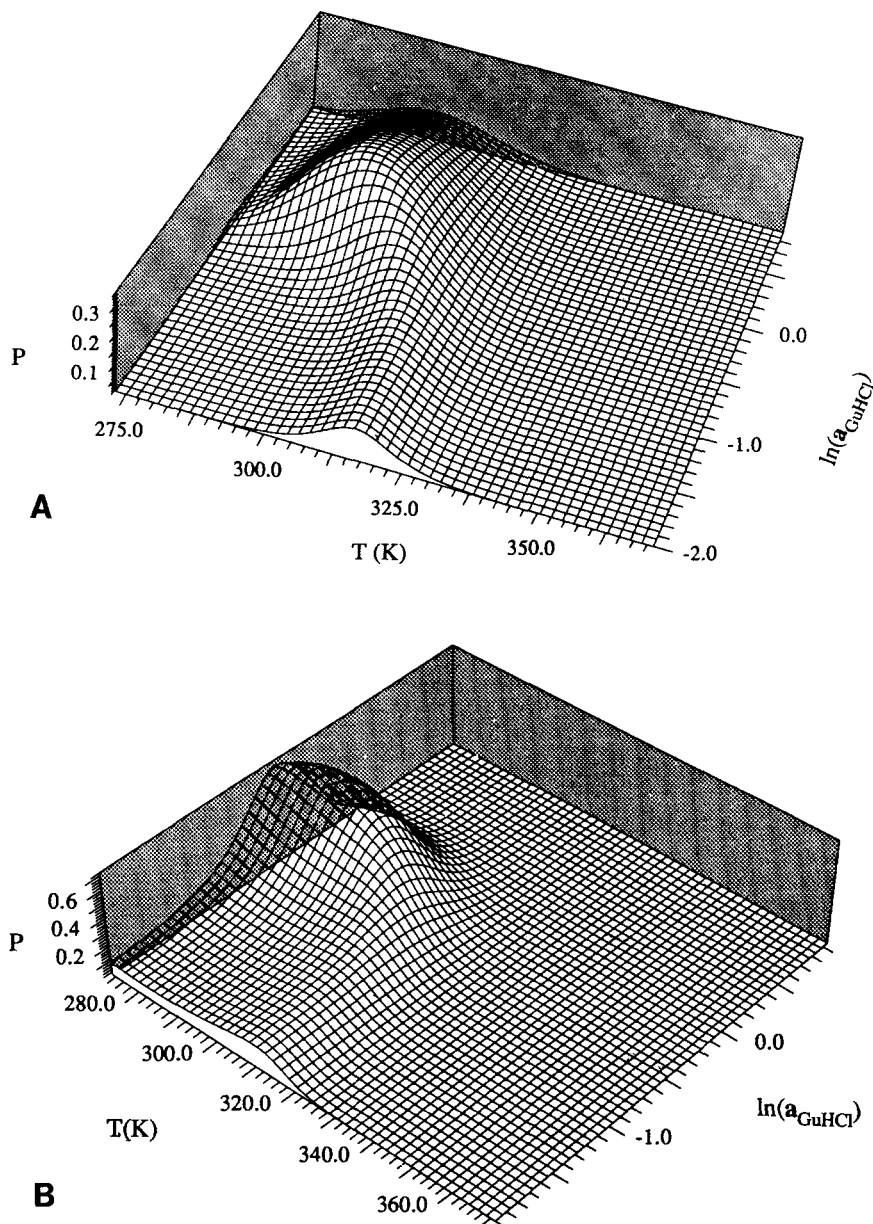


Fig. 11. Hypothetical intermediate state population surfaces for apo- $\alpha$ -LA as a function of the natural logarithm of GuHCl activity and of the absolute temperature. (A) The extreme situation in which there is no enthalpy difference or heat capacity difference between the molten globule state and the denatured state.  $\Delta H_J(25^\circ\text{C}) = \Delta H_D(25^\circ\text{C}) = 30 \text{ kcal mol}^{-1}$ ,  $\Delta C_{p,J} = \Delta C_{p,D} = 1,800 \text{ cal K}^{-1} \text{ mol}^{-1}$ . Under such conditions, J is not significantly

stabilized at any concentration of GuHCl; the population of J does not exceed 30% anywhere in the experimentally accessible domain. Significant stabilization of J is possible only for  $\Delta H_J(25^\circ\text{C}) < 20 \text{ kcal mol}^{-1}$  (see text for further discussion). In (B)  $\Delta H_J(25^\circ\text{C}) = 7 \text{ kcal mol}^{-1}$ ,  $\Delta n_J = 18$ . In this case, the population of the molten globule state is maximal ( $\sim 70\%$ ) when the GuHCl activity is 0.67 and the temperature is about  $13^\circ\text{C}$ .

crease in order to obtain significant stabilization; if  $\Delta H_J(25^\circ\text{C})$  is near  $20 \text{ kcal mol}^{-1}$   $\Delta n_J$  must equal  $\Delta n_D - 1$ . This analysis can be used to define a practical upper limit for  $\Delta H_J(25^\circ\text{C})$ .

As to a lower limit, negative  $\Delta H_J(25^\circ\text{C})$  values also can stabilize the molten globule state, at least in principle; however, on physical grounds it appears unlikely that a partly folded state has a lower enthalpy than the native state. Therefore, it can be

assumed that  $\Delta H_J(25^\circ\text{C})$  lies somewhere between 0 and  $20 \text{ kcal mol}^{-1}$ . In Figure 11B,  $\Delta H_J(25^\circ\text{C})$  has been set to  $7 \text{ kcal mol}^{-1}$ ,  $\Delta n_J$  to 18. We use these values simply because they are consistent with the analysis provided just above and with some experimental data.<sup>85</sup> For these conditions, the population of J is maximal when the GuHCl activity is 0.67 ( $[\text{GuHCl}] \sim 1.5 \text{ M}$ ) and the temperature  $10\text{--}15^\circ\text{C}$ . At lower GuHCl concentrations or lower temperatures,

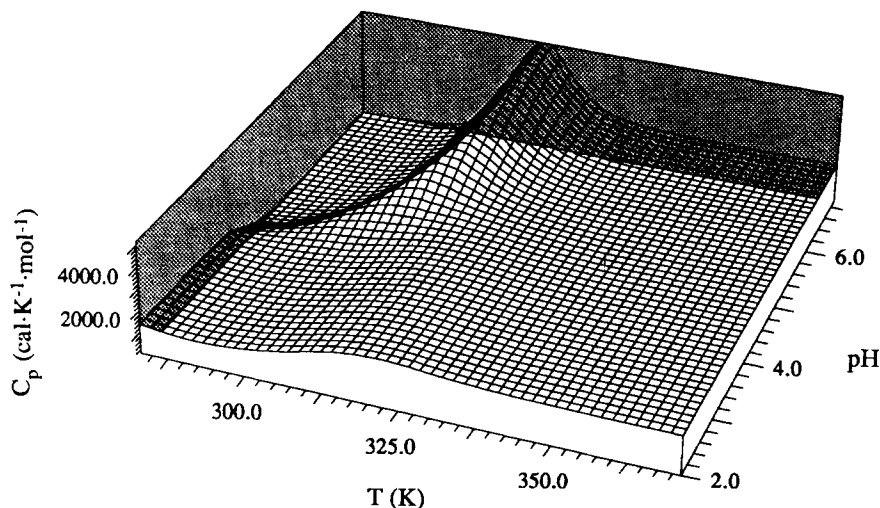


Fig. 12. Heat capacity surface for pH denaturation of  $\alpha$ -LA. The independent variables are temperature and pH. The thermodynamic parameters of the structural transition were derived from the structural model of the acid state described in the text. Protonation constants for four carboxyl groups and one histidyl groups are from ref. 48; the enthalpies of protonation of carboxyl

groups and of histidyl groups were assumed to be  $-1.0$  and  $-7.0$  kcal mol $^{-1}$ , respectively.<sup>108</sup> As the pH decreases, the transition peak gradually diminishes in size. The heat capacity surface defines the regions in which an accessible state is predominantly stable (see text for additional details).

the predominant state is the native state; whereas at higher GuHCl concentrations or higher temperatures, the denatured state becomes the most abundant one. In comparison, according to CD spectroscopic data on  $\alpha$ -LA at 25°C in a solution buffered with 50 mM cacodylate–50 mM NaCl at pH 7.0 ( $[Na^+] = 0.1$  M), the population of the molten globule state appears to be close to 0.5 at a GuHCl concentration of about 1.75 M.<sup>103</sup>

According to the structural energetics measured experimentally, at neutral pH and high salt the state stable at high temperatures is the denatured state—not the molten globule state. This situation is illustrated in Figure 11A and B and is valid for a wide range of values for the energetic parameters. Despite the existence of a residual far-UV signal in the CD spectrum of the heat denatured state of  $\alpha$ -LA, the experimental thermodynamic parameter values for the thermal unfolding of  $\alpha$ -LA at neutral pH and high salt concentration (0.2 M NaCl) are consistent with a complete or near complete solvation of all sidechains. The apparent conflict between the thermodynamic and spectroscopic data constitutes an important outstanding issue which needs to be resolved.

### Stabilization of molten globule state by pH

Stabilization of the molten globule state at low pH values can be described by a different modification of the partition function:

$$Q = 1 + \exp(-\Delta G_D/RT) \cdot \prod_{i=1}^m$$

$$(1 + K_{J,i} \cdot a_H) \cdot (1 + K_{N,i} \cdot a_H)^{-1} + \exp(-\Delta G_D/RT) \cdot \prod_{i=1}^m (1 + K_{D,i} \cdot a_H) \cdot (1 + K_{N,i} \cdot a_H)^{-1} \quad (13)$$

where the multiplication runs over all  $m$  protonizable groups;  $K_{N,i}$ ,  $K_{J,i}$ , and  $K_{D,i}$  are the protonation constants of the  $i$ th group in the N, J, and D states, respectively; we have retained the symbol  $Q$  as before; and  $a_H$  is the hydrogen ion activity. In general, the temperature dependence of the protonation constants must be considered explicitly.

We illustrate the expected behavior of this system using the thermodynamic parameters calculated for the structural model of the acid state of  $\alpha$ -LA described above and the titration parameters of Kuwajima and collaborators.<sup>48</sup> According to those authors, at room temperature, three carboxyl groups have the values  $pK_N = 3.3$ ,  $pK_A = 4.4$ , and  $pK_D = 4.4$ ; one carboxyl group has  $pK_N = 3.8$ ,  $pK_A = 4.4$ , and  $pK_D = 4.4$ ; and one histidyl group has  $pK_N = 5.8$ ,  $pK_A = 6.3$ , and  $pK_D = 6.3$ . For our calculations, the protonation enthalpies of the carboxyl and histidyl groups were set to  $-1.0$  and  $-7.0$  kcal mol $^{-1}$ , respectively.<sup>108</sup>

The expected excess heat capacity surface as a function of temperature and pH is displayed in Figure 12, which shows that the characteristic transition peak becomes less pronounced as the pH decreases and eventually disappears at pH values lower than  $\sim 4$ . The heat capacity surface graphically illustrates the regions in which the various states are predominantly stable. Specifically, at low temperatures and at pH values close to neutral, the native state predominates, giving rise to the “de-

pression" in the upper left hand side of the surface. In the same temperature range but at low pH values, the molten globule state is most stable. At temperatures above the transition peak and at pH values close to neutral, the denatured state is the predominant one. When the pH is low, increases in temperature gradually transform the molten globule state into the denatured state. Importantly, the transition in this region is *not* accompanied by a significant heat absorption peak, which agrees well with the experimental observations of Pfeil et al.<sup>87</sup> and of Dobson and colleagues.<sup>49</sup> The former have noted the absence of a heat absorption peak in the thermal unfolding of the acid state of human  $\alpha$ -LA; the latter have found a gradual thermal denaturation of the acid state of guinea pig  $\alpha$ -LA.

Attempts to simulate the behavior of the acid state using the thermodynamic parameters that model the behavior of the GuHCl molten globule were unsuccessful. This result suggests that the energetic properties of the GuHCl-induced intermediate differ significantly from those of the acid state. Above we noted that  $\Delta H_A(25^\circ\text{C})$  is close to 25 kcal mol<sup>-1</sup>, which is on the order of 10–15 kcal mol<sup>-1</sup> more than for the GuHCl-induced intermediate state. As discussed above, raising  $\Delta H(25^\circ\text{C})$  of the GuHCl-induced molten globule state to 25 kcal mol<sup>-1</sup> does not suitably simulate the characteristics of this form, even if the number of GuHCl binding sites in the intermediate is allowed to be as high as the number in the denatured state. We conclude, therefore, that the acid form may indeed have a somewhat higher intrinsic enthalpy and may be less ordered than the GuHCl-induced intermediate. The available structural data (primarily CD) provide low resolution information only and cannot discriminate potential differences between these forms. Also, it must be noted that the effects of pH are specific, the binding sites being ionizable groups located at distinct positions in the folded structure; whereas the effects of a denaturant like GuHCl are quite non-specific.

Finally, we consider the low pH stabilization of state I of apo-Mb. In a previous section we used the model of Hughson et al.<sup>58,69</sup> to calculate  $\Delta C_{p,I} \approx 850$  cal K<sup>-1</sup> mol<sup>-1</sup> and  $\Delta H_I(25^\circ\text{C}) \approx 8$  kcal mol<sup>-1</sup>. Differential scanning calorimetry studies of the stability of apo-Mb by Griko et al. indicate that at pH 5.0 the thermal unfolding is centered about 61°C and is characterized by a  $\Delta H_D$  of 53 kcal mol<sup>-1</sup> and a  $\Delta C_{p,D}$  of 1.56 kcal K<sup>-1</sup> mol<sup>-1</sup>.<sup>68</sup> At low temperatures, the pH titration curve of apo-Mb is biphasic, with inflection points near pH 3.5 and 5.5.<sup>69</sup> When the pH is above 6 the native form predominates, while at pH values less than 3 the denatured form is the most significantly populated. In the plateau region between the two inflection points ( $\sim$ pH 4) state I is stable. These results suggest that a transition from N to I involves the exposure of some buried groups

but the continued protection of others that do not become protonated until the transition to the denatured state.

Recently, pH titrations of histidine residues in sperm whale apo-Mb were monitored by Lecomte and coworkers using NMR.<sup>104</sup> Four histidyl groups were found to have anomalously low  $pK_a$ s clustered around 5 (His-24,  $pK_a < 4.8$ ; His-48,  $pK_a = 5.2$ ; His-113,  $pK_a < 5.5$ ; His-119,  $pK_a = 5.3$ ). The titration of histidines near to the heme binding site (His-64, His-93, and His-97) could not be observed by NMR. These three residues are completely buried in the native state; in horse apo-Mb they do not titrate even at the lowest pH values that could be studied (4.8).<sup>104</sup> There are no buried carboxyl groups in the folded form of myoglobin,<sup>111</sup> a situation expected to be similar in apo-Mb. If so, the low pH inflection point might be attributable to the histidines with the very low  $pK_a$  values.

The above energetic and pH titration data have been used to generate a stability surface of apo-Mb (Fig. 13). Slight variations in the magnitudes of the proton binding parameters produce only minor changes in the character of the surface. The calculated curves clearly reproduce the observed behavior of apo-Mb: at low temperatures and pH > 6, the native state is the predominant species; the population of I is greatest for pH values around 4; the denatured state is the most stable one at high temperatures and pH < 3. Figure 13A shows that the population of I decreases *gradually* with increasing temperature. The behavior of the functions at pH 4 is particularly interesting and important. Even though there is a large change in the population of I with an increase in temperature, for the same values of pH and temperature the heat capacity surface is nearly flat, as shown in Figure 13B. At pH > 4, a "strong" heat effect is observed, and the transition curves are close to those obtained using the two-state approximation. Below pH 4, the heat absorption peaks are no longer visible, even though the protein is not completely denatured. This behavior is identical to that observed experimentally.<sup>68</sup> Since experimentally derived estimates of the populations have not yet appeared in the literature, we will not attempt to refine the energetic parameter values.

To summarize, state I unfolds *gradually*: this process is not accompanied by a measurable peak in the heat capacity function. Also, at low temperatures the higher end of the cold denaturation peaks can be observed. Our analysis indicate that the calculated structural thermodynamic parameter values plus the NMR determined  $pK_a$  values can account for the observed stability behavior of apo-Mb within the context of a three-state model.

Recently, Barrick and Baldwin determined the pH-urea stability surface of apo-Mb and showed that it also can be rationalized in terms of a three state model (personal communication). Addition of urea

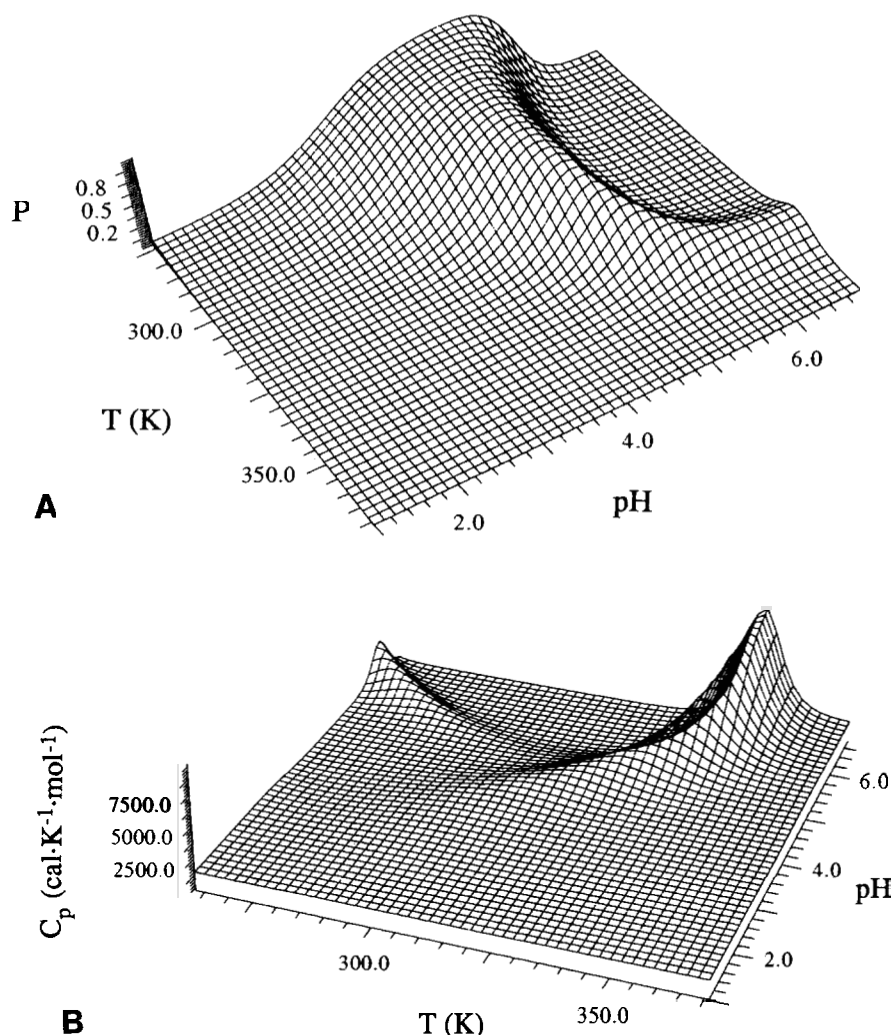


Fig. 13. Calculated stability surfaces for apo-Mb as a function of temperature and pH. (A) Population of I. (B) Excess heat capacity. Energetic parameter values used in the calculations were obtained from a structural thermodynamic analysis of the model proposed by Hughson et al. (ref. 58 and above); all other parameters (excluding  $\Delta S_i$ ) were determined experimentally (see text). Here,  $\Delta H_i(25^\circ\text{C}) = 8 \text{ kcal mol}^{-1}$ ,  $\Delta C_{p,i} = 850 \text{ cal K}^{-1} \text{ mol}^{-1}$ ,  $\Delta S_i(25^\circ\text{C}) = 11.5 \text{ cal K}^{-1} \text{ mol}^{-1}$ ;  $\Delta H_D(25^\circ\text{C}) = 28 \text{ kcal mol}^{-1}$ ,  $\Delta C_{p,D} = 1,550 \text{ cal K}^{-1} \text{ mol}^{-1}$ ,  $\Delta S_D(25^\circ\text{C}) = 63.5 \text{ cal K}^{-1} \text{ mol}^{-1}$ . The  $pK_a$ s of the histidine residues are from ref. 104. For the sim-

ulation seven histidyl groups were assumed to have anomalous  $pK_a$ s in the native state (4.5, 5.2, 5.2, 5.3, 2.5, 2.5, and 2.5). Two of the groups with  $pK_a$ s around 2.5 were assumed to remain anomalous in the I state. The  $pK_a$ s of the exposed histidyl groups were assumed to be 6. As discussed in the text, small variations in these values do not significantly affect the main features of the stability surfaces. The necessary condition to generate two pH-induced inflection points is the existence of a set of titrable groups with  $pK_a$ s around 5 and another one with  $pK_a$ s around 2.5–3.

always results in destabilization of the partly folded intermediate. This observation also points up a fundamental difference between the acid-pH and denaturant mechanisms of intermediate stabilization.

#### Stabilization of molten globule state by other factors

Recently it has become evident that intermediate states of different proteins can be stabilized by other factors. High concentrations (60% v/v) of methanol, for example, stabilize an intermediate state of ubiquitin.<sup>98</sup> At the time of this writing, however, little information is available regarding the intrinsic en-

ergetics of these states or the energetics of the interaction of these agents with proteins; therefore, no attempt will be made here to simulate their behavior.

#### CONCLUSIONS

The molten globule state is believed to represent a universal class of intermediate in protein folding; as such, knowledge of the energetics of this state is of paramount importance for understanding the nature of the folding pathway. Though much more work needs to be done, we have presented in this review a thermodynamic framework that accounts

for the observed behavior of partly folded states of different proteins under a variety of experimental conditions. A survey of experimental structural data indicates that molten globule states of different proteins exhibit similar features. These common characteristics have been used to provide a minimal structural definition of the molten globule state. There are three predominant features: (1) a relatively high content of secondary structure, (2) secondary structural elements which do not establish tertiary structure contacts satisfying the rigid geometric constraints of the native state; and (3) a sizable hydrophobic core.

The analysis demonstrated that structural differences among the native state, the molten globule state, and the denatured state are reflected in the energetics of these states. In particular, the molten globule and denatured states of several proteins were found to be energetically distinct: the enthalpy, entropy, and heat capacity differences between these states can not be zero. The absence of thermal denaturation transitions in some molten globule states appears to be more the result of the specific energetic balance for each case than a general property of the molten globule state. That is, the assumption that the molten globule state to denatured state transition is a second order phase transition is apparently a superfluous one; the behavior of molten globule states of different proteins can be explained without resorting to such a conjecture. Finally, a particular degree of cooperativity in unfolding is neither necessary nor sufficient for molten globule stabilization. The calculations presented here suggest that the stabilization of molten globule structures can be rationalized with the same statistical thermodynamic framework used to describe the stabilization of the native structure.

## ACKNOWLEDGMENTS

The authors thank Kenneth Murphy, Vincent Hilser, and Bertrand Garcia-Moreno for comments on the manuscript and helpful discussions. Robert Baldwin allowed us to examine apo-Mb data from his lab prior to its publication; for this we are grateful. Also, we thank A. Mark and W.F. van Gunsteren, who provided the atomic coordinates for the partly folded state of lysozyme. Supported by the National Institutes of Health (RR04328, GM37911, and NS24520) and the National Science Foundation (MCB-9118687).

## REFERENCES

1. Anfinsen, C.B., Scheraga, H.A. Experimental and theoretical aspects of protein folding. *Adv. Protein Chem.* 29: 205–300, 1975.
2. Levinthal, C. Are there pathways for protein folding? *J. Chim. Phys.* 65:44–45, 1968.
3. Wetlaufer, D.B. Nucleation, rapid folding, and globular intrachain regions in proteins. *Proc. Nat. Acad. Sci. U.S.A.* 70:697–701, 1973.
4. Kim, P.S., Baldwin, R.L. Intermediates in the folding reactions of small proteins. *Annu. Rev. Biochem.* 59:631–660, 1990.
5. Creighton, T.E. Characterizing intermediates in protein folding. *Curr. Biol.* 1:8–10, 1991.
6. Dobson, C.M. Characterization of protein folding intermediates. *Curr. Opin. Struct. Biol.* 1:22–27, 1991.
7. Roder, H., Baldwin, R.L. Characterizing protein folding intermediates. *Curr. Biol.* 1:218–220, 1991.
8. Dobson, C.M. Unfolded proteins, compact states and molten globules. *Curr. Opin. Struct. Biol.* 2:6–12, 1992.
9. Kim, P.S., Baldwin, R.L. Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding. *Annu. Rev. Biochem.* 51:459–489, 1982.
10. Freire, E., Biltonen, R.I. Statistical mechanical deconvolution of thermal transitions in macromolecules. I. Theory and application to homogenous systems. *Biopolymers* 17:463–479, 1978.
11. Privalov, P.L. Stability of proteins. Small globular proteins. *Adv. Protein Chem.* 33:167–241, 1979.
12. Chothia, C. Principles that determine the structure of proteins. *Annu. Rev. Biochem.* 53:537–572, 1984.
13. Kuwajima, K. A folding model of  $\alpha$ -lactalbumin deduced from the three-state denaturation mechanism. *J. Mol. Biol.* 114:241–258, 1977.
14. Dolgikh, D.A., Gilmanshin, R.I., Brazhnikov, V.E., Bychkova, G.V., Semisotnov, G.V., Venyaminov, S. Yu., Ptitsyn, O.B.  $\alpha$ -lactalbumin: Compact state with fluctuating tertiary structure? *FEBS Lett.* 136:311–315, 1981.
15. Brazhnikov, E.V., Chirgadze, Yu.N., Dolgikh, D.A., Ptitsyn, O.B. Noncooperative temperature melting of a globular protein without specific tertiary structure: Acid form of bovine carbonic anhydrase B. *Biopolymers* 24: 1899–1907, 1985.
16. Ohgushi, M., Wada, A. 'Molten-globule state': A compact form of globular proteins with mobile side chains. *FEBS Lett.* 164:21–24, 1983.
17. Brems, D.N., Plaisted, S.M., Havel, H.A., Kauffman, E.W., Stodola, J.D., Eaton, L.C., White, R.D. Equilibrium denaturation of pituitary- and recombinant-derived bovine growth hormone. *Biochemistry* 24:7662–7668, 1985.
18. Goto, Y., Fink, A.L. Conformational states of  $\beta$ -lactamase: Molten globule states at acidic and alkaline pH with high salt. *Biochemistry* 28:945–952, 1989.
19. Matthews, C.R., Crisanti, M.M. Urea-induced unfolding of the  $\alpha$  subunit of tryptophan synthase: Evidence for a multistate process. *Biochemistry* 20:784–792, 1981.
20. Silva, J.L., Silveira, C.F., Correia, A., Jr., Pontes, L. Dissociation of a native dimer to a molten globule monomer. Effects of pressure and dilution on the association equilibrium of Arc repressor. *J. Mol. Biol.* 223:545–555, 1992.
21. Dryden, D., Weir, M.P. Evidence for an acid-induced molten-globule state in interleukin-2; a fluorescence and circular dichroism study. *Biochim. Biophys. Acta* 1078:94–100, 1991.
22. Hua, Q.-X., Kochoyan, M., Weiss, M.A. Structure and dynamics of des-pentapeptide-insulin in solution: The molten globule hypothesis. *Proc. Natl. Acad. Sci. U.S.A.* 89: 2379–2383, 1992.
23. Tanford, C. Protein denaturation. *Adv. Protein Chem.* 23: 121–282, 1968.
24. Ptitsyn, O.B. Protein folding: Hypotheses and experiments. *J. Protein Chem.* 6:273–293, 1987.
25. Kuwajima, K. The molten globule state as a clue for understanding the folding and cooperativity of globular-protein structure. *Proteins: Struct., Funct., Genet.* 6:87–103, 1989.
26. Christensen, H., Pain, R.H. Molten globule intermediates and protein folding. *Eur. Biophys. J.* 19:221–229, 1991.
27. Denton, J.B., Konishi, Y., Scheraga, H. Folding of ribonuclease A from a partially disordered conformation. Kinetic study under folding conditions. *Biochemistry* 21: 5155–5163, 1982.
28. States, D.J., Dobson, C.M., Karplus, M., Creighton, T.E. Conformations of intermediates in the folding of the pancreatic trypsin inhibitor. *J. Mol. Biol.* 195:731–739, 1987.
29. Brems, D.N., Havel, H.A. Folding of bovine growth hormone is consistent with the molten globule hypothesis. *Proteins: Struct., Funct., Genet.* 5:93–95, 1989.
30. Goldberg, M.E., Semisotnov, G.V., Friguet, B., Kuwajima, K., Ptitsyn, O.B., Sugai, S. An early folding inter-



- mediate of the tryptophan synthase  $\beta_2$  subunit is a 'molten globule.' *FEBS Lett.* 263:51–56, 1990.
31. Dunker, A.K., Ensign, L.D., Arnold, G.E., Roberts, L.M. Proposed molten globule intermediates in fd phage penetration and assembly. *FEBS Lett.* 292:275–278, 1991.
  32. Palleros, D.R., Reid, K.L., McCarty, J.S., Walker, G.C., Fink, A.L. DnaK, hsp73, and their molten globules. Two different ways heat shock proteins respond to heat. *J. Biol. Chem.* 267:5279–5285, 1992.
  33. Cleland, J.L., Randolph, T.W. Mechanism of polyethylene glycol interaction with the molten globule folding intermediate of bovine carbonic anhydrase B. *J. Biol. Chem.* 267:3147–3153, 1992.
  34. Martin, J., Langer, T., Boteva, R., Schramel, A., Horwich, A.L., Hartl, F.-U. Chaperonin-mediated protein folding at the surface of groEL through a 'molten globule'-like intermediate. *Nature (London)* 352:36–42, 1991.
  35. Gething, M.-J., Sambrook, J. Protein folding in the cell. *Nature (London)* 355:33–45, 1992.
  36. van der Goot, F.G., González-Mañas, J.M., Lakey, J.H., Pattus, F. A 'molten-globule' membrane-insertion intermediate of the pore-forming domain of colicin A. *Nature (London)* 354:408–410, 1991.
  37. Ptitsyn, O.B., Pain, R.H., Semisotnov, G.V., Zervovnik, E., Razgulyaev, O.I. Evidence for a molten globule state as a general intermediate in protein folding. *FEBS Lett.* 262: 20–24, 1990.
  38. Creighton, T.E. Experimental studies of protein folding and unfolding. *Prog. Biophys. Mol. Biol.* 33:231–297, 1978.
  39. Ghelis, C. Transient conformational states in proteins followed by differential labeling. *Biophys. J.* 32:503–514, 1980.
  40. Kim, P.S. Amide proton exchange as a probe of protein folding pathways. *Methods Enzymol.* 131:136–156, 1986.
  41. Roder, H. Structural characterization of protein folding intermediates by proton magnetic resonance and hydrogen exchange. *Methods Enzymol.* 176:446–473, 1989.
  42. McKenzie, H.A., White, F.H., Jr. Lysozyme and  $\alpha$ -lactalbumin: structure, function, and interrelationships. *Adv. Protein Chem.* 41:173–315, 1991.
  43. Hiraoka, Y., Segawa, T., Kuwajima, K., Sugai, S., Murai, N.  $\alpha$ -Lactalbumin: A calcium metalloprotein. *Biochem. Biophys. Res. Commun.* 95:1098–1104, 1980.
  44. Acharya, K.R., Stuart, D.I., Walker, N.P.C., Lewis, M., Phillips, D.C. Refined structure of baboon  $\alpha$ -lactalbumin at 1.7 Å resolution. Comparison with c-type lysozyme. *J. Mol. Biol.* 208:99–127, 1989.
  45. Segawa, T., Sugai, S. Interactions of divalent metal ions with bovine, human, and goat  $\alpha$ -lactalbumins. *J. Biochem. (Tokyo)* 93:1321–1328, 1983.
  46. Hiraoka, Y., Sugai, S. Equilibrium and kinetic study of sodium- and potassium-induced conformational changes of apo- $\alpha$ -lactalbumin. *Int. J. Peptide Protein Res.* 26:252–261, 1985.
  47. Kronman, M.J. Metal-ion binding and the molecular conformational properties of  $\alpha$  lactalbumin. *Crit. Rev. Biochem. Mol. Biol.* 24:565–667, 1989.
  48. Kuwajima, K., Nitta, K., Yoneyama, M., Sugai, S. Three-state denaturation of  $\alpha$ -lactalbumin by guanidine hydrochloride. *J. Mol. Biol.* 106:359–373, 1976.
  49. Baum, J., Dobson, C.M., Evans, P.A., Hanley, C. Characterization of a partly folded protein by NMR methods: Studies of the molten globule state of guinea-pig  $\alpha$ -lactalbumin. *Biochemistry* 28:7–13, 1989.
  50. Robbins, F.M., Holmes, L.G. Circular dichroism spectra of  $\alpha$ -lactalbumin. *Biochim. Biophys. Acta* 221:234–240, 1970.
  51. Gilmanshin, R.I., Dolgikh, D.A., Ptitsyn, O.B., Finkelshtein, A.V., Shakhnovich, Ye.I. Protein globules without a unique spatial structure: Experimental data for  $\alpha$ -lactalbumins and a general model. *Biophysics (Poland)* 27: 1052–1064, 1982.
  52. Dolgikh, D.A., Abaturov, L.V., Bolotina, I.A., Brazhnikov, E.V., Bushuev, V.N., Bychkova, V.E., Gilmanshin, R.I., Lebedev, Yu.O., Semisotnov, G.V., Tiktupulo, E.I., Ptitsyn, O.B. Compact state of a protein molecule with pronounced small-scale mobility: Bovine  $\alpha$ -lactalbumin. *Eur. J. Biophys.* 13:109–121, 1985.
  53. Bierzynski, A., Kim, P.S., Baldwin, R.L. A salt bridge stabilizes the helix formed by isolated C-peptide of RNase A. *Proc. Natl. Acad. Sci. U.S.A.* 79:2470–2474, 1982.
  54. Kronman, M.J., Holmes, L.G. Inter- and intramolecular interactions of  $\alpha$ -lactalbumin. IV. Location of tryptophan groups. *Biochemistry* 4:526–532, 1965.
  55. Miranker, A., Radford, S.E., Karplus, M., Dobson, C.M. Demonstration by NMR of folding domains in lysozyme. *Nature (London)* 349:633–636, 1991.
  56. Kronman, M.J., Andreotti, R., Vitols, R. Inter- and intramolecular interactions of  $\alpha$ -lactalbumin. II. Aggregation reactions at acid pH. *Biochemistry* 3:1152–1160, 1964.
  57. Damaschun, G., Gernat, Ch., Damaschun, H., Bychkova, V.E., Ptitsyn, O.B. Comparison of intramolecular packing of a protein in native and 'molten globule' states. *Int. J. Biol. Macromol.* 8:226–230, 1986.
  58. Hughson, F.M., Wright, P.E., Baldwin, R.L. Structural characterization of a partly folded apomyoglobin intermediate. *Science* 249:1544–1548, 1990.
  59. Ewbank, J.J., Creighton, T.E. The molten globule protein conformation probed by disulphide bonds. *Nature (London)* 350:518–520, 1991.
  60. Meyer, T.E., Kamen, M.D. New perspectives on c-type cytochromes. *Adv. Protein Chem.* 35:105–212, 1982.
  61. Mathews, F.S. The structure, function and evolution of cytochromes. *Prog. Biophys. Mol. Biol.* 45:1–56, 1985.
  62. Goto, Y., Takahashi, N., Fink, A.L. Mechanism of acid-induced folding of proteins. *Biochemistry* 29:3480–3488, 1990.
  63. Goto, Y., Nishikiori, S. Role of electrostatic repulsion in the acidic molten globule of cytochrome c. *J. Mol. Biol.* 222:679–686, 1991.
  64. Kuroda, Y., Kidokoro, S.-i., Wada, A. Thermodynamic characterization of cytochrome c at low pH. Observation of the molten globule state and of the cold denaturation process. *J. Mol. Biol.* 223:1139–1153, 1992.
  65. Jeng, M.-F., Englander, S.W. Stable submolecular folding units in non-compact form of cytochrome c. *J. Mol. Biol.* 221:1045–1061, 1991.
  66. Jeng, M.-F., Englander, S.W., Elöve, G.A., Wand, A.J., Roder, H. Structural description of acid-denatured cytochrome c by hydrogen exchange and 2D NMR. *Biochemistry* 29:10433–10437, 1990.
  67. Stryer, L. "Oxygen-transporting proteins: Myoglobin and hemoglobin." In: "Biochemistry," 3rd ed. New York: Freeman, 1988: 143–174.
  68. Griko, Yu.V., Privalov, P.L., Venyaminov, S.Yu., Kutysheiko, V.P. Thermodynamic study of the apomyoglobin structure. *J. Mol. Biol.* 202:127–138, 1988.
  69. Hughson, F.M., Barrick, D., Baldwin, R.L. Probing the stability of a partly folded apomyoglobin intermediate by site-directed mutagenesis. *Biochemistry* 30:4113–4118, 1991.
  70. Privalov, P.L., Tiktupulo, E.I., Venyaminov, S.Yu., Griko, Yu.V., Makhataidze, G.I., Khechinashvili, N.N. Heat capacity and conformation of proteins in the denatured state. *J. Mol. Biol.* 205:737–750, 1989.
  71. Schellman, J. The thermodynamic stability of proteins. *Annu. Rev. Biophys. Chem.* 16:115–137, 1987.
  72. Freire, E., Murphy, K.P. The molecular basis of cooperativity in protein folding. *J. Mol. Biol.* 222, 687–698, 1991.
  73. Freire, E., Murphy, K.P., Sanchez-Ruiz, J.M., Galisteo, M., Privalov, P.L. The molecular basis of cooperativity in protein folding. Thermodynamic dissection of interdomain interactions in phosphoglycerate kinase. *Biochemistry* 31:250–256, 1992.
  74. Murphy, K.P., Bhakuni, V., Xie, D., Freire, E. Molecular basis of cooperativity in protein folding. III. Structural identification of cooperative folding units and folding intermediates. *J. Mol. Biol.* 227:293–306, 1992.
  75. Murphy, K.P., Freire, E. Thermodynamics of structural stability and cooperative folding behavior in proteins. *Adv. Prot. Chem.*, in press.
  76. Kuwajima, K., Hiraoka, Y., Ikeguchi, M., Sugai, S. Comparison of the transient folding intermediates in lysozyme and  $\alpha$ -lactalbumin. *Biochemistry* 24:874–881, 1985.
  77. Robertson, A.D., Baldwin, R.L. Hydrogen exchange in

- thermally denatured ribonuclease A. *Biochemistry* 30: 9907–9914, 1991.
78. Labhardt, A. Secondary structure in ribonuclease. I. Equilibrium folding transitions seen by amide circular dichroism. *J. Mol. Biol.* 157:331–355, 1982.
79. Biringer, R.G., Fink, A.L. Methanol-stabilized intermediates in the thermal unfolding of ribonuclease A. Characterization by  $^1\text{H}$  nuclear magnetic resonance. *J. Mol. Biol.* 160:87–116, 1982.
80. Biringer, R.G., Fink, A.L. Intermediates in the refolding of ribonuclease at subzero temperatures. 3. Multiple folding pathways. *Biochemistry* 27:315–325, 1988.
81. Murphy, K.P., Gill, S.J. Group additivity thermodynamics for dissolution of solid cyclic dipeptides into water. *Thermochim. Acta* 172:11–20, 1990.
82. Murphy, K.P., Gill, S.J. Solid model compounds and the thermodynamics of protein unfolding. *J. Mol. Biol.* 222: 699–709, 1991.
83. Hiraoka, Y., Sugai, S. Thermodynamics of thermal unfolding of bovine apo- $\alpha$ -lactalbumin. *Int. J. Peptide Protein Res.* 23:535–542, 1984.
84. Pfeil, W., Sadowski, M.L. A scanning calorimetric study of bovine and human  $\alpha$ -lactalbumin. *Studia Biophys.* 109: 163–170, 1985.
85. Xie, D., Bhakuni, V., Freire, E. Calorimetric determination of the energetics of the molten globule intermediate in protein folding: apo- $\alpha$ -lactalbumin. *Biochemistry* 30: 10673–10678, 1991.
86. Chen, B.-I., Baase, W.A., Schellman, J.A. Low-temperature unfolding of a mutant of phage T4 lysozyme. 2. Kinetic investigations. *Biochemistry* 28:691–699, 1989.
87. Pfeil, W., Bychkova, V.E., Ptitsyn, O.B. Physical nature of the phase transition in globular proteins. Calorimetric study of human  $\alpha$ -lactalbumin. *FEBS Lett.* 198:287–291, 1986.
88. Scholtz, J.M., Marqusee, S., Baldwin, R.L., York, E.J., Stewart, J.M., Santoro, M., Bolen, D.W. Calorimetric determination of the enthalpy change for the  $\alpha$ -helix to coil transition of an alanine peptide in water. *Proc. Natl. Acad. Sci. U.S.A.* 88:2854–2858, 1991.
89. Privalov, P.L., Gill, S.L. Stability of protein structure and hydrophobic interaction. *Adv. Prot. Chem.* 39:191–117, 1988.
90. Straume, M., Freire, E. Two-dimensional differential scanning calorimetry (2D DSC): simultaneous resolution of intrinsic protein structural energetics and ligand binding interactions by global linkage analysis. *Anal. Biochem.* 203:259–268, 1992.
91. Freire, E. Statistical thermodynamic analysis of the heat capacity function associated with protein folding-unfolding transitions. *Comments Mol. Cell. Biophys.* 6:123–140, 1989.
92. Privalov, P.L., Khechinashvili, N.N. A thermodynamic approach to the problem of stabilization of globular protein structure: A calorimetric study. *J. Mol. Biol.* 86:665–684, 1974.
93. Chen, B.-I., Schellman, J.A. Low-temperature unfolding of a mutant of phage T4 lysozyme. 1. Equilibrium studies. *Biochemistry* 28:685–691, 1989.
94. Bhakuni, V., Xie, D., Freire, E. Thermodynamic identification of stable folding intermediates in the B-subunit of cholera toxin. *Biochemistry* 30:5055–5060, 1991.
95. Wyman, J., Gill, S.J. The binding polynomial (nonassociating macromolecules). In: "Binding and Linkage. Functional Chemistry of Biological Macromolecules." Mill Valley, CA: University Science Books, 1990: 63–121.
96. Creamer, T.P., Rose, G.D. Side-chain entropy opposes  $\alpha$ -helix formation but rationalizes experimentally determined helix-forming propensities. *Proc. Natl. Acad. Sci. U.S.A.* 89:5937–5941, 1992.
97. Okuda, T., Sugai, S. Calorimetric study of the conformational transition of  $\alpha$ -lactalbumin induced by guanidine hydrochloride. *J. Biochem. (Tokyo)* 81:1051–1056, 1977.
98. Harding, M.M., Williams, D.H., Woolfson, D.N. Characterization of a partially denatured state of a protein by two-dimensional NMR: reduction of the hydrophobic interactions in ubiquitin. *Biochemistry* 30:3120–3128, 1991.
99. Ptitsyn, O. The molten globule state. In: "Protein Folding." Creighton, T.E., ed., New York: Freeman, 1992: 243–300.
100. Potekhin, S., Pfeil, W. Microcalorimetric studies of conformational transitions of ferricytochrome *c* in acidic solution. *Biophys. Chem.* 34:55–62, 1989.
101. Bychkova, V.E., Berni, R., Rossi, G.L., Kutysenko, V.P., Ptitsyn, O.B. Retinol-binding protein is in the molten globule state at low pH. *Biochemistry* 31:7566–7571, 1992.
102. Makhatadze, G.I., Privalov, P.L. Protein interactions with urea and guanidinium chloride. A calorimetric study. *J. Mol. Biol.* 226:491–505, 1992.
103. Ikeguchi, M., Kuwajima, K., Sugai, S.  $\text{Ca}^{2+}$ -induced alteration in the unfolding behavior of  $\alpha$ -lactalbumin. *J. Biochem. (Tokyo)* 99:1191–1201, 1986.
104. Cocco, M.J., Kao, Y.-H., Phillips, A.T., Lecomte, J.T.J. Structural comparison of apomyoglobin and metamyoglobin: pH titration of histidines by NMR spectroscopy. *Biochemistry* 31:6481–6491, 1992.
105. Hill, T. "An Introduction to Statistical Mechanics." Reading, MA: Addison-Wesley, 1960.
106. Makhatadze, G.I., Privalov, P.L. Heat capacity of proteins. I. Partial molar heat capacity of individual amino acid residues in aqueous solution: hydration effect. *J. Mol. Biol.* 213:375–384, 1990.
107. Spolar, R.S., Livingstone, J.R., Record, M.T., Jr. Use of liquid hydrocarbon and amide transfer data to estimate contribution to thermodynamic functions of protein folding from the removal of nonpolar and polar surface from water. *Biochemistry* 31:3947–3955, 1992.
108. Christensen, J.J., Hanks, R.W., Izatt, R.M. "Handbook of Heats of Mixing." New York: John Wiley, 1982.
109. Brandts, J.F. The thermodynamics of protein denaturation. I. The denaturation of chymotrypsin. *J. Am. Chem. Soc.* 86:4291–4301, 1964.
110. Kendrew, J.C., Dickerson, R.E., Strandberg, B.E., Hart, R.G., Davies, D.R., Phillips, D.C., Shore, V.C. Structure of myoglobin. A three-dimensional fourier synthesis at 2 Å resolution. *Nature (London)* 185:422–427, 1960.
111. Breslow, E., Gurd, F.R.N. Reactivity of sperm whale metmyoglobin towards hydrogen ions and p-nitrophenyl acetate. *J. Biol. Chem.* 237:371–381, 1962.
112. Tanford, C. Protein denaturation. Part C. Theoretical models for the mechanism of denaturation. *Adv. Prot. Chem.* 24:1–95, 1970.
113. Mark, A.E., van Gunsteren, W.F. Simulation of the thermal denaturation of hen egg white lysozyme: trapping the molten globule state. *Biochemistry* 31:7745–7748, 1992.
114. Yutani, K., Ogasahara, K., Kuwajima, K. The absence of the thermal transition in apo- $\alpha$ -lactalbumin in the molten globule state: A study by differential scanning microcalorimetry. *J. Mol. Biol.* 228:347–350, 1992.