Robson, K. J. H., Beattie, W., James, R. J., Cotton, R. C. H., Morgan, F. J., & Woo, S. L. C. (1984) *Biochemistry 23*, 5671-5673.
Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A., & Arnheim, H. (1985) *Science 230*, 1350-1354.
Saiki, R. K., Bugawan, T. L., Horn, G. T., Mullis, K. B., & Erlich, H. A. (1986) *Nature 324*, 163-166.
Scriver, C. R., & Clow, C. L. (1980a) *N. Engl. J. Med. 303*, 1336-1343.
Scriver, C. R., & Clow, C. L. (1980b) *N. Engl. J. Med. 303*, 1394-1440.
Smith, I., Lobascher, M. E., Stevenson, J. E., Wolff, O. H., Schmidt, H., Grubel-Kaiser, S., & Bickel, H. (1978) *Br. Med. J. 2*, 723-726.
Speer, A., Dahl, H. H., Reiss, D., Lober, G., Hanke, R., Cotton, R. G. H., & Coutelle, C. (1986) *Clin. Genet. 29*, 491.
Woo, S. L. C. (1984) *Pediatrics 74*, 412-423.
Woo, S. L. C. (1988) *Am. J. Hum. Genet. 43*, 781-783.
Woo, S. L. C., Lidsky, A. S. Güttler, F., Chandra, T., & Robson, K. J. H. (1983) *Nature 306*, 151-155.
Woo, S. L. C., Lidsky, A., Chandra, T., Güttler, F., & Robson, K. (1984) *JAMA, J. Am. Med. Assoc. 251*, 1998-2002.

Accelerated Publications

Characterization of a Partly Folded Protein by NMR Methods: Studies on the Molten Globule State of Guinea Pig α-Lactalbumin†

Jean Baum,‡ Christopher M. Dobson,* Philip A. Evans,§ and Claire Hanley
Inorganic Chemistry Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QR, England
Received August 15, 1988; Revised Manuscript Received October 24, 1988

ABSTRACT: NMR spectroscopy has been used to investigate the structure of a partially folded state of a protein, the molten globule or A-state of α-lactalbumin. The ^{1}H NMR spectrum of this species differs substantially from those of both the native and fully unfolded states, reflecting the intermediate level of order. The resolution in the spectrum is limited by the widespread overlap and substantial line widths of many of the resonances. Methods have therefore been developed that exploit the well-resolved spectrum of the native protein to probe indirectly the A-state. A number of resonances of the A-state have been found to be substantially shifted from their positions in the spectrum of the unfolded state and have been identified through magnetization transfer with the native state, under conditions where the two states are interconverting. The most strongly perturbed residues in the A-state were found to be among those that form a hydrophobic core to the native structure. A number of amides were found to be highly protected from solvent exchange in the A-state. These have been identified through pH-jump experiments, which label them in the spectrum of the native protein. They were found to occur mainly in segments that are helical in the native structure. These results enable a model of the A-state to be proposed in which significant conformational freedom exists but where specific elements of native-like structure are preserved.

The properties of partially organized states of proteins, in which only a subset of the native folding interactions may be present, are of fundamental importance in relation to our understanding of the nature of protein folding. The cooperativity of folding is such that these species often exist only transiently in the course of folding (Tanford, 1968; Kim & Baldwin, 1982), so that their detailed characterization may be feasible only if some form of trapping is possible (Creighton, 1978; Ghelis, 1980; Kim, 1986; Roder, 1988). Recently, however, it has been reported that several proteins can exist, under certain conditions, in partially structured states which are stable at equilibrium (Kuwajima, 1977; Dolgikh et al., 1981; Brazhnikov et al., 1985; Denton et al., 1981; Ohgushi & Wada, 1983; States et al., 1987). A number of such species have been characterized as "molten globule states", and these appear to be compact states exhibiting a high level of secondary structure (Dolgikh et al., 1981, 1985; Ohgushi & Wada, 1983). This description has been established from extensive experimental studies, but in order to understand fully the nature of these species, and particularly the manner in which they are related to native structures, it will be necessary to characterize the behavior of individual residues within a molten globule structure. We have been investigating the application of NMR spectroscopy to this problem in the specific case of the molten globule state of α-lactalbumin.

α-Lactalbumins are globular proteins of molecular weight around 14000 produced in the lactating mammary gland (Hill & Brew, 1975; Hall & Campbell, 1986). They show close sequence homology with c-type lysozymes, and a recent X-ray crystallographic study of baboon α-lactalbumin confirms that in the native state they are also closely similar in conformation (Stuart et al., 1986; Phillips et al., 1987). The α-lactalbumins

† This work was supported by the U.K. Science and Engineering Research Council. J.B. acknowledges receipt of a University of California President's Fellowship and a Fulford Junior Research Fellowship, Somerville College, Oxford. P.A.E. acknowledges an SERC Postdoctoral Fellowship and C.H. a Dee Graduate Scholarship, St. Hugh's College. C.M.D. is a member of the Oxford Centre for Molecular Sciences.
‡ Present address: Department of Chemistry, Rutgers University, New Brunswick, NJ 08903.
§ Present address: Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, England.

0006-2960/89/0428-0007$01.50/0

are, however, quite different from lysozymes in their biological role, in their tight binding of Ca^{2+} ions (Hiraoka et al., 1980), and in their folding properties. While lysozymes conform rather closely to the classical two-state model for cooperative folding (Tanford et al., 1966; Privalov & Khechinashivili, 1974; Dobson & Evans, 1984), it is well established that α-lactalbumins do not. A state resembling unfolded lysozyme is obtained for α-lactalbumins only in the presence of high concentrations of chemical denaturants (Kuwajima, 1977; Dolgikh et al., 1985). At lower denaturant concentrations, or on thermal or acid-induced unfolding, the molten globule state is generated in which considerable conformational order is retained (Dolgikh et al., 1985).

Several studies by ^1H NMR of α-lactalbumins in their native states have been reported (Bradbury & Norton, 1975; Berliner & Kaptein, 1981; Koga & Berliner, 1985; Kuwajima et al., 1986). The spectra are typical of globular proteins of this size and as such are potentially amenable to extensive assignment and analysis using methods which, for example, have recently permitted the virtually complete assignment of the spectra of hen and human lysozymes (Redfield & Dobson, 1988, and unpublished results). The spectrum of bovine α-lactalbumin in the molten globule state has also been reported (Dolgikh et al., 1985; Kuwajima et al., 1986). It differs substantially from the spectrum of the protein in its native state and is relatively poorly resolved. Conventional NMR approaches to assignment and structure determination rely on well-resolved spectra and on well-defined conformations; their direct applicability to studies of α-lactalbumin in the molten globule state is therefore likely to be restricted. We demonstrate in this paper, however, an approach whereby specific structural features of the molten globule state can be defined indirectly by methods that correlate resonances in the poorly resolved spectrum of the molten globule state with those in the well-resolved spectrum of the native state. The early results of this approach applied to guinea pig α-lactalbumin, which exploits and extends a strategy recently used to study the unfolded state of lysozyme (Wedin et al., 1982; Evans, 1986), are discussed in terms of their implications for our understanding of the molten globule state and its relationship to the native state.

Materials and Methods

Guinea pig α-lactalbumin (GPLA) was purified from whole milk obtained from Porcellus Animal Breeding, Sussex, by gel filtration of the whey on Sephadex G100 (Brew & Campbell, 1967), followed by ion exchange on DEAE-Bio-Gel A against a sodium chloride gradient and extensive dialysis. Deuterated urea was prepared by exchanging the labile protons of Analar urea (obtained from the BDH Chemical Co.) in D_2O and lyophilizing.

Samples were unbuffered and ranged in concentration from 1 to 3 mM for 1D NMR spectra and 5 mM for 2D spectra. For pH-jump hydrogen-exchange studies, the protein was initially dissolved in dilute solution at pH 2 in D_2O (35 mg/25 mL of D_2O) and left for various intervals during which time exchange took place from the A-state. The pH was then carefully raised to restore the native structure, and the sample was lyophilized. Finally, the protein was redissolved in 0.5 mL of D_2O and the pH adjusted as necessary prior to accumulation of spectra.

NMR experiments were performed on spectrometers of the Oxford Centre for Molecular Sciences. These were Bruker 500- and 300-MHz machines equipped with Aspect 3000 computers and a home-built 500-MHz spectrometer equipped with an Oxford Instruments Co. magnet, a GE/Nicolet 1280 computer, and a Bruker probe. Phase-sensitive NOESY experiments were carried out by using either the method of time-proportional phase increments (Marion & Wüthrich, 1983) or the method of States et al. (1982) with 10% random variation in the mixing time. Data sets collected by using the States method consisted of 256 t_1 increments of 4K data points, and those collected by using the TPPI method were of 512 t_1 increments of 2K data points. Various methods of resolution enhancement were used including trapezoidal multiplication and shifted sine bell; where data sets were to be directly compared, the same resolution enhancement was used in each. Zero filling was normally used to give a digital resolution of 3.5 Hz/point. All 2D spectra are presented as contour plots, and since both the native NOEs and the chemical-exchange peaks appear in the positive contour levels, only these are shown.

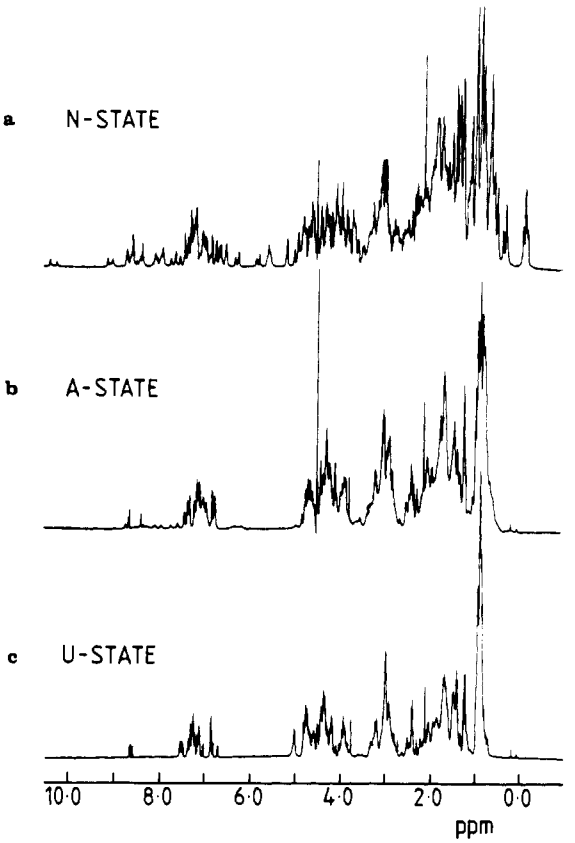

FIGURE 1: 500-MHz ^1H NMR spectra recorded at 52 °C of guinea pig α-lactalbumin at (a) pH 5.4 (the native state), (b) pH 2.0 (the molten globule state), and (c) pH 2.0 in the presence of 9 M urea (the unfolded state). In each case the protein was dissolved in ^2H_2O for 6 h at 20 °C prior to acquisition of the spectrum.

Results

Spectrum of the Molten Globule State. The ^1H NMR spectrum of guinea pig α-lactalbumin in its native (N) state is shown in Figure 1a. The wide dispersion of chemical shifts is characteristic of globular proteins and reflects the highly specific interresidue interactions within the compact folded structure. These features can be seen particularly clearly in the aromatic region of the spectrum, which is shown in Figure 2a, where the chemical shifts range from approximately 5.5 to 8.0 ppm. The spectra shown in Figures 1c and 2c, by contrast, are typical of the unfolded state of a protein and are characterized by only marginal deviation from the spectrum predicted for an ideal random coil (Bundi & Wüthrich, 1979). This state was obtained by exposure of the α-lactalbumin to 9 M urea at pH 2 and will be referred to as the U-state.

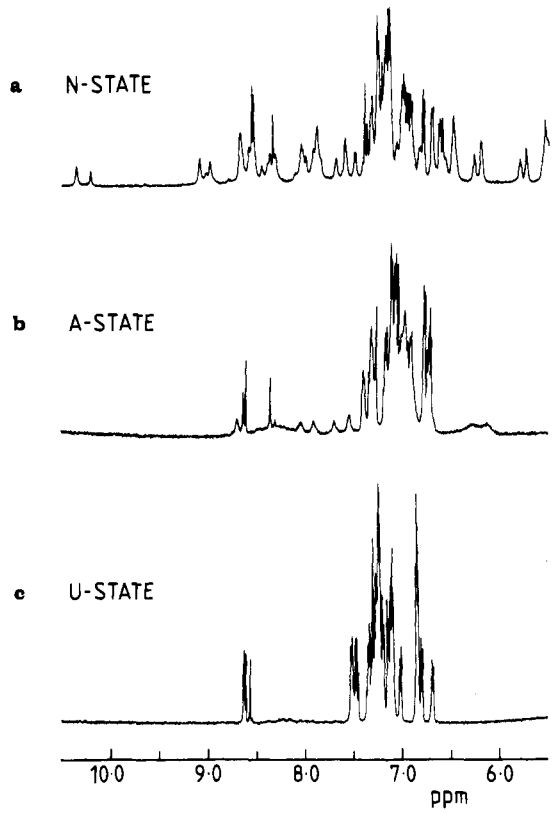

FIGURE 2: Expansion of the low-field regions of the spectra from Figure 1, showing resonances from the protons of aromatic residues and from amide protons that have not exchanged with deuterons from the ^2H_2O solvent.

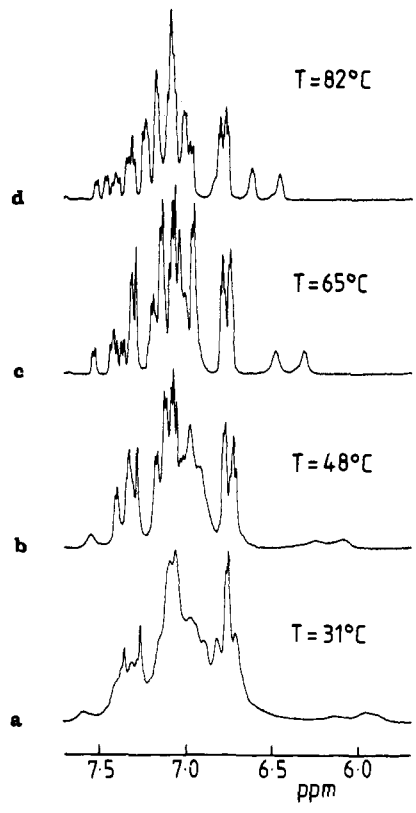

FIGURE 3: The aromatic region of the 500-MHz ^1H NMR spectrum of the A-state of α-lactalbumin (pH 2.0 in ^2H_2O) at (a) 31 °C, (b) 48 °C, (c) 65 °C, and (d) 82 °C. The amide protons were completely exchanged for deuterons prior to acquisition by equilibration for several minutes at the highest temperature.

In the absence of urea at pH 2, the spectrum observed is very different from those of either the N- or U-states (Figures 1b and 2b). Under these conditions the molten globule state, which we will refer to as the A-state, exists. The chemical shifts of many resonances in the A-state spectrum are close to those in the U-state spectrum but others are markedly shifted. In addition, many peaks are much broader than those in either the N- or U-state spectra. These features are particularly clearly seen for the pair of resonances at 6.15 and 6.3 ppm in Figure 2b. When the temperature is increased, these two resonances are observed to shift progressively to lower field and also ultimately to sharpen up, revealing them to be a pair of two proton intensity doublets (Figure 3). These resonances are mutually coupled and are readily assigned to the ring protons of a single tyrosine residue. Similar temperature-dependent effects are observed for other perturbed resonances, and the spectrum of the A-state progressively approaches that of the U-state at higher temperatures. It is interesting to note, however, that even at temperatures as high as 75 °C the chemical shifts of many resonances are still significantly perturbed relative to random coil values. Similar behavior is observed on titration of the A-state with the chemical denaturant urea. Here, however, the resonances have chemical shifts close to random coil values at high concentrations of denaturant, at which point the protein is said to be in the U-state. An important feature of both the temperature and urea dependence is that the changes in the spectra are gradual. No abrupt transition, characteristic of a cooperative transition (Tanford, 1968), is observed.

The broadening of many individual resonances of the A-state is particularly evident in the aromatic region of the spectrum, as seen in Figure 2. Broadening of resonances in protein NMR spectra is generally associated either with chemical exchange processes or with restricted rotational motion. Spectra of the A-state at a variety of temperatures were obtained both at 500 and 300 MHz and the line widths compared. The resolved tyrosine doublets are convenient to observe, and it was found that at temperatures up to about 65 °C the lines were markedly broader at 500 MHz than at 300 MHz; at higher temperatures the line widths became narrower and more closely comparable at both fields. Similar behavior was evident for other resonances, although severe overlap makes it difficult to measure the line widths of individual peaks. Such field-dependent broadening is typical of intermediate exchange processes and would be consistent with interconversion between a number of different conformations of the molten globule at a rate slower than about 10^3 s^{-1}. The temperature dependence was found to be complex, however, and the nature of such interconverting conformers cannot yet be defined. The possibility that slower overall tumbling of the A-state also contributes to the broadening has not been eliminated. Measurements of the radius of gyration suggest, however, that this tumbling would probably not be significantly slower for the A-state than for the native protein (Dolgikh et al., 1985; Gast et al., 1986), indicating that this is unlikely to be a major contribution to the observed line widths. Molecular association has been ruled out as a possible cause of the line broadening by two methods. First, the line widths were found to be independent of concentration over the range 0.02–2 mM. Second, equilibrium sedimentation measurements for α-lactalbumins from several species at protein concentrations used in the NMR experiments confirm that the major species present is a monomer (Dolgikh et al., 1985; C. Hanley, unpublished results).

A further interesting feature of the spectra in Figure 2 is the contrast between patterns of hydrogen exchange. Each of these spectra was obtained after the protein had been dis-

solved in D_2O for 6 h under the conditions specified. From these, together with 2D phase-sensitive COSY spectra, the number of remaining amide hydrogens that have not been substituted by deuterons can be counted. In the N-state there are 20–25 amides, and in the A-state there are 8–10; the U-state, obtained by adding urea at pH 2, has no remaining amide hydrogens after this time. The resonances in the A-state persist for several days at room temperature, and some of those of the native state persist for longer still. The hydrogen-exchange behavior of the A-state therefore appears in this sense intermediate between that of the N- and U-states.

The character of the chemical shift dispersion, broad lines, and hydrogen-exchange behavior demonstrates the existence of some level of residual folded structure in the A-state of α-lactalbumin. The shifts of the resonances of specific protons away from those observed in the U-state are likely to result from close interactions with other residues, particularly aromatic residues. The very slowly exchanging amide hydrogens show the presence of elements of structure that greatly reduce solvent accessibility. The disappearance of these characteristic effects upon addition of concentrated denaturant further confirms the notion that the molten globule includes elements of persistent folded structure.

The nature of the spectrum of the A-state affords little prospect of resonance assignment by conventional methods. Alternative strategies for detailed study of this state via the well-resolved and dispersed resonances of the native state are therefore essential. In the following sections two approaches devised to achieve this are described.

Magnetization-Transfer Experiments. Magnetization-transfer methods can correlate resonances of different states where these coexist at equilibrium and interconvert at rates comparable to, or faster than, the nuclear spin–lattice relaxation rates (Forsen & Hoffman, 1963; Jeener et al., 1979; Campbell et al., 1978). Reversible thermal unfolding of α-lactalbumin occurs above approximately 65 °C at pH 7.4. The unfolded state under these conditions gives a spectrum which is qualitatively similar to that of the A-state at pH 2, although the perturbations from the U-state are somewhat less pronounced than those observed at pH 2 and lower temperatures. Results from other spectroscopic techniques have also demonstrated the similarity of these states (Dolgikh et al., 1985). We can therefore conclude that the thermally unfolded state has a molten globule structure akin to that of the species present at pH 2, and therefore we subsequently refer to this form also as the A-state. At lower pH the midpoint of unfolding occurs at lower temperatures; for example, at pH 3.5 it is close to 40 °C. The perturbations attributable to molten globule structure are greater at this lower temperature, but magnetization-transfer studies are easier at pH 7.4 and higher temperature where the interconversion rate is higher and the A-state lines are less broad. Our present experiments have therefore been carried out at the higher pH.

Figure 4 shows the aromatic region of the spectrum of α-lactalbumin at pH 7.4 and 65 °C, close to the midpoint of the unfolding transition. The line widths of several resonances are markedly greater than found either above or below this temperature. The broadness is attributed to intermediate exchange effects associated with the interconversion of the N- and A-states; the rate is estimated from the exchange broadening to be approximately 7 s^{-1}. The 2D chemical-exchange spectrum obtained with a mixing time of 40 ms is also shown in Figure 4. Both NOE and exchange cross peaks can occur in such an experiment (Jeener et al., 1979), but here they were readily distinguished from one another by performing the same experiment at 55 °C, just below the transition zone of unfolding, at which temperature the protein is fully in its native state and only NOE peaks are expected to occur.

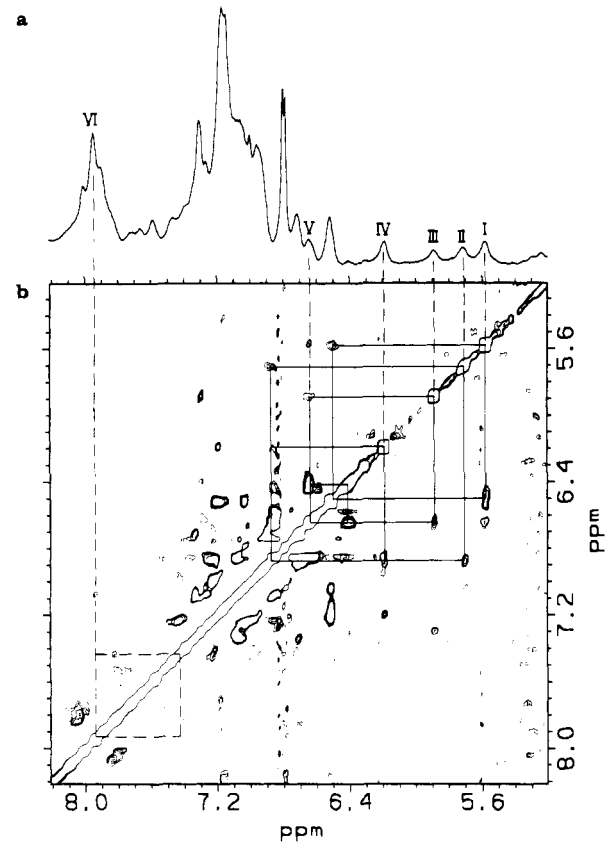

FIGURE 4: (a) Low-field region of the 500-MHz ^1H NMR spectrum of α-lactalbumin at pH 7.4 in H_2O, recorded at 68 °C. This temperature is close to the midpoint of the transition between the N- and A-states of the protein. Most of the resonances downfield of 7.4 ppm arise from the amide protons. (b) The corresponding region of the two-dimensional exchange correlated spectrum obtained with a mixing time of 40 ms. The cross peaks indicated arise from aromatic protons tentatively assigned to (I) Y103 C(2,6)H, (II) W26 C6H, (III) W26 C5H, (IV) F31 C(2,6)H, (V) Y103 C(3,5)H, and (VI) W104 C4H. The native state resonances are indicated above the 1D spectrum. In the case of proton VI the cross peak was visible only at lower contour level.

Despite the rather broad lines resulting from the exchange effects, we have been able to use this experiment to correlate the strongly perturbed resonances evident in the aromatic region of the spectrum of the A-state with those of the N-state. The exchange cross peaks observed most clearly were those from aromatic proton resonances which in the N-state have been identified from 2D COSY experiments as belonging to the protons of a tyrosine, a phenylalanine, and two tryptophan residues; see Figure 4. No crystal structure of GPLA is available, but nuclear Overhauser experiments conducted on the N-state have been analyzed by using a model based on the crystal coordinates of hen lysozyme (Handoll, 1985) and baboon α-lactalbumin (Stuart et al., 1986), both of which are presumed to be closely similar, to predict proton proximities in the guinea pig protein. These studies are consistent with the assignment of these four aromatic residues to Tyr 103, Phe 31, Trp 26, and Trp 104; sequential assignment studies are presently in progress to confirm these assignments. An important observation is that these resonances, with the largest deviations from random coil shifts in the A-state, correlate well with those having the largest shifts in the N-state. This strongly implies similarity between at least some features of

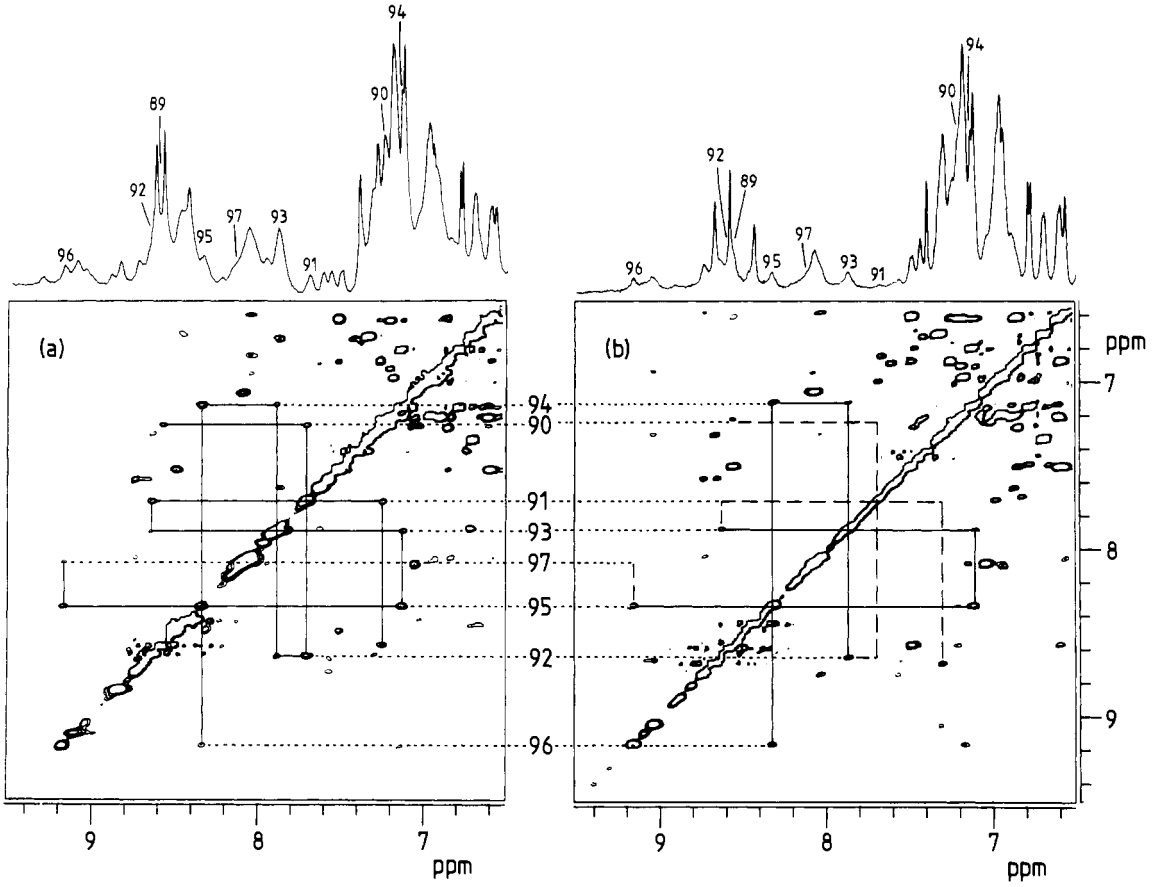

FIGURE 5: Low-field regions of the NOESY spectra of α-lactalbumin at 35 °C freshly dissolved (a) in ^2H_2O at pH 5.4 and (b) in ^2H_2O at pH 5.4 following the pH-jump experiment. The sequential connectivities between residues 90 and 97 are shown in both cases. The broken lines in (b) indicate connectivities from (a) that are not observed in the spectra following the pH jump; these are between residues 90–91, 91–92, and 96–97.

the structures of the N- and A-states. Specifically, the highly perturbed resonances that have been identified are all located in the "hydrophobic box" region in the native structure, suggesting that this clustering of aromatic residues persists in the molten globule state.

Hydrogen-Exchange Experiments. The spectrum shown in Figure 1b reveals that a group of amide protons remains unexchanged after exposure to D_2O at pH 2 for several hours. The very slow exchange of these amides with solvent indicates that in the A-state certain regions of the protein backbone are highly protected from solvent, which is itself a remarkable observation for a protein not in its globular state. It is of great interest to assign this slowly exchanging subset in the A-state but once again conventional methods cannot readily be used. Magnetization transfer, used to assign the chemically shifted aromatic resonances, also proved difficult for this purpose because hydrogen-exchange rates were too fast under the optimal conditions for this procedure. Instead, a method has been used whereby the slowly exchanging amides in the A-state are assigned by labeling them in the spectrum of the native protein.

The pH-jump hydrogen-exchange experiment used to achieve this consisted of two parts. First, the protein was dissolved in D_2O at pH 2 for 10 h, during which time hydrogen exchange took place. Second, the protein was restored to the native state by carefully raising the pH to 5.4. NMR experiments then revealed that the subset of amide protons that had not exchanged out at pH 2 could now be observed in the spectrum of the native protein. Moreover, these amides were found to be strongly resistant to further exchange from within the native structure. Thus, those amides that are protected from solvent exchange in the molten globule are all likewise protected in the folded state, providing further firm evidence that the two states are structurally closely related. This experiment thus offers the possibility of assigning this subset of amides in the N-state spectrum and so identifying, indirectly, regions where there is stable structure in the A-state.

Figure 5a shows the amide region of the NOESY spectrum of α-lactalbumin freshly dissolved in D_2O at pH 5.4. Prominent among the cross peaks in this region is the characteristic pattern of NH_i–NH_{i+1} correlations (Wüthrich, 1986) of a specific α-helical region, and this has been indicated in the figure. We have been able to assign these resonances by sequential methods (Wüthrich, 1986) to the helix encompassing residues 89–96; spin systems were analyzed via phase-sensitive COSY and single and double relayed coherence transfer spectroscopy in H_2O. The peptide sequence corresponding to the α-helix connectivities seen in Figure 5a was Ile-U-U-Val-U-U-Ile-Leu, where U indicates a residue as yet unidentified. The full helix, observed only in H_2O, contains the sequence Thr-U-U-Ile-U-U-Val-U-U-Ile-Leu-U. This sequence occurs uniquely in the protein in the segment comprising residues 86–97; the part of the helix seen in D_2O in Figure 5a is therefore assigned to residues 89–96.

After the pH-jump experiment, all but one of the NH_i–NH_{i+1} cross peaks in the 89–96 helix remain (see Figure 5b), showing that amides from this region have been protected from exchange in the A-state. The amide hydrogen of Cys 91 has exchanged in the A-state after 10 h: hence the two missing cross peaks at 7.7–7.25 and 7.7–8.65 ppm and the missing resonance at 7.7 ppm in the 1D spectrum. By contrast, in the spectrum of the freshly dissolved protein the intensity of Cys

91 NH appears comparable to the others in the helix. The peaks at 8.65–7.3 and 8.50–7.55 ppm are attributed to histidine C2–C4 couplings, and those at ca. 7.7–6.9 ppm are residual cross peaks of side-chain NH_2 protons of glutamine or asparagine. The remaining cross peaks in this region indicate that other short stretches of neighboring amides are also slowly exchanging in the A-state; for example, the residues with NH resonances at 8.72, 7.58, and 8.65 ppm are adjacent to one another. Finally, in COSY experiments performed after the pH jump all but one out of the seven cross peaks detectable in the fingerprint region were found to correspond to an amide that gives a cross peak in the NH_i–NH_{i+1} region of the NOESY spectrum. The implications of these observations are discussed below.

Discussion

The NMR spectrum of the A-state of α-lactalbumin has been shown to have overall properties quite different from those typical of either native or unfolded proteins. The existence of elements of folded structure within the molten globule is reflected in the deviations from random coil chemical shifts and in the slowly exchanging amide protons. The spread of chemical shifts is much less than in the native state, however, and many of the resonances are rather broad. Both of these features are indicative of fluctuations between conformations within the molten globule structure which are much greater than those occurring within the native protein. In this paper, we have described methods whereby the behavior of individual residues of the A-state can be investigated by taking advantage of the much more highly resolved spectrum of the native state. In particular, magnetization-transfer and pH-jump hydrogen-exchange experiments have been employed to identify the chemically shifted aromatic resonances as well as a very slowly exchanging subset of amide protons of the A-state.

The experiments described in this paper have been carried out with guinea pig α-lactalbumin while many of the studies using other techniques have used the more easily available bovine protein. Our studies have shown that the spectra from the bovine, human, and guinea pig α-lactalbumins in the A-state strongly resemble one another. The overall appearance of the spectra, notably the prominent tyrosine resonances at high field of the aromatic envelope, and the spectral changes observed as the temperature and the concentration of urea were varied, are all closely similar. Guinea pig α-lactalbumin was chosen as the species to be studied in detail because, under our conditions, hydrogen-exchange kinetics were observed to be slowest for the guinea pig protein, making it most suitable for the techniques employed. Further, work is in progress on the heterologous expression of the cloned gene for this particular protein, which should permit protein engineering studies to be performed in the future (P. A. Evans, unpublished results).

The hydrogen-exchange experiments, in which 8–10 amides were found to be slowly exchanging, show that certain regions of the polypeptide backbone are highly protected from solvent. These observations are of particular interest in the light of the results of circular dichroism and infrared experiments, from which a helical content of 30–40% in the A-state of α-lactalbumins has been estimated (Dolgikh et al., 1985). The implication is that, just as in native globular proteins, the secondary structure does not uniformly protect amide hydrogens from solvent exchange but rather that there are specific regions of particular stability which fluctuate rather little at room temperature, while other regions are more labile and afford only marginal protection. We have determined that all but one of the most slowly exchanging amide hydrogens in the A-state give rise to NH_i–NH_{i+1} cross peaks in the NOESY spectrum after the pH jump, which is consistent with their location in segments of the protein that are helical in the native state. One particular long stretch of helix has already been assigned to residues 89–96. The strong implication is that these elements of secondary structure are conserved in the A-state. In fact, the majority of NH_i–NH_{i+1} correlations in the spectrum of freshly dissolved α-lactalbumin in the native state are also visible after the pH jump, indicating a close relationship between the two states in terms of the stability of elements of helical structure. However, most of the other amides that are slowly exchanging in the N-state, identifiable both in 1D and COSY spectra, are not preserved after the pH jump. These are presumably located in nonhelical regions, and indeed we have already traced some to the main β-sheet of the protein. The clear qualitative difference in exchange behavior demonstrates that these parts of the molecule are much more labile in the molten globule state than in the native state.

Results from high-angle diffuse X-ray scattering experiments have suggested that residues are rather densely packed in the interior of the molten globule structure (Timchenko et al., 1986; Damaschun et al., 1986), but substantial averaging of the side-chain environment is nevertheless indicated by a sharp loss of intensity in the near-UV CD spectra compared to those of native proteins (Dolgikh et al., 1981, 1985). The NMR data are consistent with this view. The large deviations from random coil chemical shifts of many resonances in the spectrum of the N-state arise from the close proximity of protons to specific residues, particularly aromatic residues, in the folded structure. These shifts are strongly dependent on orientation as well as distance; their substantial diminution in the A-state spectrum indicates widespread averaging of interresidue interactions. It does not necessarily mean, however, that the structure is markedly less compact than that of the native state. Further, the extent of this averaging is not uniform, and some resonances show quite large deviations from random coil shifts. This shows that some interactions between individual residues are relatively specific and persistent.

Conclusion

The results presented in this paper have permitted a considerably more detailed picture of the molten globule state of a protein to be pieced together than has hitherto been possible. For α-lactalbumin we are able to conclude that there are specific elements of native-like secondary structure conserved in the molten globule state; in particular, one major helical region, including residues 89–96, has been identified from the hydrogen-exchange experiments. Second, the helical sections of the molten globule state, including the extended sequence 89–96, are extremely stable. This contrasts with the observations of helical segments in small peptides, where there appears to be generally rapid interconversion with a significant population of unstructured forms (Bierzynski et al., 1982; Dyson et al., 1988), but resembles conclusions drawn from the study of kinetic intermediates in protein refolding experiments (Roder et al., 1988; Udgaonkar & Baldwin, 1988). The helix in α-lactalbumin that contains residues 89–96 is amphipathic, and like all the helical segments in the protein it borders, in the native state, on the hydrophobic box region of the structure. The magnetization-transfer experiments reported here reveal that the most highly perturbed residues of the molten globule state, as judged from chemical shift values, include aromatic residues located in this region. We therefore conclude that this is a highly structured part of the molecule; the stability of the 89–96 helical structure may thus be a consequence of

a variety of side-chain interactions within this region of the protein. Even in this region of the molecule, however, there is considerably enhanced dynamic averaging of side-chain interactions relative to that found in the native state of the protein. By contrast to the hydrophobic core of the molecule, other regions of the molten globule state are evidently much less structured, including at least a major part of the β-sheet region found in the native structure. In particular, the small chemical shift dispersion of the majority of the resonances in the NMR spectrum of the molten globule state shows that there is considerable conformational disorder in much of the molecule. The molten globule state therefore appears intermediate in structure between the native and the fully unfolded state of the protein and, as such, should provide insight both into the factors stabilizing structural elements of globular proteins and perhaps into structural features likely to be found in intermediate states on protein folding pathways.

ACKNOWLEDGMENTS

We thank D. I. Stuart and K. R. Acharya for providing coordinates of the baboon α-lactalbumin crystal structure and for generating a model for the guinea pig protein structure. We also thank J. Heritage for assistance with the ultracentrifugation experiments. We are grateful to O. B. Ptitsyn and C. Redfield for useful discussions.

REFERENCES

Berliner, L. J., & Kaptein, R. (1981) *Biochemistry 20*, 799.
Bierzynski, A., Kim, P. S., & Baldwin, R. L. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 2470.
Bradbury, J. H., & Norton, R. S. (1975) *Eur. J. Biochem. 53*, 387.
Brazhnikov, E. V., Chigadze, D. A., Dolgikh, D. A., & Ptitsyn, O. B. (1985) *Biopolymers 24*, 1899.
Brew, K., & Campbell, P. N. (1967) *Biochem. J. 102*, 258.
Bundi, A., & Wüthrich, K. (1979) *Biopolymers 18*, 285.
Campbell, I. D., Dobson, C. M., Ratcliffe, R. G., & Williams, R. J. P. (1978) *J. Magn. Reson. 29*, 397.
Creighton, T. E. (1978) *Prog. Biophys. Mol. Biol. 33*, 231.
Damaschun, G., Gernet, C., Damaschun, H., Bychkova, V. E., & Ptitsyn, O. B. (1986) *Int. J. Biol. Macromol. 8*, 226.
Denton, J. B., Konishi, Y., & Scheraga, H. (1982) *Biochemistry 21*, 5155.
Dobson, C. M., & Evans, P. A. (1984) *Biochemistry 23*, 4267.
Dolgikh, D. A., Gilmanshin, R. I., Brazhnikov, E. V., Bychkova, V. E., Semisotnov, G. V., Venyaminov, S. Y., & Ptitsyn, O. B. (1981) *FEBS Lett. 136*, 311.
Dolgikh, D. A., Abaturov, L. V., Bolotina, I. A., Brazhnikov, E. V., Bychkova, V. E., Bushuev, V. N., Gilmanshin, R. I., Lebedev, Y. O., Semisotnov, G. V., Tiktopulo, E. I., & Ptitsyn, O. B. (1985) *Eur. Biophys. J. 13*, 109.
Dyson, H. J., Rance, M., Houghten, R. A., Wright, P. E., & Lerner, R. A. (1988) *J. Mol. Biol. 201*, 201.
Evans, P. A. (1986) D. Phil. Thesis, Oxford University.
Forsén, S., & Hoffman, R. A. (1963) *J. Chem. Phys. 39*, 2892.
Gast, K., Zirwer, D., Welfle, H., Bychkova, V. E., & Ptitsyn, O. B. (1986) *Int. J. Biol. Macromol. 8*, 231.
Ghelis, C. (1980) *Biophysical Discussions: Proteins and Nucleoproteins, Structure, Dynamics, & Assembly*, Rockefeller University Press, New York.
Hall, L., & Campbell, P. N. (1986) *Essays Biochem. 22*, 1.
Handoll, H. H. G. (1985) D. Phil. Thesis, Oxford University.
Hill, R. L., & Brew, K. (1975) *Adv. Enzymol. Relat. Areas Mol. Biol. 43*, 411.
Hiraoka, Y., Segawa, T., Kuwajima, K., Sugai, S., & Murai, N. (1980) *Biochem. Biophys. Res. Commun. 95*, 1098.
Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546.
Kim, P. S. (1986) *Methods Enzymol. 131*, 136.
Kim, P. S., & Baldwin, R. L. (1982) *Annu. Rev. Biochem. 51*, 549.
Koga, K., & Berliner, L. J. (1985) *Biochemistry 24*, 7257.
Kuwajima, K. (1977) *J. Mol. Biol. 114*, 241.
Kuwajima, K., Harushima, Y., & Sugai, S. (1986) *Int. J. Pept. Protein Res. 27*, 18.
Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967.
Ohgushi, M., & Wada, A. (1983) *FEBS Lett. 164*, 21.
Phillips, D. C., Acharya, K. R., Handoll, H. H. G., & Stuart, D. I. (1987) *Biochem. Soc. Trans. 15*, 737.
Privalov, P. L., & Khechinashivili, N. N. (1974) *J. Mol. Biol. 86*, 665.
Redfield, C., & Dobson, C. M. (1988) *Biochemistry 27*, 122.
Roder, H. (1988) *Methods Enzymol.* (in press).
Roder, H., Elove, G. A., & Englander, S. W. (1988) *Nature 335*, 700.
States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286.
States, D. J., Dobson, C. M., Karplus, M., & Creighton, T. E. (1987) *J. Mol. Biol. 195*, 731.
Stuart, D. I., Acharya, K. R., Walker, N. P. C., Smith, S. G., Lewis, M., & Phillips, D. C. (1986) *Nature 324*, 84.
Tanford, C. (1968) *Adv. Protein Chem. 23*, 121.
Tanford, C., Pain, R. H., & Otchin, N. S. (1966) *J. Mol. Biol. 15*, 489.
Timchenko, A. A., Dolgikh, D. A., Damaschun, H., & Damaschun, G. (1986) *Stud. Biophys. 112*, 201.
Udgaonkar, J. B., & Baldwin, R. L. (1988) *Nature 335*, 694.
Wedin, R. E., Delepierre, M., Dobson, C. M., & Poulsen, F. M. (1982) *Biochemistry 21*, 1098.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.