Characterization of the Molten Globule of Human Serum Retinol-Binding Protein Using NMR Spectroscopy[†,‡]

Lesley H. Greene,[§] Ramani Wijesinha-Bettoni, and Christina Redfield*

Department of Biochemistry and Chemistry Research Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QU, U.K.

Received February 3, 2006; Revised Manuscript Received May 31, 2006

ABSTRACT: The molten globule state is a partially folded conformer of proteins that has been the focus of intense study for more than two decades. This non-native fluctuating conformation has been linked to protein-folding intermediates, to biological function, and more recently to precursors in amyloid fibril formation. The molten globule state of human serum retinol-binding protein (RBP) has been postulated previously to be involved in the mechanism of ligand release (Ptitsyn, O. B., et al. (1993) *FEBS Lett. 317*, 181−184). Conserved residues within RBP have been identified and proposed to be key to folding and stability, although a link to a molten globule state has not previously been shown (Greene, L. H., et al. (2003) *FEBS Lett. 553*, 39−44). In this work, a detailed characterization of the acid-induced molten globule of RBP is presented. Using stopped-flow fluorescence spectroscopy in the presence of 8-anilino-1-naphthalene sulfonic acid (ANS), we show that RBP populates a state with molten-globule-like characteristics early in refolding. To gain insight into the structural features of the molten globule of RBP, we have monitored the denaturant-induced unfolding of this ensemble using NMR spectroscopy. The transition at the level of individual residues is significantly more cooperative than that found previously for the archetypal molten globule, α-lactalbumin (α-LA); this difference may be due to a predominantly *β*-sheet structure present in RBP in contrast to the α-helical nature of the α-LA molten globule.

In the 1980s the molten globule state was first observed by partially unfolding a protein under acidic conditions (*1*, *2*). The presence of a nativelike secondary structure and the lack of a well-defined tertiary structure, revealed using spectroscopic methods such as circular dichroism, led to the suggestion that this state might have some relationship to intermediates in protein folding. The ability to study a stable equilibrium form of a partially folded protein that bears a similarity to a transient kinetic intermediate drew a great deal of excitement and attention from protein chemists. Through the 1990s and up to the present, a significant amount of data has been accumulated that suggests that these equilibrium molten globule states do in fact have key features in common with protein-folding intermediates (*3*−*6*). The role of the molten globule has also been expanded to include physiological processes (*7*−*11*) and as precursors to amyloid fibril formation (*12*, *13*). The experimental and computational study of molten globules continues to be an active area of research within the field of protein science. Elucidating the structural features of these partially folded conformers provides important insights into the determinants of protein topology, the mechanism of protein folding, and biological function.

Obtaining residue-specific information for molten globule states can be best achieved using the power of NMR[1] spectroscopy (*14*). However, because of the severe line broadening arising from conformational fluctuations throughout the molten globule on a ms to *μ*s time scale, the use of NMR spectroscopy remains quite challenging. One approach to the elucidation of the structural characteristics of the molten globule involves the stepwise unfolding of this species using urea or guanidine HCl. This has enabled a detailed picture, at the level of individual residues, of the α-lactalbumin (α-LA) molten globule to emerge (*15*, *16*). Other successful avenues involving NMR spectroscopy include hydrogen−deuterium exchange studies to identify the hydrogen-bonded secondary structure (*17*, *18*), the use of chemical shift, short-range NOE, and ^{15}N relaxation data to identify regions of persistent secondary structure (*19*, *20*), raising the temperature to modify the complex dynamic properties of the molten globule (*19*, *21*), the use of peptide models to probe local and long-range interactions (*22*), comparative studies between species variants (*23*) and site-directed mutagenesis to identify persistent hydrophobic interactions (*24*), and photochemically induced dynamic nuclear polarization to monitor the solvent accessibility of

[†] This work was supported by the U.K. Biotechnology and Biological Sciences Research Council (Grant 43/B17889). L.H.G. thanks the National Science Foundation for an International Research Fellows Award (#0000597).

[‡] BMRB Accession Number 7149.

* To whom correspondence should be addressed. Tel: +44 (0)1865 275330. Fax: +44 (0)1865 275259. Email: christina.redfield@bioch.ox.ac.uk.

[§] Current address: University College London, Department of Biochemistry and Molecular Biology, Biomolecular Structure and Modelling Group, Gower Street, London WC1E 6BT, U.K.

[1] Abbreviations: NMR, nuclear magnetic resonance; α-LA, α-lactalbumin; RBP, retinol-binding protein; ANS, 8-anilino-1-naphthalene sulfonic acid; CD, circular dichroism; HSQC, heteronuclear single quantum coherence; *β*LG, *β*-lactoglobulin.

10.1021/bi060229c CCC: $33.50
Published on Web 07/13/2006

tyrosine, tryptophan, and histidine residues (*25*). These studies have provided important insights into the structural characteristics of the molten globule states of a variety of proteins.

The molten globule of the human serum retinol-binding protein (RBP), a member of the lipocalin superfamily (*26*), is of particular interest. The lipocalins are a very large, divergent, and functionally diverse group of proteins that share a common eight-stranded up-and-down β-barrel topology that ranks among the 20 most prevalent folds in the protein universe (Figure 1A) (*27*). Lipocalins are widespread throughout the eukaryotic and prokaryotic kingdoms, suggesting that they are an ancient line. A conserved sequence signature within the lipocalins has been postulated to be key to folding and native-state stability (*28*). Physiologically, RBP is a carrier protein that transports all-trans retinol (vitamin A) in complex with transthyretin within the blood stream to target cells (*29*, *30*). The potential interaction with a cell receptor is currently an area of active research (*31*). RBP also forms an acid-induced molten globule state that may be relevant to the mode of ligand release, although the detailed mechanism of retinol release remains unresolved (*32*, *33*). The partially folded form of RBP has recently been the subject of molecular dynamics simulations, which provided insights into the ensemble of structures sampled by the molten globule (*34*). In this article, we report a study aimed at characterizing the equilibrium molten globule state of RBP to provide a foundation to explore a potential functional link as well as to investigate the determinants of protein folding and topology. We provide evidence from ANS fluorescence that during kinetic refolding RBP appears to populate a state with molten-globule-like characteristics. NMR spectroscopy is employed to follow, at the level of individual residues, the urea-induced unfolding of the RBP molten globule. We find that the RBP molten globule unfolds in a more cooperative manner than the archetypal molten globule formed by α-lactalbumin (*15*); this difference may be due to a predominantly β-sheet structure present in RBP in contrast to the α-helical nature of the α-LA molten globule.

MATERIALS AND METHODS

Sample Preparation. Recombinant human serum RBP was expressed in *E. coli* as inclusion bodies, purified, and refolded as described previously (*35*). In this work, however, the protocol was modified to express the protein using M9 minimal medium supplemented either with ^{15}N ammonium chloride or with ^{15}N ammonium chloride and deuterium oxide. The ^{15}N ammonium chloride and deuterium oxide were purchased from CK Gas and Fluorochem, respectively. For expression using deuterium oxide, the *E. coli* BL21-(DE3) cells were grown over several days in increasing ratios of $^{2}H_2O/^{1}H_2O$ until growth in 100% $^{2}H_2O$ was obtained. Deuterated urea was prepared by dissolving high quality urea (Sigma Ultra) in $^{2}H_2O$ and then lyophilizing; this procedure was repeated three times.

Circular Dichroism, Fluorescence, and Stopped-Flow Studies. Unlabeled protein was used for the circular dichroism (CD) and fluorescence studies. The urea-induced unfolding of the RBP molten globule at pH 2 and 20 °C in 5 mM sodium phosphate buffer was monitored by far-UV CD using a Jasco J-720 spectropolarimeter; 600 points measured at 222 nm were averaged over 60 s for urea concentrations ranging from 0 to 8 M. The protein concentration was 0.5 mg/mL, and a cuvette with a 0.1 cm path length was used.

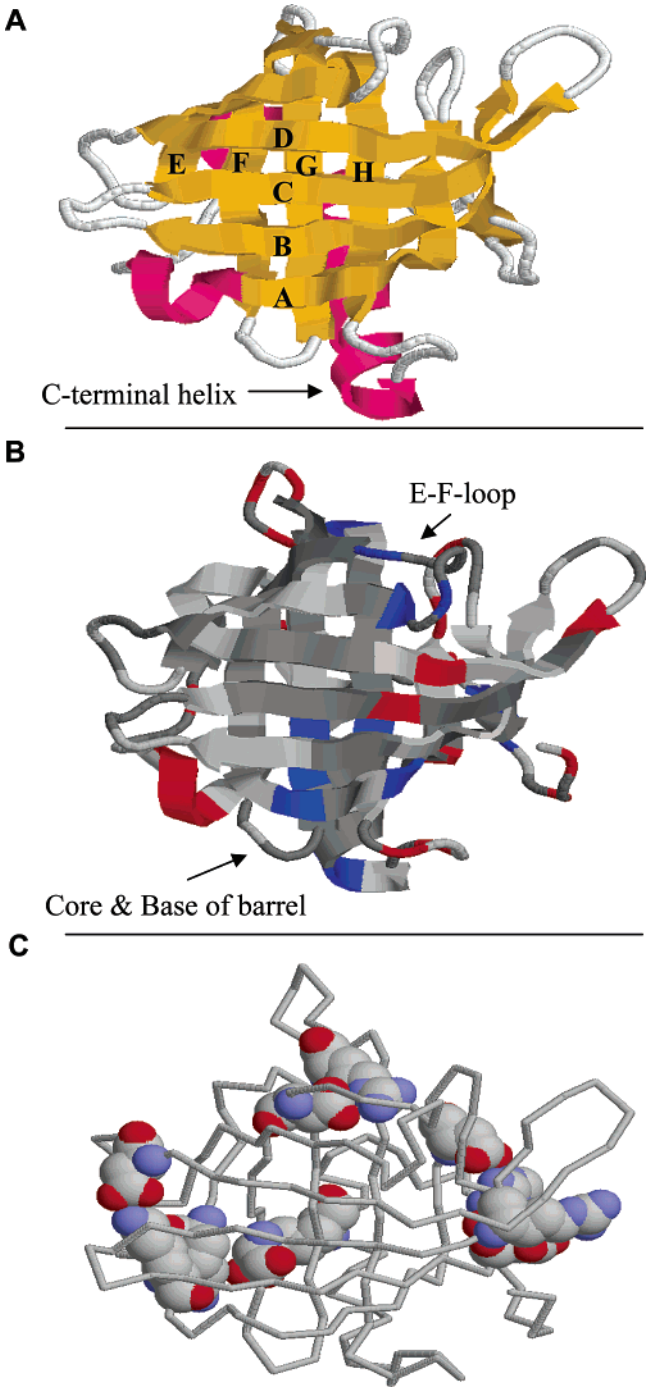

FIGURE 1: Schematic representations of the structure of apo human serum retinol-binding protein. (A) The 1.7 Å resolution crystal structure of native RBP (*35*). The eight β-strands are shown in yellow and the α-helices in magenta. (B) The residues are colored in red (≤4.7 M urea), blue (≥5.7 M urea), and dark gray (4.8−5.6 M urea) on the basis of the groupings of stability depicted in Figure 6B. The residues that could not be monitored because of spectral overlap are shown in light gray. (C) The residues involved in salt bridges that are likely to be disrupted at low pH are shown on the α-carbon backbone as light gray space fills; the red and blue coloring represents oxygen and nitrogen, respectively. The RBP structure (pdb code 1JYD) (*35*) is visualized using Rasmol (version 2.7.1).

The interaction of 8-anilino-1-naphthalene sulfonic acid (ANS) with RBP was monitored by fluorescence spectros-

copy at pH 2 and at pH 7 in 0.4 M and 4 M guanidine HCl. Fluorescence measurements were performed using a Perkin-Elmer Model LS-50B luminescence spectrometer using a cuvette with a 1 cm path length. ANS was purchased from Fluka. The protein and ANS concentrations were 2 μM. An excitation wavelength of 340 nm was used for these experiments, and the emission spectrum was measured between 400 and 620 nm with a 5 nm slit width. All spectra are the average of five scans collected at 20 °C.

Stopped-flow fluorescence spectroscopy was conducted to monitor the interaction of RBP with ANS during the refolding of RBP using an Applied Photophysics stopped-flow system. The protein (238 μM) was initially denatured in 4 M guanidine HCl at pH 7. Refolding was initiated by a 10-fold dilution; the refolding buffer contained 20 mM Bis-Tris (pH 7.0) and 24 μM ANS. The excitation wavelength was 340 nm with a 530 nm cutoff filter; although this cutoff wavelength (530 nm) is above the λ_{max} for ANS emission (~475 nm), a significant fluorescence signal from ANS is still observed. Slit widths were 2 nm, and the temperature was maintained at 20 °C. An average of nine shots was fitted to a double-exponential equation using SigmaPlot, version 8.0 (SSPS, Inc.).

NMR Studies. Resonance assignments for the 21 kDa protein in 8 M urea at pH 2 and 30 °C were obtained using the standard sequential assignment methodology. A lack of spectral resolution usually characterizes the spectra of unfolded proteins; the ^{1}H chemical shift dispersion of H^N is small (~1 ppm). However, the ^{15}N chemical shift dispersion is still fairly large because of the spread of random coil ^{15}N shifts. The resonance overlap in the ^{15}N dimension is also reduced by the sequence dependence of the random coil shifts, which leads to changes in the chemical shift of up to 5 ppm (*36*). The resonance assignments were obtained using a 3D gradient-enhanced ^{15}N-edited NOESY−HSQC spectrum (*37*, *38*) with a mixing time of 200 ms recorded for ^{15}N-labeled RBP in 8 M urea at pH 2 and 30 °C. This spectrum was collected at 750 MHz with 8 scans per increment using 128 complex t_1(^{1}H) values, 42 complex t_2-(^{15}N) values, and 512 complex points in the acquisition dimension, t_3 (^{1}H). Sweep widths of 8130.08, 6756.76, and 1440.9 Hz were used in F_3, F_1, and F_2, respectively. The assignments were confirmed using a 3D HSQC−NOESY−HSQC spectrum (*39*) obtained using ^{15}N-^{2}H labeled RBP. This experiment was collected at 600 MHz with 12 scans per increment using 55 complex t_1(^{15}N) values, 32 complex t_2(^{15}N) values, and 512 complex points in the acquisition dimension, t_3 (^{1}H). Sweep widths of 6493.5, 1201.9, and 1201.9 Hz were used in F_3, F_1, and F_2, respectively. The assignments obtained at 30 °C were transferred to lower temperatures using HSQC spectra collected at 30, 25, and 20 °C. $^1H^N$ and ^{15}N chemical shifts for RBP in 8 M urea at pH 2 and 20 °C have been deposited in the BMRB (accession number 7149).

Unfolding experiments in urea were performed at 750 MHz using 0.16 mM samples of ^{15}N,^{2}H-labeled RBP at pH 2 and 20 °C in 95% ^1H_2O/5% ^2H_2O. Spectra were collected at urea concentrations ranging from 0 to 10 M. The desired urea concentration was obtained by mixing appropriate volumes of stock solutions of 0.16 mM RBP in 0 M and 10 M urea containing 5 mM sodium phosphate at pH 2. This ensured that the concentration of protein remained constant throughout the titration. The urea concentrations were determined using a hand-held refractometer (Atago R5000). RBP has been shown previously to be largely monomeric at low ionic strength (5 mM sodium phosphate) at pH 2 (*40*). NMR diffusion measurements carried out using the pulse-gradient spin longitudinal echo diffusion sequence (*41*) show that at pH 2 RBP is ~15% expanded compared to native RBP at pH 7; this value is consistent with previous reports (*33*). Two-dimensional gradient-enhanced ^{15}N-^{1}H HSQC spectra (*38*, *42*) were recorded at 750 MHz for each urea concentration. These spectra consisted of 128 complex t_1 increments of 1 K complex data points. Twenty-four scans were collected for each t_1 increment. Sweep widths of 8130.08 and 1501.5 Hz were used in the F_2(^{1}H) and F_1(^{15}N) dimensions, respectively. All NMR spectra were processed using Felix 2.3 (Accelrys, San Diego, CA).

The apparent midpoint of the unfolding transition for each resolved residue was determined as follows. Peak heights were obtained for each resolved peak in the HSQC spectra at each urea concentration using an in-house peak-picking program; if a peak was not observed above the noise level, then the intensity was assigned as zero. For each residue, peak heights were scaled relative to the peak height measured in 10 M urea. Peak heights were also corrected for the broadening effect that results from the increased viscosity as the concentration of urea is increased. The peak heights were fitted using a standard sigmoidal function:

$$I_{obs} = I_{MG} + \frac{(I_U - I_{MG})}{(1 + \exp(-([\text{urea}]\text{-}midpt)/width))} \quad (1)$$

I_{obs} is the scaled and corrected peak height observed at each concentration of urea ([urea]). I_{MG} is the peak height in the molten globule state and is assumed in this study to be 0. I_U is the peak intensity in the fully unfolded state; this parameter is fitted and is found to be 0.95 ± 0.01. The apparent midpoint of unfolding, the urea concentration at which the peak has half its maximum intensity, is defined as *midpt*; this parameter is fitted and is found to range from 4.2 to 6.1 M urea with an average value of 5.1 ± 0.2 M urea for the 118 residues monitored. The range of urea concentrations in which the observed peak intensity varies significantly is defined as *width*; this parameter is fitted and is found to have an average value of 1.01 ± 0.03. Errors in the fitted apparent unfolding midpoints (*midpt*), arising from baseline noise in the HSQC spectra, were calculated using 500 Monte Carlo simulations; these errors ranged from 0.003 to 0.22 M urea with an average error of 0.02 M urea.

RESULTS

Circular Dichroism and Fluorescence Studies. The circular dichroism (CD) spectrum of RBP at pH 2 has been shown in previous studies to be characteristic of a partially folded molten globule state (*40*); the far-UV CD spectrum shows the presence of significant nativelike β-sheet secondary structure, whereas the lack of near-UV CD signal indicates the absence of fixed aromatic residues in an asymmetric environment. We have observed similar CD spectra for recombinant RBP at pH 2 (not shown). A comparison of the far-UV CD spectra at pH 2 and 7 indicates a 16% increase in helicity, a 5% decrease in β-structure, and an 11% decrease in random coil structure in the molten globule

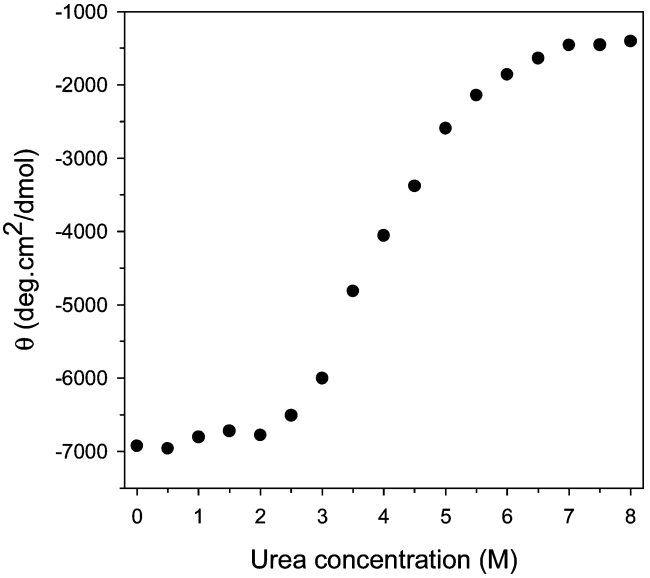

FIGURE 2: Urea-induced unfolding of the pH 2 molten globule state of RBP monitored by CD ellipticity at 222 nm and 20 °C.

compared to those of the native state (*35*). Molecular dynamics simulations of the partially folded form of RBP suggest that the additional helicity may arise from the conversion of the E−F β-hairpin to a helix (*34*). To probe the stability of the low pH molten globule, a urea titration monitored by far-UV CD was conducted (Figure 2). The observed ellipticity as a function of urea concentration shows a sigmoidal curve, although the baseline for the unfolded state is not as clearly delineated as that observed for the denaturation of apo-RBP at pH 7.0 (*35*). The unfolding of the molten globule occurs over a somewhat wider range of urea concentration (2.5−6.5 M with a midpoint at 3.9 M urea) than generally observed for native states. This may indicate a decrease in the coopertivity of unfolding of the RBP molten globule and is consistent with previous observations for other partially folded states including α-lactalbumin at pH 2 (*14*, *15*).

The molten globule of RBP has been postulated to be involved in the mechanism of release of its ligand, retinol (*32*, *33*). Protein-folding intermediates have also been shown to have key features in common with molten globules (*3*−*6*). A stopped-flow fluorescence refolding study of RBP in the presence of ANS has been conducted to identify whether molten-globule-like features exist along the refolding pathway of RBP. Partially folded proteins such as molten globules interact with the hydrophobic dye ANS because of the exposure of large numbers of hydrophobic residues on the protein surface (*43*); this can be monitored by the observation of enhanced fluorescence for ANS. RBP at pH 2 has been shown previously to interact strongly with ANS, whereas no interaction was observed under native or fully denaturing conditions (*40*). The refolding of RBP, unfolded in 4 M guanidine HCl at the start of the experiment, was initiated by a 10-fold dilution of RBP into a Bis-Tris buffer containing ANS at pH 7; the fluorescence of ANS was monitored above 530 nm as a function of the refolding time (Figure 3A). The observed refolding rates of RBP are similar to those obtained using stopped-flow intrinsic fluorescence and circular dichroism in previous studies (*28*). The fluorescence intensity of ANS shows a significant enhancement

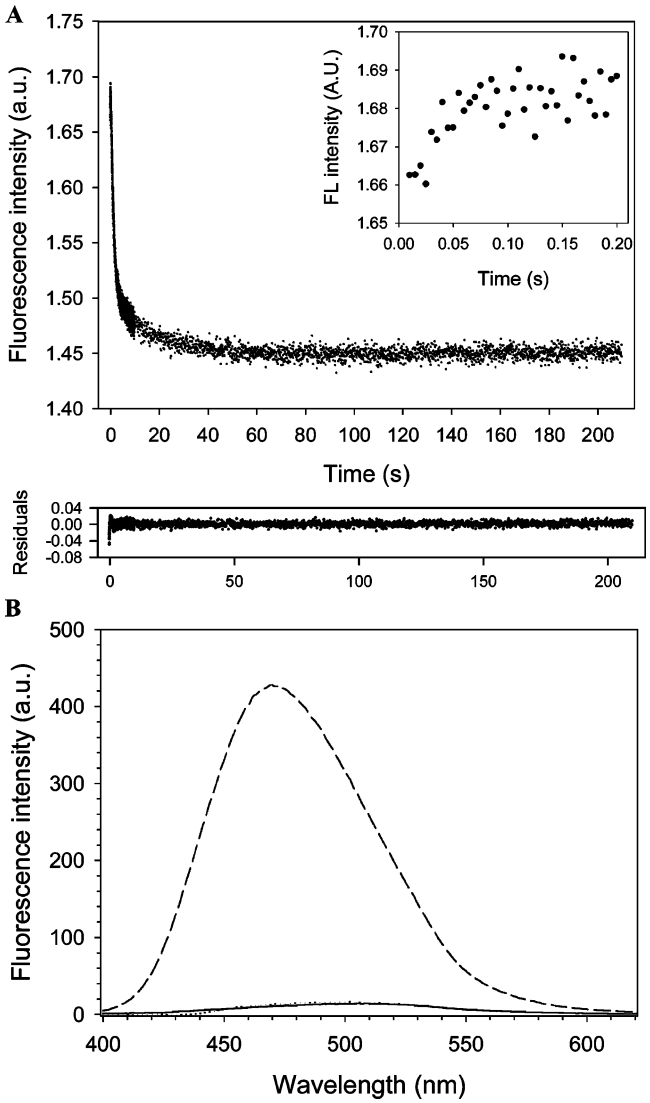

FIGURE 3: Interaction of ANS with RBP monitored by fluorescence spectroscopy. (A) The fluorescence intensity of ANS was monitored as a function of time following the initiation of refolding of RBP. The calculated amplitudes and rates from the dissociation of ANS are as follows: $A_1 = 0.217 \pm 0.001$, $k_1 = 0.731 \pm 0.006$ s^{-1}, $A_2 = 0.048 \pm 0.001$, and $k_2 = 0.056 \pm 0.002$ s^{-1}. The ANS concentration relative to protein concentration after dilution was in a 1:1 ratio. The decrease in the ANS fluorescence during refolding demonstrates the presence of a transient species with molten globule-like characteristics. The inset is a magnification of the data points up to 200 ms, although proper interpretation requires further detailed experiments. (B) The fluorescence spectra of ANS in the presence of native RBP at pH 7.0 in 0.4 M guanidine HCl (....), RBP unfolded in 4 M guanidine HCl at pH 2 (—), and RBP at pH 2 (---). All three samples contained 2 μM RBP and 2 μM ANS. The almost 20-fold enhancement of the ANS fluorescence in the presence of RBP at pH 2 is characteristic of the behavior of ANS in the presence of a molten globule. Both the kinetic and equilibrium studies were conducted at 20 °C.

early in the refolding of RBP (Figure 3A); this enhancement is characteristic of that observed for ANS in the presence of a transient molten globule (*6*, *43*−*47*). For comparison, the steady-state fluorescence of ANS in the presence of RBP at pH 7 in 4 M guanidine HCl, the starting point of the refolding experiment, and in 0.4 M guanidine HCl, the end point of the refolding experiment, are shown in Figure 3B; no enhancement of the fluorescence of ANS is observed under either of these conditions. Thus, we postulate that at the onset

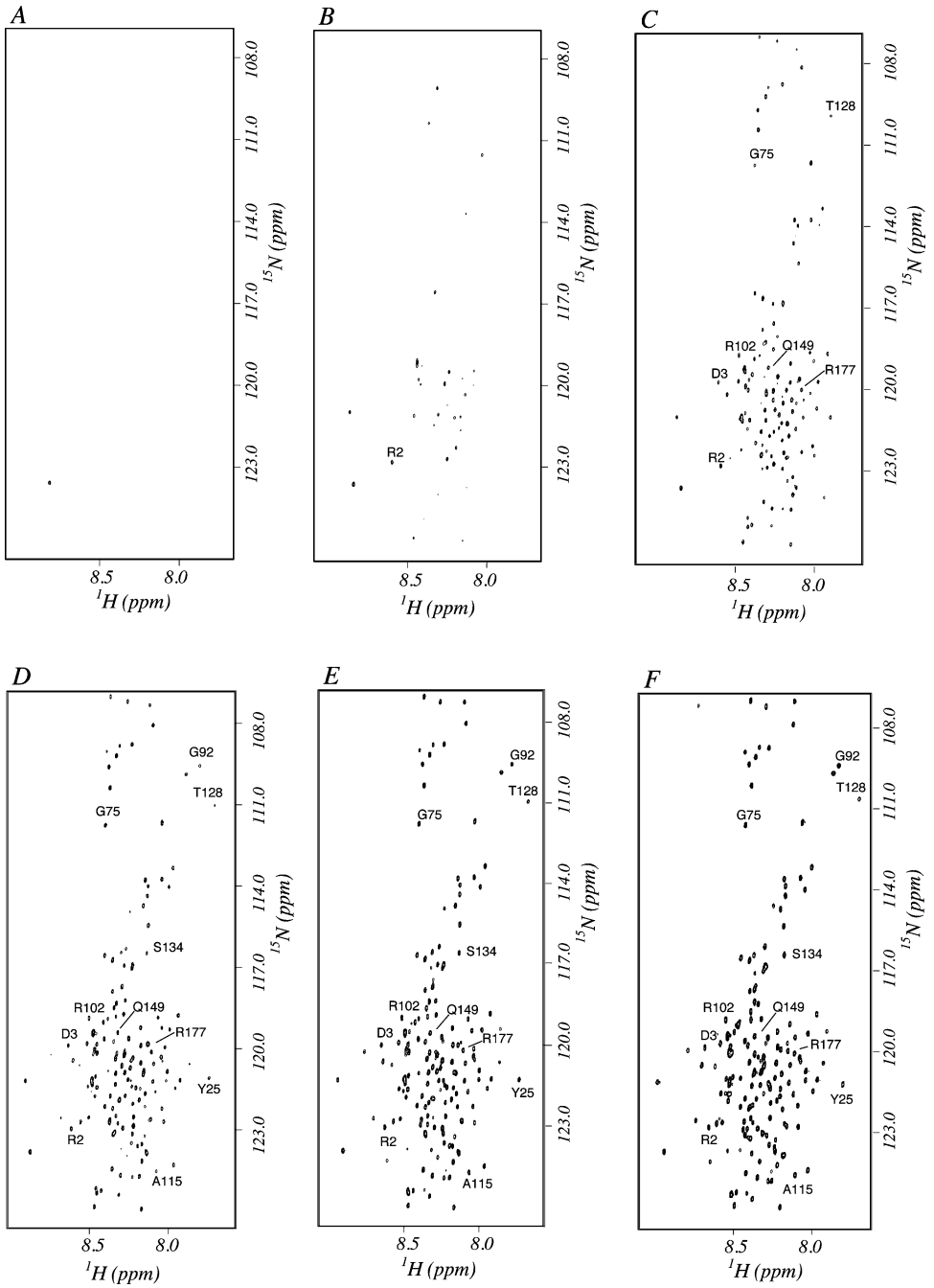

FIGURE 4: Progressive appearance of the unfolded NMR resonances as the RBP molten globule is unfolded. Shown are representative ^{15}N-^{1}H HSQC spectra from the urea titration of the RBP molten globule at pH 2 in (A) the absence of urea, (B) 3 M urea, (C) 4 M urea, (D) 5 M urea, (E) 6 M urea, and (F) 8 M urea, all at 20 °C. Residues whose unfolding behavior is plotted in Figure 6A are labeled.

of the refolding of RBP a transient species with molten globule-like characteristics is formed; these molten-globule-like characteristics disappear as the well-ordered, native hydrophobic core of RBP is formed. Although we cannot ascertain at this point whether the molten-globule-like species observed during refolding has detailed characteristics in common with the equilibrium molten globule of RBP at pH 2, this result suggests that a more detailed characterization of the pH 2 molten globule may provide insights into the folding pathway of RBP.

Urea Unfolding of Individual Residues Monitored by NMR Spectroscopy. To gain residue-specific information, we have used NMR spectroscopy to monitor the unfolding of the RBP molten globule in urea at pH 2. 2D HSQC spectra were collected for urea concentrations ranging from 0 to 10 M, and the appearance of cross-peaks from specific residues was analyzed; the choice of urea concentrations was based on the CD data shown in Figure 2. To improve the resolution of the crowded 2D HSQC spectra, ^{15}N,^{2}H-labeled RBP was used in the urea titration experiments. In the HSQC spectrum of the RBP molten globule formed at pH 2 and 20 °C, only one peak corresponding to E1 (which is preceded in recombinant RBP by M0) is visible (Figure 4A). This is typical of ^{15}N-^{1}H HSQC spectra of molten globules; the conformational heterogeneity and complex dynamical properties of the molten globule lead to extreme line broadening and the absence of peaks from the HSQC spectrum (*14*). The number of cross-peaks observed increases as urea is added; 32 peaks are observed in 3 M urea, 121 peaks are observed in 4 M urea, and 154 peaks are observed in 5 M

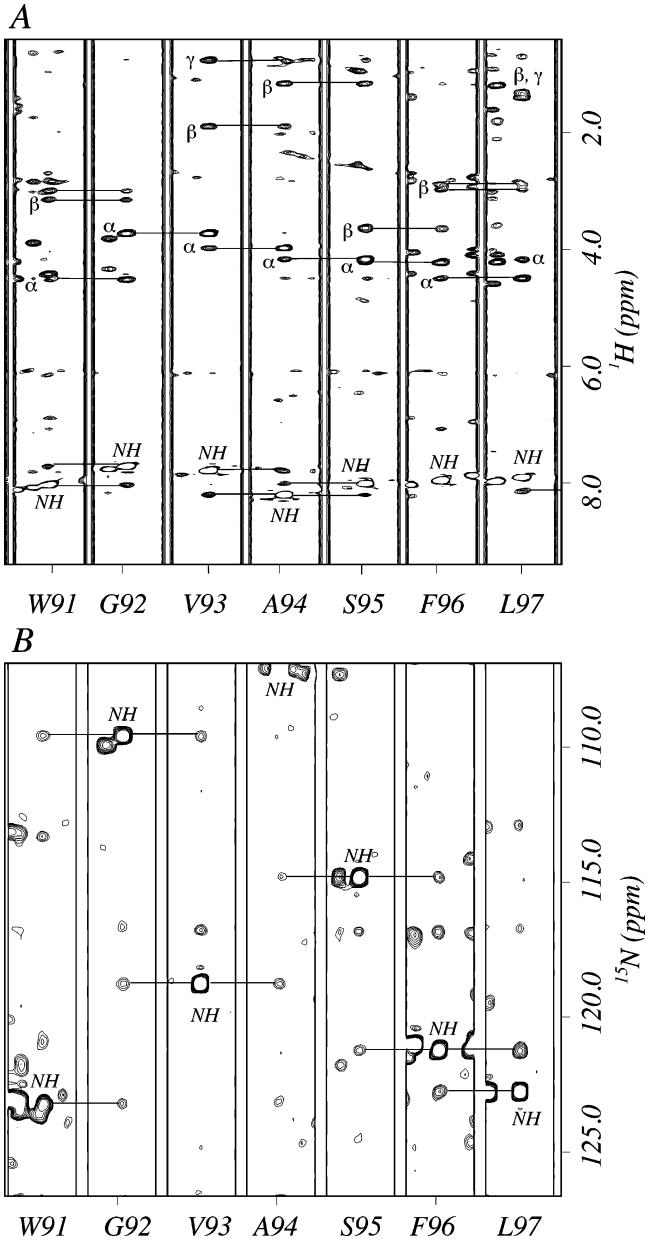

FIGURE 5: Strip plots showing 3D NOE data for RBP in 8 M urea at pH 2 and 30 °C. The strips correspond to residues 91−97 and are labeled with their assignments. Sequential NOE connectivities are shown by horizontal lines. (A) 3D ^{15}N-edited NOESY−HSQC spectrum. (B) 3D HSQC−NOESY−HSQC spectrum. The diagonal peaks are labeled NH.

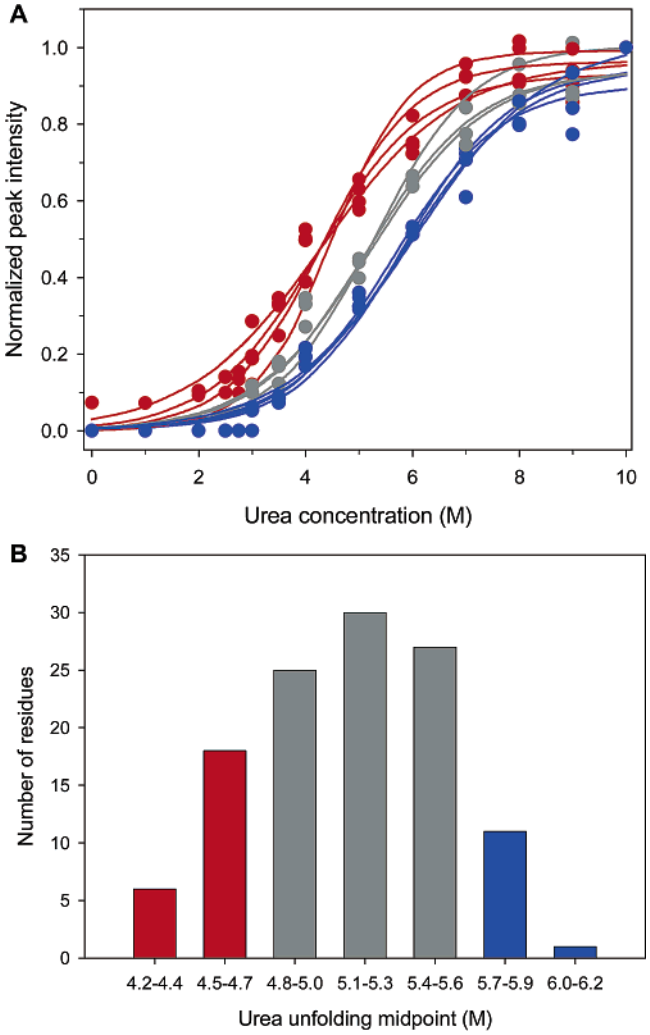

FIGURE 6: Representative urea titration curves for individual residues monitored by NMR spectroscopy. (A) The scaled and corrected peak height is plotted as a function of urea concentration for residues 2, 3, 25, 75, 92, 102, 115, 128, 134, 149, and 177. Experimental data are shown with filled circles and the fits are shown with solid lines. Residues 2, 3, 128, and 177, which show a below average apparent unfolding midpoint, are shown in red. Residues 25, 92, 115, and 134, which show an above average apparent midpoint are shown in blue. Residues 75, 102 and 149 which show close to average behavior are shown in dark gray. (B) Histogram showing the distribution of apparent unfolding midpoints for the 118 residues that could be monitored throughout the urea titration.

urea (Figure 4B−D). The peaks that are observed in the HSQC spectra in increasing concentrations of urea correspond to peaks from the unfolded form of RBP. The increase in peak intensity corresponds to an increase in the population of unfolded RBP with increasing urea concentration. From the spectrum shown in Figure 4F, it is clear that all of the expected peaks are observed in 8 M urea.

The spectrum of RBP unfolded in 8 M urea was assigned using ^{15}N-edited 3D NOESY−HSQC, and 3D HSQC−NOESY−HSQC experiments as described in Materials and Methods (Figure 5). Where overlap of the H^N chemical shifts of neighboring residues made assignment difficult (eg., S95-F96-L97 in Figure 5A), the HSQC−NOESY−HSQC spectrum (Figure 5B) helped in confirming the assignments (because the ^{15}N chemical shifts of the amides of S95, F96, and L97 are not the same). Assignments for 174 of the expected 176 main chain ^{15}N-$^1H^N$ HSQC peaks have been obtained.

The assignments obtained for RBP in 8 M urea can be used to assign peaks observed in the HSQC spectra collected at other urea concentrations. As a result, the unfolding behavior of the RBP molten globule can be monitored at the level of individual residues. Resonances for 118 of the 176 backbone amides in RBP could be followed clearly throughout the titration after the initial appearance of a peak at a particular urea concentration; these were, therefore, amenable to a more detailed analysis. To analyze the unfolding transitions quantitatively, the scaled and corrected HSQC peak heights were plotted as a function of urea concentration (Figure 6A), and the apparent midpoints of unfolding were calculated from these plots as described in

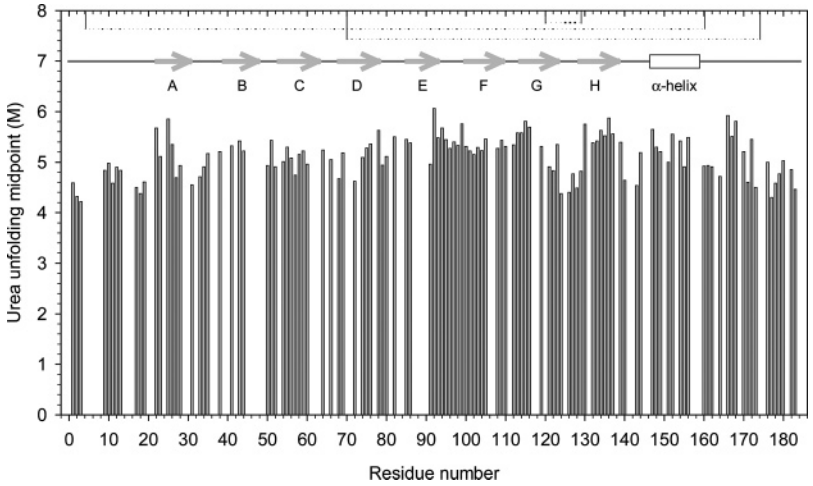

FIGURE 7: Apparent unfolding midpoints are plotted as a function of amino acid sequence. The secondary structure is shown schematically as well as the location of the three disulfide bonds (....).

Materials and Methods. The average apparent midpoint of unfolding is 5.1 M urea, and the range is limited between 4.2 and 6.1 M urea. It is interesting to note that the average apparent unfolding midpoint of 5.1 M urea is significantly higher than the unfolding midpoint of 3.9 M urea obtained from the CD data (Figure 2). A similar observation has been made previously for the unfolding of α-LA at pH 2 (*15*). This suggests that in both RBP and α-LA the loss of regular secondary structure may precede the complete unfolding of the polypeptide chain.

The errors in the HSQC peak heights were used to estimate errors in the apparent unfolding midpoints; these errors ranged from 0.003 to 0.22 M with an average value of 0.02 M. Thus, the range of observed unfolding midpoints, 4.2−6.1 M urea, is significantly larger than the estimated errors; this may indicate a degree of noncooperativity in the unfolding of the RBP molten globule. We have grouped the apparent midpoints into three stability groups (Figures 6B and 1B). A less stable group (apparent midpoint ≤4.7 M urea), a more stable group (apparent midpoint ≥5.7 M urea), and an intermediate stability group. In Figure 7, the calculated apparent unfolding midpoints are plotted as a function of sequence. Residues at the *N*-terminus of RBP, prior to β-strand A, and residues at the *C*-terminus, after residue 175, are found in the less stable group. In addition, residues in many of the loops, particularly the G−H loop, and residues directly preceding and following the *C*-terminal helix are found to be less stable. In contrast, the residues located in β-strand A, the second face of the barrel (β-strands G and H), the *C*-terminal helix, and the E−F loop are found in the more stable group (Figure 1B).

DISCUSSION

Human serum RBP undergoes a transition to a molten globule at pH 2 (*40*). Circular dichroism spectra indicate that there is a 16% increase in helicity, a 5% decrease in β-structure and an 11% decrease in random coil structure, and the overall loss of fixed tertiary structure at pH 2 (*35*). Stopped-flow ANS fluorescence studies suggest that RBP refolds through a transient species with molten-globule-like characteristics (Figure 3A). This observation is based on the increase in binding of the fluorescent dye ANS at the onset of refolding; ANS is well known to interact with the exposed hydrophobic surfaces of molten globules.

The urea-induced unfolding of individual residues in the RBP molten globule has been monitored by 2D HSQC NMR spectroscopy. The apparent unfolding midpoints for 118 residues have been calculated from the observed peak heights in HSQC spectra collected as a function of urea concentration. These apparent midpoints range from 4.2 to 6.1 M urea, indicating a small degree of noncooperativity in the unfolding behavior. There appears to be a small enhancement in the stability of β-strand A, the second face of the barrel (β-strands G and H), the *C*-terminal helix, and the E−F loop (Figure 1B). Interestingly, the first three of these structural regions contain the majority of evolutionarily conserved residues (*28*). The latter is a region that we propose is more stable because of the conversion from a loop to a helix at low pH; this may facilitate ligand release as well as dissociation from transthyretin. Molecular dynamics simulations of RBP support the proposal of a conversion of this region to a helix (*34*).

The kinetic refolding pathway and the molten globule state of β-lactoglobulin (βLG), another member of the lipocalin superfamily that shows extensive structural homology to RBP (*26*), have been studied in detail. The kinetic folding pathway of bovine βLG has been investigated using ultrarapid mixing techniques in conjunction with hydrogen-exchange labeling to identify the stable secondary structure formed early in folding (*48*). The F, G, and H β-strands and the *C*-terminal helix were found to form hydrogen bonds early in folding and were identified as the folding core in bovine βLG (*48*). A 27-residue peptide derived from the G−H β-strand region of the bovine βLG sequence shows nativelike hydrophobic interactions suggesting that the G and H strands may act as a folding initiation site in βLG (*49*). Hydrogen exchange studies for equine βLG at pH 1.5 show significant protection for residues in the A, F, G, and H β-strands and in the *C*-terminal helix in the molten globule that is formed at low pH (*50*). It is interesting to note that regions found to be most stable in the RBP molten globule, namely β-strands A, G, and H and the *C*-terminal helix, have also been found to be involved in important interactions in the folding core and molten globule of the homolgous protein βLG.

The unfolding of the RBP molten globule at pH 2 has been found to be significantly more cooperative than the unfolding of the archetypal molten globule formed by human α-lactalbumin at pH 2 (*15*). α-LA is a 14 kDa α + *β* protein whose molten globule is one of the best characterized. Although most residues in RBP have an unfolding midpoint in the range of 4−6 M urea, in α-LA, unfolding occurs between approximately 2 and 10 M urea with many residues not completely unfolded even in 10 M urea. This difference may be the result of the different secondary structure compositions and topologies of the two proteins. α-LA contains two subdomains with a large difference in stability between the α and *β* regions; the highly helical α-subdomain is much more stable than the *β*-subdomain (*15*, *16*). Within the α-domain, the individual helices show different stabilities. This marked noncooperativity may be possible because helices involve relatively short-range interactions and the unfolding of one helix may not perturb the other helices (*51*, *52*). In contrast, RBP has a single domain *β*-barrel topology; it may be difficult to unfold one strand of the *β*-sheet without causing disruption to other strands because of the longer-range hydrogen-bond interactions inherent in a *β*-sheet protein. A more cooperative unfolding transition may be a common feature of molten globules with a *β*-barrel topology.

At low pH, it is expected that salt bridges, which involve aspartic and glutamic acid residues, will be disrupted because of the protonation of the side chain carboxyl groups of these residues. Ptitsyn et al. have considered the role that the loss of salt bridges at low pH may play in the release of the ligand retinol from holo-RBP (*32*). Holo-RBP has several salt bridges located at the face of the retinol-binding pocket; the loss of these salt bridges, involving D39-R60, K87-D103, and D102-R121, may provide a mechanism for the release of retinol at low pH, and the resulting positively charged region on the surface of RBP may ensure the proper orientation of RBP relative to a negatively charged cell membrane (*32*).

The loss of salt bridges at low pH may play a more general role in the transition from the well-defined native structure to a less well-defined molten globule structure (*32*). The salt bridges in RBP provide important long-range interactions that help to stabilize the relative orientations of different regions of secondary structure in the well-ordered native tertiary structure. In particular, they link the *N*- and *C*-termini to the *β*-barrel (K12-D108, K17-D79, and D31-R166), the *C*-terminal helix to the second face of the *β*-barrel (D108-R155), and the two sides of the *β*-barrel (R19-D112) together (Figure 1C). The disruption of these salt bridges at low pH leads to a loss of important long-range contacts resulting in a destabilization of the native tertiary structure and a partial unfolding of RBP to a molten globule state. The lower stability observed for some *N*- and *C*-terminal residues may result from the loss of this first group of salt bridges. Interestingly, a study of the crystal structures of bovine RBP at pH values ranging from 2 to 9 shows a general increase in the distances between salt bridge pairs and an increase in the volume of the retinol-binding cavity as pH is lowered (*53*). Well-defined tertiary structure was not lost in this study because pH changes were achieved by soaking crystals of native holo-RBP (generated at pH 7) in lower pH buffers. However, the general trends observed are consistent with a loosening of the tertiary structure at low pH.

Ptitsyn et al. have proposed that the loss of the salt bridge formed between D102 and R121 may be involved in the release of retinol (*32*). The loss of this salt bridge may facilitate the formation of the proposed helical structure in the E−F loop by loosening the structural constraints on *β*-strand F (*34*). In a protein engineering study conducted by Sundaram and co-workers, it was shown that the E−F loop in RBP is involved in binding transthyretin and retinol retention inside the barrel (*31*). Interestingly, in two homologues of RBP, *β*-lactoglobulin and tear lipocalin, changes in the E−F loop have been proposed to be linked to ligand binding (*54*, *55*).

Understanding the mechanism of ligand release for RBP may also have ramifications for understanding the role of the lipocalins as carriers of hydrophobic ligands more generally. Most members of this superfamily are involved in the transport of a wide range of ligands in animals and plants, including in addition to retinol, sapid molecules, odorants, pigments for insect coloration, and progesterone (*56*). It has been proposed that another superfamily member, tear lipocalin, undergoes partial unfolding as a mechanism for lipid delivery (*57*, *58*). Lipocalins have also recently become the target of studies to re-engineer the barrel to carry novel ligands (*59*, *60*). Further biophysical studies, particularly those employing NMR spectroscopy, will significantly facilitate our understanding of the relationship between the molten globule state of the lipocalins and their folding and function.

ACKNOWLEDGMENT

We thank Dr. Jane Clarke (University of Cambridge) for the use of the stopped-flow apparatus for these studies and members of our group in Oxford for helpful discussions during the course of this work. We are grateful to Professor Keith Brew for the inspiration to study the lipocalin superfamily.

REFERENCES

1. Dolgikh, D. A., Gilmanshin, R. I., Brazhnikov, E. V., Bychkova, V. E., Semisotnov, G. V., Venyaminov, S. Y., and Ptitsyn, O. B. (1981) Alpha-lactalbumin: compact state with fluctuating tertiary structure, *FEBS Lett. 136*, 311−315.
2. Ohgushi, M., and Wada, A. (1983) Molten-globule state − a compact form of globular-proteins with mobile side-chains, *FEBS Lett. 164*, 21−24.
3. Ptitsyn, O. B., Pain, R. H., Semisotnov, G. V., Zerovnik, E., and Razgulyaev, O. I. (1990) Evidence for a molten globule state as a general intermediate in protein folding, *FEBS Lett. 262*, 20−24.
4. Jennings, P. A., and Wright, P. E. (1993) Formation of a molten globule intermediate early in the kinetic folding pathway of apomyoglobin, *Science 262*, 892−896.
5. Yao, J., Chung, J., Eliezer, D., Wright, P. E., and Dyson, H. J. (2001) NMR structural and dynamic characterization of the acid-unfolded state of apomyoglobin provides insights into the early events in protein folding, *Biochemistry 40*, 3561−3571.
6. Sandberg, A., Leckner, J., and Karlsson, B. G. (2004) Apo-azurin folds via an intermediate that resembles the molten-globule, *Protein Sci. 13*, 2628−2638.
7. Bose, H. S., Whittal, R. M., Baldwin, M. A., and Miller, W. L. (1999) The active form of the steroidogenic acute regulatory protein, StAR, appears to be a molten globule, *Proc. Natl. Acad. Sci. U.S.A. 96*, 7250−7255.
8. Twigg, P. D., Parthasarathy, G., Guerrero, L., Logan, T. M., and Caspar, D. L. D. (2001) Disordered to ordered folding in the regulation of diphtheria toxin repressor activity, *Proc. Natl. Acad. Sci. U.S.A. 98*, 11259−11264.

9. Vandergoot, F. G., Gonzalezmanas, J. M., Lakey, J. H., and Pattus, F. (1991) A molten-globule membrane-insertion intermediate of the pore-forming domain of colicin-A, *Nature 354*, 408−410.
10. Dunker, A. K., Ensign, L. D., Arnold, G. E., and Roberts, L. M. (1991) Proposed molten globule intermediates in Fd phage penetration and assembly, *FEBS Lett. 292*, 275−278.
11. Kirkitadze, M. D., Barlow, P. N., Price, N. C., Kelly, S. M., Boutell, C. J., Rixon, F. J., and McClelland, D. A. (1998) The herpes simplex virus triplex protein, V*P*23, exists as a molten globule. *Journal of Virology 72*, 10066−10072.
12. Safar, J., Roller, P. P., Gajdusek, D. C., and Gibbs, C. J. (1994) Scrapie amyloid (prion) protein has the conformational characteristics of an aggregated molten globule folding intermediate, *Biochemistry 33*, 8375−8383.
13. Dumoulin, M., Canet, D., Last, A. M., Pardon, E., Archer, D. B., Muyldermans, S., Wyns, L., Matagne, A., Robinson, C. V., Redfield, C., and Dobson, C. M. (2005) Reduced global cooperativity is a common feature underlying the amyloidogenicity of pathogenic lysozyme mutations, *J. Mol. Biol. 346*, 773−788.
14. Redfield, C. (2004) Using nuclear magnetic resonance spectroscopy to study molten globule states of proteins, *Methods 34*, 121−132.
15. Schulman, B. A., Kim, P. S., Dobson, C. M., and Redfield, C. (1997) A residue-specific NMR view of the non-cooperative unfolding of a molten globule, *Nat. Struct. Biol. 4*, 630−634.
16. Redfield, C., Schulman, B. A., Milhollen, M. A., Kim, P. S., and Dobson, C. M. (1999) alpha-lactalbumin forms a compact molten globule in the absence of disulfide bonds, *Nat. Struct. Biol. 6*, 948−952.
17. Baum, J., Dobson, C. M., Evans, P. A., and Hanley, C. (1989) Characterization of a partly folded protein by NMR methods: studies on the molten globule state of guinea-pig alpha-lactalbumin, *Biochemistry 28*, 7−13.
18. Hughson, F. M., Wright, P. E., and Baldwin, R. L. (1990) Structural characterization of a partly folded apomyoglobin intermediate, *Science 249*, 1544−1548.
19. Eliezer, D., Yao, J., Dyson, H. J., and Wright, P. E. (1998) Structural and dynamic characterization of partially folded states of apomyoglobin and implications for protein folding, *Nat. Struct. Biol. 5*, 148−155.
20. Eliezer, D., Chung, J., Dyson, H. J., and Wright, P. E. (2000) Native and non-native secondary structure and dynamics in the pH 4 intermediate of apomyoglobin, *Biochemistry 39*, 2894−2901.
21. Ramboarina, S., and Redfield, C. (2003) Structural characterisation of the human alpha-lactalbumin molten globule at high temperature, *J. Mol. Biol. 330*, 1177−1188.
22. Demarest, S. J., Fairman, R., and Raleigh, D. P. (1998) Peptide models of local and long-range interactions in the molten globule state of human alpha-lactalbumin, *J. Mol. Biol. 283*, 279−291.
23. Chowdhury, F. A., and Raleigh, D. P. (2005) A comparative study of the alpha-subdomains of bovine and human alpha-lactalbumin reveals key differences that correlate with molten globule stability, *Protein Sci. 14*, 89−96.
24. Wang, Y., and Shortle, D. (1996) A dynamic bundle of four adjacent hydrophobic segments in the denatured state of staphylococcal nuclease, *Protein Sci. 5*, 1898−1906.
25. Mok, K. H., Nagashima, T., Day, L. J., Hore, P. J., and Dobson, C. M. (2005) Multiple subsets of side-chain packing in partially folded states of alpha-lactalbumins, *Proc. Natl. Acad. Sci. U.S.A. 102*, 8899−8904.
26. Pervaiz, S., and Brew, K. (1985) Homology of beta-lactoglobulin, serum retinol-binding protein, and protein Hc, *Science 228*, 335−337.
27. Holm, L., and Sander, C. (1998) Dictionary of recurrent domains in protein structures, *Proteins: Struct., Funct., Bioinf. 33*, 88−96.
28. Greene, L. H., Hamada, D., Eyles, S. J., and Brew, K. (2003) Conserved signature proposed for folding in the lipocalin superfamily, *FEBS Lett. 553*, 39−44.
29. Naylor, H. M., and Newcomer, M. E. (1999) The structure of human retinol-binding protein (RBP) with its carrier protein transthyretin reveals an interaction with the carboxy terminus of RBP, *Biochemistry 38*, 2647−2653.
30. McCammon, M. G., Scott, D. J., Keetch, C. A., Greene, L. H., Purkey, H. E., Petrassi, H. M., Kelly, J. W., and Robinson, C. V. (2002) Screening transthyretin amyloid fibril inhibitors: Characterization of novel multiprotein, multiligand complexes by mass spectrometry, *Structure 10*, 851−863.
31. Sundaram, M., van Aalten, D. M. F., Findlay, J. B. C., and Sivaprasadarao, A. (2002) The transfer of transthyretin and receptor-binding properties from the plasma retinol-binding protein to the epididymal retinoic acid-binding protein, *Biochem. J. 362*, 265−271.
32. Ptitsyn, O. B., Zanotti, G., Denesyuk, A. L., and Bychkova, V. E. (1993) Mechanism of Ph-induced release of retinol from retinol-binding protein, *FEBS Lett. 317*, 181−184.
33. Bychkova, V. E., Dujsekina, A. E., Fantuzzi, A., Ptitsyn, O. B., and Rossi, G. L. (1998) Release of retinol and denaturation of its plasma carrier, retinol-binding protein, *Folding Des. 3*, 285−291.
34. Paci, E., Greene, L. H., Jones, R. M., and Smith, L. J. (2005) Characterization of the molten globule state of retinol-binding protein using a molecular dynamics simulation approach, *FEBS Journal 272*, 4826−4838.
35. Greene, L. H., Chrysina, E. D., Irons, L. I., Papageorgiou, A. C., Acharya, K. R., and Brew, K. (2001) Role of conserved residues in structure and stability: Tryptophans of human serum retinol-binding protein, a model for the lipocalin superfamily, *Protein Sci. 10*, 2301−2316.
36. Braun, D., Wider, G., and Wuthrich, K. (1994) Sequence-corrected N-15 random coil chemical-shifts, *J. Am. Chem. Soc. 116*, 8466−8469.
37. Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M., and Clore, G. M. (1989) Overcoming the overlap problem in the assignment of H-1-NMR spectra of larger proteins by use of 3-dimensional heteronuclear H-1-N-15 Hartmann−Hahn multiple quantum coherence and nuclear Overhauser multiple quantum coherence spectroscopy: application to interleukin-1-beta, *Biochemistry 28*, 6150−6156.
38. Kay, L. E., Keifer, P., and Saarinen, T. (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity, *J. Am. Chem. Soc. 114*, 10663−10665.
39. Frenkiel, T., Bauer, C., Carr, M. D., Birdsall, B., and Feeney, J. (1990) Hmqc-Noesy-Hmqc, a 3-dimensional NMR experiment which allows detection of nuclear Overhauser effects between protons with overlapping signals, *J. Magn. Reson. 90*, 420−425.
40. Bychkova, V. E., Berni, R., Rossi, G. L., Kutyshenko, V. P., and Ptitsyn, O. B. (1992) Retinol-binding protein is in the molten globule state at low pH, *Biochemistry 31*, 7566−7571.
41. Jones, J. A., Wilkins, D. K., Smith, L. J., and Dobson, C. M. (1997) Characterisation of protein unfolding by NMR diffusion measurements, *J. Biomol. NMR 10*, 199−203.
42. Bodenhausen, G., and Ruben, D. J. (1980) Natural abundance N-15 NMR by enhanced heteronuclear spectroscopy, *Chem. Phys. Lett. 69*, 185−189.
43. Semisotnov, G. V., Rodionova, N. A., Razgulyaev, O. I., Uversky, V. N., Gripas, A. F., and Gilmanshin, R. I. (1991) Study of the molten globule intermediate state in protein folding by a hydrophobic fluorescent-probe, *Biopolymers 31*, 119−128.
44. Engelhard, M., and Evans, P. A. (1995) Kinetics of interaction of partially-folded proteins with a hydrophobic fluorescent dye: evidence that molten globule character is maximal in early folding intermediates, *Protein Sci. 4*, 1553−1562.
45. Maki, K., Cheng, H., Dolgikh, D. A., Ramachandran, S. M. C. and Roder, H. (2004) Early events during folding of wild-type staphylococcal nuclease and a single-tryptophan variant studied by ultrarapid mixing, *J. Mol. Biol. 338*, 383−400.
46. Chapeaurouge, A., Johansson, J. S., and Ferreira, S. T. (2002) Folding of a de novo designed nativelike four-helix bundle protein, *J. Biol. Chem. 277*, 16478−16483.
47. Clark, P. L., Liu, Z. P., Rizo, J., and Gierasch, L. M. (1997) Cavity formation before stable hydrogen bonding in the folding of a beta-clam protein, *Nat. Struct. Biol. 4*, 883−886.
48. Kuwata, K., Shastry, R., Cheng, H., Hoshino, M., Batt, C. A., Goto, Y., and Roder, H. (2001) Structural and kinetic characterization of early folding events in beta-lactoglobulin, *Nat. Struct. Biol. 8*, 151−155.
49. Ragona, L., Catalano, M., Zetta, L., Longhi, R., Fogolari, F., and Molinari, H. (2002) Peptide models of folding initiation sites of bovine beta-lactoglobulin: Identification of nativelike hydrophobic interactions involving G and H strands, *Biochemistry 41*, 2786−2796.
50. Kobayashi, T., Ikeguchi, M., and Sugai, S. (2000) Molten globule structure of equine beta-lactoglobulin probed by hydrogen exchange, *J. Mol. Biol. 299*, 757−770.

51. Schulman, B. A., and Kim, P. S. (1996) Proline scanning mutagenesis of a molten globule reveals noncooperative formation of a protein's overall topology, *Nat. Struct. Biol. 3*, 682−687.
52. Quezada, C. M., Schulman, B. A., Froggatt, J. J., Dobson, C. M., and Redfield, C. (2004) Local and global cooperativity in the human alpha-lactalbumin molten globule, *J. Mol. Biol. 338*, 149−158.
53. Calderone, V., Berni, R., and Zanotti, G. (2003) High-resolution structures of retinol-binding protein in complex with retinol: pH-induced protein structural changes in the crystal state, *J. Mol. Biol. 329*, 841−850.
54. Ragona, L., Fogolari, F., Catalano, M., Ugolini, R., Zetta, L., and Molinari, H. (2003) EF loop conformational change triggers ligand binding in beta-lactoglobulins, *J. Biol. Chem. 278*, 38840−38846.
55. Gasymov, O. K., Abduragimov, A. R., Yusifov, T. N., and Glasgow, B. J. (2004) Interstrand loops CD and EF act as pH-dependent gates to regulate fatty acid ligand binding in tear lipocalin, *Biochemistry 43*, 12894−12904.
56. Flower, D. R. (1996) The lipocalin protein family: Structure and function, *Biochem. J. 318*, 1−14.
57. Gasymov, O. K., Abduragimov, A. R., Gasimov, E. O., Yusifov, T. N., Dooley, A. N., and Glasgow, B. J. (2004) Tear lipocalin: potential for selective delivery of rifampin, *Biochim. Biophys. Acta-Molecular Basis of Disease 1688*, 102−111.
58. Gasymov, O. K., Abduragimov, A. R., Yusifov, T. N., and Glasgow, B. J. (1998) Structural changes in human tear lipocalins associated with lipid binding, *Biochim. Biophys. Acta-Protein Structure and Molecular Enzymology 1386*, 145−156.
59. Schlehuber, S., and Skerra, A. (2002) Tuning ligand affinity, specificity, and folding stability of an engineered lipocalin variant − a so-called 'anticalin' − using a molecular random approach, *Biophys. Chem. 96*, 213−228.
60. Beste, G., Schmidt, F. S., Stibora, T., and Skerra, A. (1999) Small antibody-like proteins with prescribed ligand specificities derived from the lipocalin fold, *Proc. Natl. Acad. Sci. U.S.A. 96*, 1898−1903.

BI060229C