Equilibrium and Kinetic Studies on Folding of the Authentic and Recombinant Forms of Human α-Lactalbumin by Circular Dichroism Spectroscopy†

Tapan K. Chaudhuri,‡ Munehito Arai, Tomoki P. Terada,§ Teikichi Ikura,|| and Kunihiro Kuwajima*

Department of Physics, School of Science, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku Tokyo 113-0033, Japan

Received July 26, 2000; Revised Manuscript Received October 10, 2000

ABSTRACT: The equilibrium and kinetics of the unfolding and refolding of authentic and recombinant human α-lactalbumin, the latter of which had an extra methionine residue at the N-terminus, were studied by circular dichroism spectroscopy, and the results were compared with the results for bovine and goat α-lactalbumins obtained in our previous studies. As observed in the bovine and goat proteins, the presence of the extra methionine residue in the recombinant protein remarkably destabilized the native state, and the destabilization was entirely ascribed to an increase in the rate of unfolding. The thermodynamic stability of the native state against the unfolded state was lower, and the thermodynamic stability of the molten globule state against the unfolded state was higher for the human protein than for the other α-lactalbumins previously studied. Thus, the population of the molten globule intermediate was higher during the equilibrium unfolding of human α-lactalbumin by guanidine hydrochloride. Unlike the molten globule states of the bovine and goat proteins, the human α-lactalbumin molten globule showed remarkably more intense circular dichroism ellipticity than the native state in the far-ultraviolet region below 225 nm. During refolding from the unfolded state, human α-lactalbumin thus exhibited overshoot kinetics, in which the α-helical peptide ellipticity exceeded the native value when the molten globule folding intermediate was formed in the burst phase. The subsequent folding involved reorganization of nonnative secondary structures. It should be noted that the rate constant of the major refolding phase was approximately the same among the three types of α-lactalbumin and that the rate constant of unfolding was accelerated 18−600 times in the human protein, and these results interpreted the lower thermodynamic stability of this protein.

α-Lactalbumin, present in the whey of mammalian milk, is a small globular protein of 123 amino acid residues and is homologous to c-type lysozyme (*1*, *2*). The high-resolution X-ray crystallographic structure of authentic α-lactalbumin has been identified for several different mammalian species, including human, goat, guinea pig, and baboon (*3*−*6*). The X-ray structure of recombinant α-lactalbumin, which was expressed in *Escherichia coli* (*E. coli*)[1] and refolded, was also determined for bovine and goat specimens and was shown to be identical with the structure of the authentic proteins (*6*, *7*). The overall structure of α-lactalbumin can be divided into two subdomains (Figure 1). The α-domain encompasses two peptide regions, 1−37 and 84−123, and is comprised of four α-helical elements (A, B, C, and D helices). The β-domain encompasses one peptide region, 38−83, and is comprised of a three-stranded β-sheet and a 3_{10}-helix (*3*−*7*).

α-Lactalbumin has often been used as a model protein to study the mechanism of protein folding (*8*−*11*). This protein most typically exhibits a molten globule intermediate as an early kinetic folding intermediate as well as an equilibrium unfolding intermediate (*8*, *10*−*12*). As detection and characterization of such folding intermediates are required in the protein folding studies, the properties of α-lactalbumin prove to be quite useful for such studies. The molten globule state of the protein has a nativelike secondary structure and is compact as regards shape. However, it does not have a specific tertiary structure (i.e., specific side-chain packing), and the α-domain is more organized than the β-domain in the α-lactalbumin molten globule (*11*): the α- and β-domains are shown in dark-gray and white, respectively, in Figure 1.

Previous folding studies of α-lactalbumin have used bovine, goat, and human α-lactalbumins either in the authentic or recombinant form (*7*, *8*, *13*, *14*). Comparatively little attention has been paid to possible differences in folding behavior that may arise from species differences between α-lactalbumins and also from differences between the authentic and recombinant forms. Only recently, the effects of the extra N-terminal methionine residue on the folding behavior of the recombinant protein have been recognized for bovine and goat α-lactalbumins (*7*, *15*, *16*). It has been

† Supported by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, and Science of Japan. T.K.C. was a postdoctoral fellow of the Japan Society for the Promotion of Science.

* To whom correspondence should be addressed. Phone: +81-3-5841-4128. Fax: +81-3-5841-4512. E-mail: kuwajima@phys.s.u-tokyo.ac.jp.

‡ Present address: Howard Hughes Medical Institute and Department of Genetics, Yale University School of Medicine, New Haven, CT 06510.

§ Present Address: Graduate School of Human Informatics, Nagoya University, Nagoya 464-8601, Japan.

|| Present address: Kirin Brewery Co. Ltd., 1-13-5 Fukuura, Kanazawa-ku, Yokohama City, Kanagawa 236-0004, Japan.

[1] Abbreviations: *E. coli*, *Escherichia coli*; CD, circular dichroism; GdnHCl, guanidine hydrochloride; UV, ultraviolet; N, the native state; I, the intermediate state; U, the unfolded state; HPLC, high performance liquid chromatography.

10.1021/bi001735j CCC: $19.00
Published on Web 11/18/2000

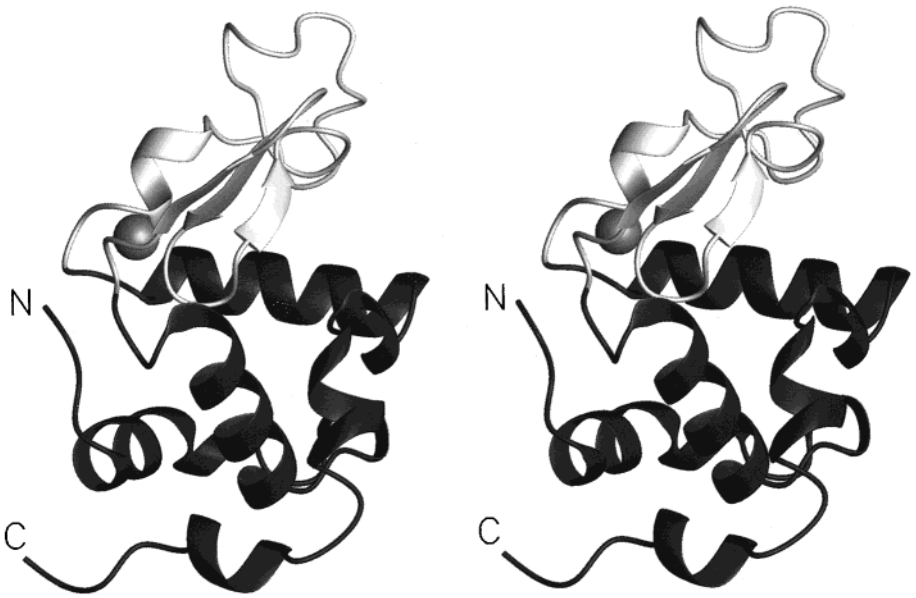

FIGURE 1: Stereoview in the ribbon model of human α-lactalbumin (PDB code 1A4V) (*43*). The α-domain and *β*-domain of the molecule are shown in dark-gray and white, respectively. A gray sphere indicates the bound calcium ion at the stronger binding site. The N- and C-termini are denoted by N and C, respectively. The figure was drawn using MOLMOL 2K.1 (*44*).

shown that the presence of the extra N-terminal methionine residue in the *E. coli* expressed recombinant proteins remarkably destabilizes their native state. Because α-lactalbumin is a typical model globular protein used in folding studies, detailed investigation of the possible differences among different α-lactalbumin species is important.

In the present study, we focus on the equilibrium and kinetics of the unfolding and refolding of authentic and recombinant human α-lactalbumin using circular dichroism (CD) spectroscopy. The results are compared with results obtained in our previous studies of corresponding bovine and goat proteins (*7*, *17*−*19*). We show that the native state of the human protein is remarkably less stable than that of either of the other α-lactalbumins studied previously. Moreover, this destabilization results in the higher population of the molten globule intermediate at a moderate concentration (∼2 M) of guanidine hydrochloride (GdnHCl). These results are considered together with results from other studies in a discussion of the mechanism of α-lactalbumin folding in light of the known three-dimensional structures of human, bovine, and goat α-lactalbumins.

MATERIALS AND METHODS

Materials. Authentic human α-lactalbumin was prepared from fresh human milk by a previously described method (*20*). Recombinant human α-lactalbumin was expressed in *E. coli* BL21(DE3)pLysS as inclusion bodies by the use of an expression plasmid, pHLAC (*13*), which was the generous gift of Peter Kim. The expression and purification of the recombinant protein were carried out by the same methods described previously for the use of the goat protein (*7*). All other materials used were the same as those described in another report (*7*).

Chemical Analysis of Authentic and Recombinant α-Lactalbumin. N-Terminal sequencing of the authentic and recombinant proteins was carried out using an automated Applied Biosystem sequencer model 477a, which was equipped with a model 120A on-line PTH amino acid analyzer. Mass spectrometric analyses of the authentic and recombinant proteins were carried out by the MALDI-TOF-MS mass spectrometric method, as described previously (*7*). The only chemical difference observed between the two proteins was the presence of an extra N-terminal methionine residue in the recombinant protein.

Equilibrium CD Spectra. Equilibrium CD spectra were taken on a Jasco J-720 spectropolarimeter using an optical cuvette with a path-length of 1.0 mm for measurements in the far-ultraviolet (UV) region and 10.0 mm for measurements in the near-UV region (*7*, *10*). The CD spectra of the protein were measured in 50 mM sodium cacodylate, 50 mM NaCl and 1 mM $CaCl_2$ at pH 7.0 and 25 °C. The solutions for the GdnHCl-induced equilibrium unfolding studies were prepared in the same buffer containing 1 mM $CaCl_2$.

Kinetic CD Measurements. Refolding and unfolding reactions of authentic and recombinant proteins were induced by GdnHCl concentration jumps, which were performed by a stopped-flow CD apparatus (UNISOKU Inc., Japan) installed in the cell compartment of the J-720 spectropolarimeter (*7*, *10*). All kinetics were measured in the presence of 50 mM sodium cacodylate, 50 mM NaCl, and 1 mM $CaCl_2$ at pH 7.0 and 25 °C. The dead time of the stopped-flow CD apparatus was 25 ms when a 4 mm cuvette was used. The concentration of the protein stock solution was 1.5−2.0 mg/mL. The initial protein solutions before the concentration jump contained 0 and 5.5 M GdnHCl for unfolding and refolding experiments, respectively. The diluent solution contains the same buffer (50 mM sodium cacodylate, 50 mM NaCl, and 1 mM $CaCl_2$ at pH 7.0 and 25 °C) and an appropriate concentration of GdnHCl. The two solutions were mixed with a mixing ratio of 1:10 (protein:diluent).

RESULTS

CD Spectra of Recombinant and Authentic α-Lactalbumin. The far- and near-UV CD spectra of authentic and recombinant human α-lactalbumin were measured in a 50 mM

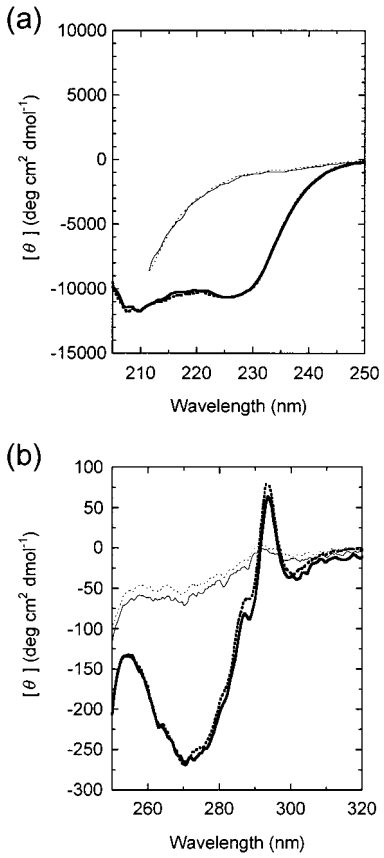

FIGURE 2: (a) Far- and (b) near-UV CD spectra of authentic and recombinant human α-lactalbumin at pH 7.0 and 25 °C (50 mM sodium cacodylate, 50 mM NaCl, and 1 mM $CaCl_2$). In panels a and b, thick and thin lines show the spectra of the native and the unfolded states, respectively, and solid and dotted lines show the spectra of the authentic and recombinant proteins, respectively. The spectra of the unfolded state was taken at 6 M GdnHCl in 50 mM sodium cacodylate, 50 mM NaCl, and 1 mM $CaCl_2$.

sodium cacodylate buffer that contained 50 mM NaCl and 1 mM $CaCl_2$, at pH 7.0 and 25 °C (Figure 2a and b). In both the aromatic and peptide regions, the spectra of the two proteins in the native state are coincident with each other, indicating that the proteins have very similar secondary and tertiary structures. A previous report by Peng et al. (*21*) has shown that the recombinant human α-lactalbumin exhibits full activity in stimulating the galactose transfer reaction of lactose synthase (*1*). Thus, this previous finding, when considered in the context of the present data, clearly indicates that the recombinant protein was correctly folded into the native structure. The CD spectra in the presence of 6 M GdnHCl of the authentic and recombinant proteins also coincided with each other, and no residual structure was observed at this denaturant concentration (Figure 2). This finding indicates that the denatured states of both of the proteins were fully unfolded.

Equilibrium Unfolding. GdnHCl-induced equilibrium unfolding transitions of authentic and recombinant human α-lactalbumin were observed in the peptide and aromatic CD spectra. Figure 3 shows the unfolding transition curves of the two proteins measured by the CD ellipticities at 222 and 270 nm. Although the transition curves of the proteins measured at 222 nm are close to each other above 2 M GdnHCl, they are remarkably different at concentrations below 2 M GdnHCl. Furthermore, the curve for the

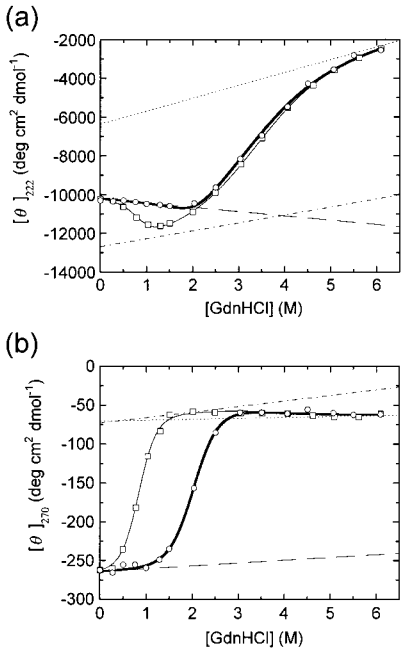

FIGURE 3: GdnHCl-induced equilibrium unfolding transition curves of authentic (circles) and recombinant (squares) human α-lactalbumin measured by the CD ellipticities at 222 nm (a) and 270 nm (b) at pH 7.0 and 25 °C (50 mM sodium cacodylate, 50 mM NaCl, and 1 mM $CaCl_2$). In panels a and b, long-dashed, dash-and-dot, and dotted lines show the baselines of the native, the intermediate, and the unfolded states, respectively. Thick and thin solid lines are the theoretical unfolding transition curves of authentic and recombinant α-lactalbumin, respectively, fitted to eq 4.

recombinant protein shows a minimum at a concentration of 1.2 M GdnHCl. The transition curves measured at 270 nm are not coincident with each other, and apparently, the unfolding transition of the recombinant protein occurs at a lower concentration of GdnHCl.

It can also be observed in Figure 3 that when the unfolding transition curves measured at the different wavelengths are compared in the same protein, they do not coincide with each other, and the transition curve measured at 222 nm is observed at a higher concentration of GdnHCl in both the authentic and recombinant proteins. This is clear evidence for the presence of at least one stable intermediate in the denaturation process. The structure of the aromatic residues that contribute to the Cotton effect at 270 nm is apparently disrupted at an earlier stage of the process, whereas the secondary structure of the polypeptide chain measured at 222 nm is apparently destroyed at a relatively late stage.

According to these results, it is reasonable to assume a three-state mechanism in which a stable intermediate state (I) is populated during the equilibrium unfolding transition of human α-lactalbumin. This intermediate (I) may correspond to the molten globule state as well-documented in equilibrium unfolding studies of bovine α-lactalbumin. Thus, the proposed mechanism is

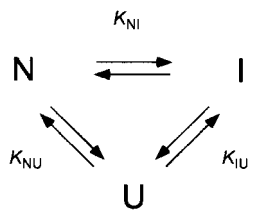

Table 1: Equilibrium Unfolding Parameters of Human, Bovine and Goat α-Lactalbumins (pH 7.0, 0.1 M Na^+−1 mM $CaCl_2$, 25 °C)[a]

		$\Delta G_{NU}^{H_2O}$ (kcal/mol)	m_{NU} (kcal/mol M)	$\Delta G_{NI}^{H_2O}$ (kcal/mol)	m_{NI} (kcal/mol M)	$\Delta G_{IU}^{H_2O}$ (kcal/mol)	m_{IU} (kcal/mol M)
human	(authentic)	6.9 ± 0.5	2.9 ± 0.2	4.5 ± 0.4	2.2 ± 0.2	2.4 ± 0.7	0.77 ± 0.28
	(recombinant)	5.3 ± 0.4	4.0 ± 0.2	2.7 ± 0.1	3.2 ± 0.2	2.7 ± 0.4	0.81 ± 0.27
bovine	(authentic)	7.2	2.4			1.42	0.75
goat	(authentic)	13.8 ± 0.7	4.4 ± 0.2			0.66 ± 0.05[b]	0.35 ± 0.02[b]
	(recombinant)	10.4 ± 0.5	3.9 ± 0.2			0.70 ± 0.05[b]	0.33 ± 0.01[b]

[a] $\Delta G_{IU}^{H_2O} = \Delta G_{NU}^{H_2O} - \Delta G_{NI}^{H_2O}$, and $m_{IU} = m_{NU} - m_{NI}$. The data for the bovine protein were from (*17*), and those for the goat protein from (*7*, *19*). [b] Urea-induced unfolding of the molten globule state at pH 2.0 and 20 °C.

where K_{NI}, K_{IU}, and K_{NU} are the equilibrium constants for the N ⇌ I, I ⇌ U, and N ⇌ U transitions, respectively. The observed ellipticity of the protein [$\theta_{obs}(c)$] at any concentration of the denaturant is given by the sum of the contributions from the three states as

$$\theta_{obs}(c) = \theta_N f_N(c) + \theta_I f_I(c) + \theta_U f_U(c) \quad (1)$$

where $f_N(c)$, $f_I(c)$, and $f_U(c)$ are the fractions of the three states at a GdnHCl concentration of c ($f_N + f_I + f_U = 1$), and θ_N, θ_I, and θ_U are the ellipticity values of the pure N, I, and U states, respectively. The f_N, f_I, and f_U are related to the equilibrium constants, K_{NI}, and K_{NU}, of the unfolding transitions from N to I and N to U, respectively, and hence are related to the corresponding free energy changes, ΔG_{NI} and ΔG_{NU}, as follows:

$$f_N = 1/(1 + K_{NI} + K_{NU}) = 1/[1 + \exp(-\Delta G_{NI}/RT) + \exp(-\Delta G_{NU}/RT)]$$

$$f_I = K_{NI}/(1 + K_{NI} + K_{NU}) = \exp(-\Delta G_{NI}/RT)/[1 + \exp(-\Delta G_{NI}/RT) + \exp(-\Delta G_{NU}/RT)] \quad (2)$$

$$f_U = K_{NU}/(1 + K_{NI} + K_{NU}) = \exp(-\Delta G_{NU}/RT)/[1 + \exp(-\Delta G_{NI}/RT) + \exp(-\Delta G_{NU}/RT)]$$

where R and T are the gas constant and the absolute temperature, respectively. For many globular proteins, the free energy changes of unfolding are known to vary approximately linearly with c (*22*), such that

$$\Delta G_{NI} = \Delta G_{NI}^{H_2O} - m_{NI}c$$

$$\Delta G_{NU} = \Delta G_{NU}^{H_2O} - m_{NU}c \quad (3)$$

where $\Delta G_{NI}^{H_2O}$, and $\Delta G_{NU}^{H_2O}$ are the ΔG_{NI} and ΔG_{NU} at 0 M GdnHCl, respectively, and m_{NI} and m_{NU} represent the dependence of the respective free energy changes on c and thus cooperativity indexes of the transitions. From eqs 1−3, $\theta_{obs}(c)$ is given by

$$\theta_{obs}(c) = \frac{\theta_N + \theta_I \exp[-(\Delta G_{NI}^{H_2O} - m_{NI}c)/RT] + \theta_U \exp[-(\Delta G_{NU}^{H_2O} - m_{NU}c)/RT]}{1 + \exp[-(\Delta G_{NI}^{H_2O} - m_{NI}c)/RT] + \exp[-(\Delta G_{NU}^{H_2O} - m_{NU}c)/RT]} \quad (4)$$

In general, θ_N, θ_I, and θ_U are also dependent on c, and we assume a linear dependence on c, as $\theta_N = a_1 + a_2c$, $\theta_I = a_3 + a_4c$, and $\theta_U = a_5 + a_6c$, where a_i ($i = 1-6$) are constants.

The data presented in Figure 3 were analyzed on the basis of eq 4 by the method of nonlinear least-squares. In this analysis, we performed the global fitting, in which all the transition curves measured at different wavelengths (222 and 270 nm) for the authentic and recombinant proteins were fitted simultaneously, on the assumption that the θ_N, θ_I, and θ_U were common to the authentic and recombinant proteins (*23*). The fitting variables in the least-squares analysis were thus two sets of $\Delta G_{NI}^{H_2O}$, $\Delta G_{NU}^{H_2O}$, m_{NI}, and m_{NU} for the authentic and recombinant proteins and two sets of a_1-a_6 for the transition curves at the two different wavelengths (222 and 270 nm). The best fit values of the equilibrium unfolding parameters thus obtained are listed in Table 1.

The solid lines in Figure 3 represent the theoretical curves drawn with the parameter values of Table 1. The theoretical curves show excellent agreement with the experimental data, indicating the validity of the three-state interpretation. Therefore, only the three states, namely, the N, I, and U states, were sufficient and no other states were required for interpreting the unfolding transition of both authentic and recombinant human α-lactalbumin. The results clearly indicate that the N state of the authentic protein is more stable than that of the recombinant protein by 1.6−1.8 kcal/mol (Table 1). However, there was no difference observed between the two proteins as regards the stability of the I state. As shown in our previous study of goat α-lactalbumin (*7*), the destabilization of the recombinant protein may arise from the presence of the extra N-terminal methionine residue in the protein (see Discussion).

From the theoretical curves, the apparent fractions of I formed during the unfolding transition were calculated using eq 2 (Figure 4). The maximum fractional populations of the I state for the authentic and recombinant proteins are observed at about 2.5 and 1.5 M GdnHCl and are about 60 and 90%, respectively. The lower population of the I state in the authentic protein is a consequence of its enhanced N-state stability (see below). The unfolding transition curve of the recombinant protein shows an initial decrease of the CD ellipticity value below 1 M GdnHCl. This curve was attributed to the accumulation of the I state, as the I state had a more intense CD signal than the N state. Because the I state showed less accumulation in the authentic protein due to its enhanced N-state stability, we were unable to observe the decreasing trend of the ellipticity at 222 nm at the beginning of the unfolding transition of the authentic protein.

Kinetics of Refolding and Unfolding. To investigate how the destabilization of the N state of the recombinant protein is reflected in the kinetics of its unfolding and refolding, we

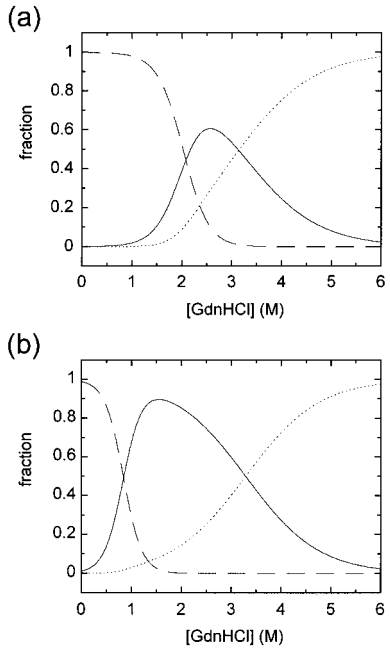

FIGURE 4: GdnHCl concentration dependence of the fractions of the native (long-dashed line), the intermediate (solid line), and the unfolded states (dotted line) of authentic (a) and recombinant (b) human α-lactalbumin at pH 7.0 and 25 °C (50 mM sodium cacodylate, 50 mM NaCl, and 1 mM $CaCl_2$). The fraction of each state was determined by eq 2.

performed refolding and unfolding kinetic studies of the two proteins using stopped-flow CD. The unfolding and refolding reactions were induced by concentration jumps of GdnHCl from 0 to 3.0 M and from 5.5 to 0.5 M, respectively. The reactions were monitored by the ellipticity change at 222 nm at pH 7.0 and 25 °C. The kinetic progress curves are shown in Figure 5, and the data were fitted by the nonlinear least-squares method with the following equation,

$$\theta(t) = \theta(\infty) + \Delta\theta_{obs} \sum_i \alpha_i \exp(-k_i t) \qquad (5)$$

where $\theta(t)$ and $\theta(\infty)$ are the observed values of the ellipticity at time t and infinite time, respectively, $\Delta\theta_{obs}$ is the observed total amplitude $[\theta(0) - \theta(\infty)]$, and k_i and α_i are the apparent first-order rate constant and the fractional amplitude, respectively, of the ith kinetic phase.

The kinetic progress curve for unfolding of the authentic protein fit well into a single-exponential equation, and the apparent rate constant and the amplitude of the ellipticity change of the kinetic process were 0.88 (±0.03) s^{-1} and −1786 (±29) deg cm^2 $dmol^{-1}$, respectively. The unfolding of the recombinant protein occurred too quickly to measure it accurately with a good signal-to-noise ratio. The kinetic study of the unfolding of recombinant α-lactalbumin monitored by stopped-flow fluorescence demonstrated that the rate constant for the unfolding reaction for the change in GdnHCl concentration from 0 to 3.0 M was about 71 s^{-1} (to be published elsewhere).

The refolding progress curves for the authentic and recombinant proteins are well-represented by a single exponential process. The apparent rate constants for the refolding process of the proteins are presented in Table 2. The rate constants for refolding reaction are essentially the same in the case of both the authentic and recombinant

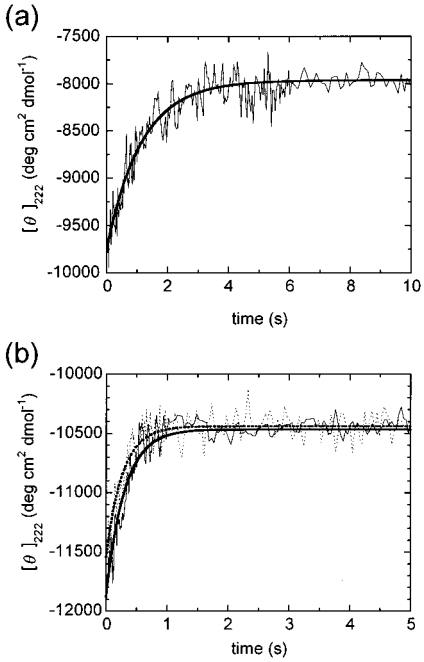

FIGURE 5: Kinetic (a) unfolding and (b) refolding curves of authentic (solid lines) and recombinant (dotted lines) human α-lactalbumin monitored by the CD ellipticity at 222 nm. The unfolding (a) and refolding (b) reactions were initiated by the GdnHCl concentration jump from 0 to 3.0 M and from 5.5 to 0.5 M, respectively, at pH 7.0 and 25 °C (50 mM sodium cacodylate, 50 mM NaCl, and 1 mM $CaCl_2$). Thick lines are fitting curves assuming a single-exponential function. In panel a, only the curves for the authentic protein are shown, because the unfolding reaction of the recombinant protein occurred too rapidly to be accurately measured by stopped-flow CD (see text).

Table 2: Rate Constants for Refolding and Unfolding Reactions of Human, Bovine, and Goat α-Lactalbumin (pH 7.0, 0.1 M Na^+−1 mM $CaCl_2$, 25 °C)[a]

		refolding at 0.5 M GdnHCl (s^{-1})	unfolding at 3 M GdnHCl (s^{-1})
human	(authentic)	3.3 ± 0.1	0.88 ± 0.03
	(recombinant)	3.8 ± 0.1	71
bovine	(authentic)	5.9	0.050
goat	(authentic)	4.9 ± 0.3	
	(recombinant)	5.7 ± 0.4	0.124

[a] The rate constants of refolding for the bovine and goat proteins are those of the major refolding phase. The data for the bovine protein were from (*18*), and those for the goat protein from (*7*, *37*).

proteins. Thus, these results demonstrate that the destabilization of the native state of recombinant human α-lactalbumin is explained by the increased unfolding rate of the protein.

It is of note that the refolding progress curve behaves in such a way that the ellipticity increases during the refolding process (Figure 5b). On the other hand, the equilibrium transition curve in Figure 3a shows a decrease in the ellipticity upon refolding (Figure 3a). Therefore, there is a transient refolding intermediate that accumulates within the dead time of the measurement (25 ms) and has more intense CD ellipticity (i.e., the lower absolute value) than the N state. The more intense CD ellipticity of the intermediate may be brought about by the formation of a non-native secondary structure within the burst phase of refolding.

Comparison of the Equilibrium Unfolding and the Transient Refolding Intermediates. Figure 6 shows the far- and

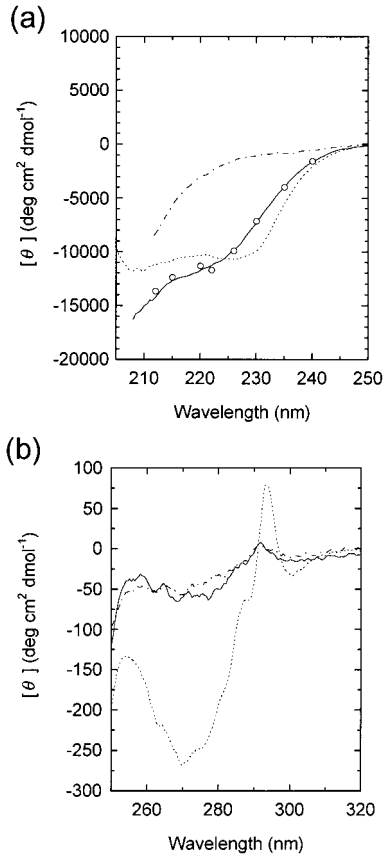

FIGURE 6: (a) Far- and (b) near-UV CD spectra of the recombinant human α-lactalbumin at 1.5 M GdnHCl (solid lines) at pH 7.0 and 25 °C (50 mM sodium cacodylate, 50 mM NaCl, and 1 mM $CaCl_2$). The spectra of the native (dotted lines) and the unfolded states (dash-and-dot lines) are also shown for comparison. Open circles in panel a are obtained by extrapolating the refolding curves measured at various wavelengths to zero time of the reaction. The refolding reaction was initiated by a GdnHCl concentration jump from 5.5 to 0.5 M at pH 7.0 and 25 °C.

near-UV CD spectra of the recombinant protein at a concentration of 1.5 M GdnHCl, where the population of the intermediate state is at a maximum (Figure 4a). In the far-UV region, the CD ellipticity is more negative than that in the N state below 225 nm. The near-UV CD spectra of the recombinant protein under these conditions, however, shows only minor Cotton effects, indicating the absence of the specific tertiary structure of aromatic side chains.

To compare the transient refolding intermediate with the equilibrium I state, we used stopped-flow CD in order to perform kinetic refolding experiments on the recombinant protein at various wavelengths. The refolding was initiated by the GdnHCl concentration jump from 5.5 to 0.5 M in the presence of 1 mM $CaCl_2$ at pH 7.0 and 25 °C. From the kinetic refolding curves, we obtained ellipticity by extrapolation at time zero of refolding; the ellipticity values of the transient intermediate thus obtained were plotted against the wavelength, as shown in Figure 6. The results show that the CD spectra of the transient refolding intermediate coincided with those of the equilibrium I state, suggesting that the two intermediate states are identical to each other.

Gel Electrophoresis and Ion-Exchange Chromatography. Differences between the properties of authentic and recombinant α-lactalbumin were also observed by gel electrophoresis and ion-exchange chromatography. The electrophoretic mobility of the recombinant protein in a nondenaturing polyacrylamide gel at pH 9.4 was significantly larger than that of the authentic protein. Elution profiles of the authentic and recombinant proteins by anion-exchange high-performance liquid chromatography (HPLC) using a RESOURCE Q column (Pharmacia Biotech) with a linear gradient from 0 to 0.5 M NaCl in 10 mM $NaH_2PO_4-Na_2HPO_4$ buffer at pH 7.0 showed that the retention time was longer for the recombinant protein (31 min) than for the authentic one (24 min). Both of these results indicate that the recombinant protein is more negatively charged.

DISCUSSION

The equilibrium and kinetics of the unfolding and refolding of authentic and recombinant human α-lactalbumin, the latter of which has an extra methionine residue at the N-terminus, have been studied by CD spectroscopy. Many previous studies have used bovine, goat, and human α-lactalbumins, either in the authentic or in the recombinant form, as model proteins to study the mechanism of protein folding (*7*, *8*, *10*, *13*, *24*). α-Lactalbumin exhibits the molten globule state as an equilibrium unfolding intermediate and as an early kinetic folding intermediate (*8*, *11*, *12*), and this property of the protein is useful for elucidating the folding mechanism. However, previous studies have not considered potential differences in the unfolding behavior among different α-lactalbumin species nor those between the authentic and recombinant forms within a single species. Only recently, the effect of the extra N-terminal methionine residue on the unfolding behavior has been recognized for bovine and goat α-lactalbumins (*7*, *15*, *16*). The present results, when considered together with our previous studies, highlight the differences in the unfolding behavior among bovine, goat, and human α-lactalbumins. In the following discussion, we first describe the differences between different α-lactalbumins, and then the discussion proceeds to focus on the effects of the N-terminal methionine in human α-lactalbumin. Finally, the folding mechanism of this protein will be considered. We discuss these results in light of the known X-ray structures of human and other α-lactalbumins.

Unfolding Equilibria of Human, Bovine, and Goat α-Lactalbumins. Because α-lactalbumin is a Ca^{2+}-binding protein (*25*), its N state is stabilized by the Ca^{2+} ion. As a result, both bovine and goat α-lactalbumins in the presence of 1 mM $CaCl_2$ undergo a cooperative two-state unfolding transition between N and U when treated with GdnHCl (*7*, *15*, *17*). Their molten globule intermediates have been observed only in the absence of the stabilizing Ca^{2+} ion at a moderate concentration of GdnHCl (~2 M) at neutral pH (*17*). In contrast, the molten globule state of human α-lactalbumin is apparently more stable and significantly populated in the presence of 1 mM $CaCl_2$ (Figure 4).

The equilibrium unfolding parameters of bovine and goat α-lactalbumins have been obtained quantitatively in our previous studies (*7*, *17*). These parameters are compared with those of the human protein, shown in Table 1. The thermodynamic stability of the molten globule (I) state against the fully unfolded (U) state is 0.7−1.4 kcal/mol for bovine and goat α-lactalbumins (*17*, *19*), and the thermodynamic stability of the I state against the U state is more stable for the human protein (2.4−2.6 kcal/mol) (Table 1).

The thermodynamic stability of the N state against the U state is lower for the human protein (i.e., 6.9 and 5.3 kcal/mol for the authentic and recombinant forms, respectively) than it is in the case of the bovine and goat proteins (i.e., 7.2−13.8 kcal/mol, depending on the protein species and the authentic or recombinant form) (*7*, *17*). Therefore, the higher equilibrium population of the molten globule intermediate in the human protein arises from the lower stability of its N state and the higher stability of its molten globule state. A study of chimeric proteins of bovine and human α-lactalbumins recently reported by Mizuguchi et al. (*26*) has shown that among three portions of the α-lactalbumin molecule, residues 1−34, 35−85, and 86−123, the first portion (residues 1−34) is most important for the stability of the human α-lactalbumin molten globule state.

Unlike the molten globule states of bovine and goat α-lactalbumins, the human α-lactalbumin molten globule shows remarkably more intense CD ellipticity than the N state in the far-UV region below 225 nm. This more intense CD ellipticity resulted in an unusual minimum in the unfolding transition curve measured by the ellipticity at 222 nm in recombinant human α-lactalbumin (Figure 3a). Because the CD ellipticity in the far-UV region mainly arises from the α-helical structure, it appears that the non-native α-helical structure is more populated in the human protein.

Effect of the N-Terminal Methionine Residue. As observed previously for bovine and goat α-lactalbumins (*7*, *15*, *16*), the extra methionine residue was present at the N-terminus of recombinant human α-lactalbumin, and the presence of this extra methionine residue remarkably destabilized the N state, whereas it did not significantly affect the stability of the molten globule intermediate. Our previous study of the goat protein has shown that this destabilization is mainly brought about by a conformational entropy effect of the extra residue in the U state (*7*). The difference in the stabilization free energy between the authentic and recombinant proteins is −1.5 to −3.5 kcal/mol for both the human and goat proteins (*7*), suggesting that the same destabilization mechanism may be employed for the two proteins.

The presence of the extra methionine residue makes the N state of the protein more negatively charged, as indicated by changes in mobility in the native gel electrophoresis and in retention time in ion-exchange chromatography. These effects of the methionine residues were also observed in the goat α-lactalbumin previously studied, but these were more significant in the human protein (*7*). Because these effects disappeared in the fully unfolded state in 8 M urea (data not shown), the structural folding to native structure in the presence of the extra methionine at the N-terminus affects the electrostatic interactions of the protein and renders it more negatively charged.

Although the destabilization of the N state by the extra N-terminal methionine residue has been found previously for bovine and goat α-lactalbumins (*7*, *15*, *16*), and now for human α-lactalbumin, it is not specific for α-lactalbumin. The effect of the extra N-terminal methionine has thus important implications regarding the stability of recombinant proteins expressed in *E. coli*. Whether the presence of the N-terminal methionine destabilizes the N state as observed in α-lactalbumin or not, is, however, dependent on protein species. For example, recombinant hen egg-white and human lysozymes contain the N-terminal methionine residue and have lower solubility and stability than the authentic forms (*27*−*29*). Similarly, recombinant apomyoglobin, immunoglobulin V_L domains, and maltose-binding protein expressed in *E. coli* contain the extra N-terminal methionine residue and are less stable than the authentic proteins (*30*−*32*). On the other hand, the presence of the extra N-terminal methionine or the extension or truncation of the N-terminal residues does not interfere with the N-state stability in certain other globular proteins (*33*, *34*). In recombinant ribonuclease A and circularly permuted recombinant ribosomal S6 protein, the extra N-terminal methionine is even known to stabilize the N state (*35*, *36*).

Folding Mechanism of Human α-Lactalbumin. From the present results, the kinetic folding of human α-lactalbumin consists of two successive stages: (i) formation of the molten globule folding intermediate (U→I), and (ii) formation of the N state (I→N). The second stage is rate limiting and involves the highest-free energy transition state of folding. The destabilization of the N state by the N-terminal methionine in the recombinant protein is entirely ascribed to an increase in the unfolding rate, and there is no significant difference in the refolding rate constant between the authentic and recombinant proteins, suggesting that the native structure around the N-terminus has not yet been organized in the transition state of folding. All of these results are thus consistent with those observed in our previous studies of bovine and goat α-lactalbumins (*7*, *9*, *18*, *37*).

However, in contrast to the bovine and goat proteins, human α-lactalbumin exhibits overshoot kinetics of refolding, in which the α-helical peptide ellipticity exceeds the native value in the burst-phase and then decreases to the native ellipticity (Figure 5b). Thus, the second stage of refolding involves a reorganization of the nonnative secondary structure, and this reorganization is more prominent in the human protein.

When we compare the rate constants of unfolding and refolding among human, bovine, and goat α-lactalbumins (Table 2), the rate constant of the major refolding phase is approximately the same among the three proteins, but the rate constant of unfolding is selectively accelerated 18−600 times in the human protein. This apparently indicates that the stability of the transition state of refolding as compared with the stability of the U state is approximately the same among the three proteins and that the N state is selectively destabilized in the human protein (*18*, *38*). The destabilization of the N state of the human protein results in the acceleration of the unfolding rate because the free energy of activation of unfolding decreases to the same extent as that caused by the free-energy increase (destabilization) of the N state.

Structural Interpretations. There are 29 and 30 amino acid replacements between goat and human α-lactalbumins and between bovine and human α-lactalbumins, respectively (Figure 7) (*39*−*42*). Although most of the replacements are located in the surface or flexible loop regions of the α-lactalbumin molecule, the replacement of Trp26 (bovine and goat) by a leucine residue in the human protein occurs within the major hydrophobic core of the α-domain of the molecule. The reduction in residue size at position 26 is compensated for by an increase in the side-chain size at position 30 [methionine in the human protein compared with threonine (bovine) or alanine (goat)]. Otherwise, the side-chain packing is subtly changed and the hydrophobic core

```
        1        10         20         30         40         50         60         70
Human   KQFTKCELSQ LLKDIDGYGG IALPELICTM FHTSGYDTQA IVENNESTEY GLFQISNKLW CKSSQVPQSR
Goat    E-L----VF- K---LKD--- VS---WV--A ---------- --Q--D---- -----N--I- --DD-N-H--
Bovine  E-L----VFR E---LK---- VS---WV--T ---------- --Q--D---- -----N--I- --DD-N-H-S
            A-helix              B-helix               β       β      β

                 80         90        100        110        120
Human   NICDISCDKF LDDDITDDIM CAKKILDIKG IDYWLAHKAL CTEKLEQWLC EKL
Goat    ---N------ ----L----V -------KV- -N-------- -S---D---- ---
Bovine  ---N------ ----L----- -V-----KV- -N-------- -S---D---- ---
                      C-helix             D-helix
```

FIGURE 7: Amino acid sequences of human, goat and bovine α-lactalbumin. Amino acids different from those of human α-lactalbumin are shown in the sequences of goat and bovine α-lactalbumin. Secondary structure regions are also shown at the bottom of the sequences (A-helix, 5−11; B-helix, 23−34; C-helix, 86−98; D-helix, 105−110; and *β*-sheets, 41−43, 48−50, and 55−56).

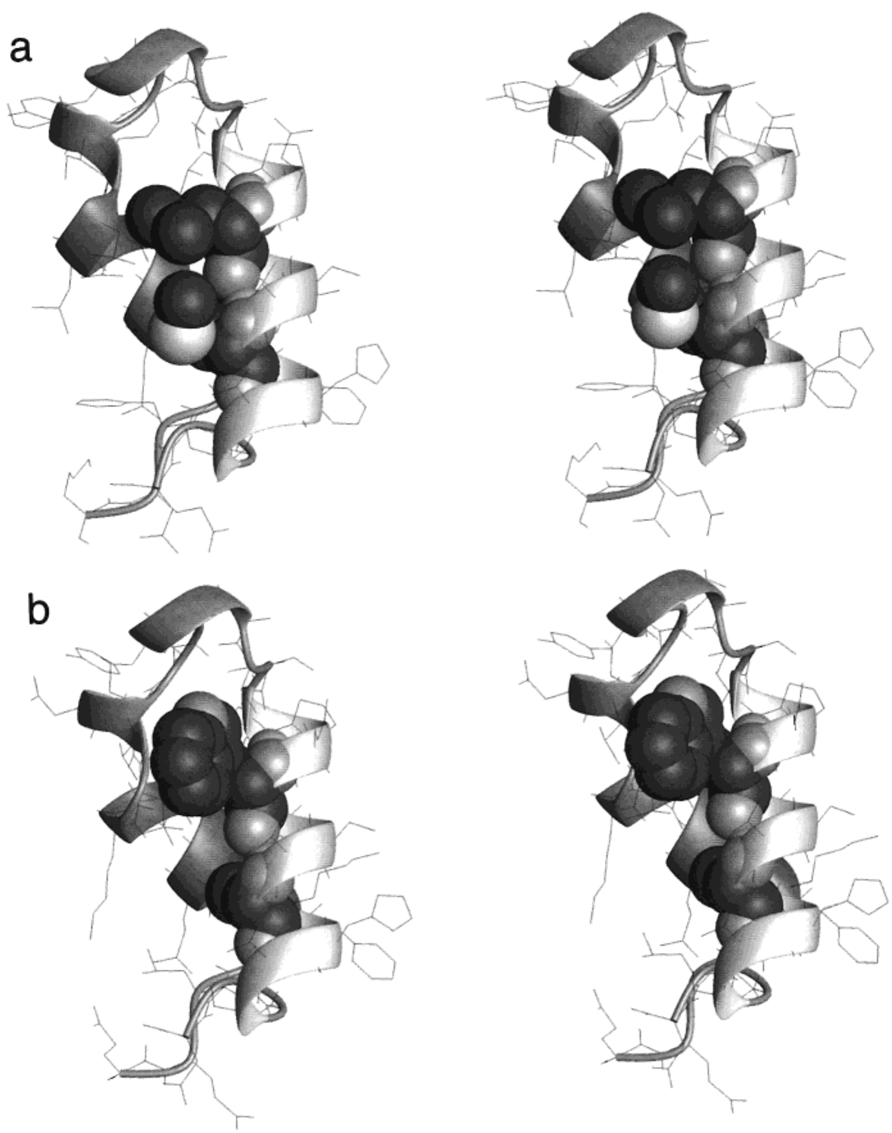

FIGURE 8: Stereoviews of the region of the A and B helices of (a) human α-lactalbumin (PDB code 1A4V) (*43*) and (b) goat α-lactalbumin (the molecule A in a PDB entry 1HFY) (*6*). The A and B helices are shown by dark-gray and white ribbons, respectively. The heavy atoms of residues 26 (Leu for human and Trp for goat) and 30 (Met for human and Ala for goat) are represented by the CPK model, and the other residues are shown by the wire model. The figure was drawn using MOLMOL 2K.1 (*44*).

is less tightly packed in the human protein (*4*−*6*). As a result, the region between the A and B helices has remarkably different backbone topology (Figure 8).

It is thus very likely that these differences in the structural architecture of the α-domain hydrophobic core may be responsible for the decreased N-state stability and the

increased unfolding rate in human α-lactalbumin. This inference is consistent with our recent study of mutants of human and goat α-lactalbumin, in which Thr29, which is also located in the hydrophobic core of the α-domain, has been replaced by isoleucine (unpublished data). This hydrophobic substitution at position 29 increased the N-state stability of the goat protein by as much as 3.5 kcal/mol, and this stabilization increase was almost entirely ascribed to an decrease in the unfolding rate constant (*37*). Amino acid substitutions in the hydrophobic core of the α-domain remarkably change the N-state stability of α-lactalbumin, and these changes in stability are ascribable to changes in the unfolding rate constant without effecting a significant change in the refolding rate, which represents the rate of refolding from the molten globule intermediate to the native state. It is thus likely that the hydrophobic core of the α-domain does not form a folding initiation site in α-lactalbumin, although it is the most hydrophobic portion of the α-lactalbumin molecule.

ACKNOWLEDGMENT

We thank Professor Peter S. Kim for his generous gift of the pHLAC plasmid. We are also grateful to Professor H. Kataoka and Dr. S. Nagata (Department of Biotechnology, Faculty of Agriculture & Life Sciences, University of Tokyo) for their useful discussions and assistance in amino acid sequence analysis and mass spectrometric measurements.

REFERENCES

1. Hill, R. L., and Brew, K. (1975) *Adv. Enzymol. Relat. Areas Mol. Biol. 43*, 411−490.
2. McKenzie, H. A., and White, F. H., Jr. (1991) *Adv. Protein Chem. 41*, 173−315.
3. Acharya, K. R., Stuart, D. I., Walker, N. P. C., Lewis, M., and Phillips, D. C. (1989) *J. Mol. Biol. 208*, 99−127.
4. Acharya, K. R., Ren, J., Stuart, D. I., Phillips, D. C., and Fenna, R. E. (1991) *J. Mol. Biol. 221*, 571−581.
5. Ren, J., Stuart, D. I., and Acharya, K. R. (1993) *J. Biol. Chem. 268*, 19292−19298.
6. Pike, A. C. W., Brew, K., and Acharya, K. R. (1996) *Structure 4*, 691−703.
7. Chaudhuri, T. K., Horii, K., Yoda, T., Arai, M., Nagata, S., Terada, T. P., Uchiyama, H., Ikura, T., Tsumoto, K., Kataoka, H., Matsushima, M., Kuwajima, K., and Kumagai, I. (1999) *J. Mol. Biol. 285*, 1179−1194.
8. Kuwajima, K. (1989) *Proteins 6*, 87−103.
9. Kuwajima, K. (1996) *FASEB J. 10*, 102−109.
10. Arai, M., and Kuwajima, K. (1996) *Folding Des. 1*, 275−287.
11. Arai, M., and Kuwajima, K. (2000) *Adv. Protein Chem. 53*, 209−282.
12. Kuwajima, K. (1996) in *Circular dichroism and the conformational analysis of biomolecules* (Fasman, G. D., Ed.) pp 159−182. Plenum, New York.
13. Peng, Z., and Kim, P. S. (1994) *Biochemistry 33*, 2136−2141.
14. Schulman, B. A., Redfield, C., Peng, Z. Y., Dobson, C. M., and Kim, P. S. (1995) *J. Mol. Biol. 253*, 651−657.
15. Ishikawa, N., Chiba, T., Chen, L. T., Shimizu, A., Ikeguchi, M., and Sugai, S. (1998) *Protein Eng. 11*, 333−335.
16. Veprintsev, D. B., Narayan, M., Permyakov, S. E., Uversky, V. N., Brooks, C. L., Cherskaya, A. M., Permyakov, E. A., and Berliner, L. J. (1999) *Proteins 37*, 65−72.
17. Ikeguchi, M., Kuwajima, K., and Sugai, S. (1986) *J. Biochem. (Tokyo) 99*, 1191−1201.
18. Kuwajima, K., Mitani, M., and Sugai, S. (1989) *J. Mol. Biol. 206*, 547−561.
19. Uchiyama, H., Perez-Prat, E. M., Watanabe, K., Kumagai, I., and Kuwajima, K. (1995) *Protein Eng. 8*, 1153−1161.
20. Nozaka, M., Kuwajima, K., Nitta, K., and Sugai, S. (1978) *Biochemistry 17*, 3753−3758.
21. Peng, Z., Wu, L. C., and Kim, P. S. (1995) *Biochemistry 34*, 3248−3252.
22. Pace, C. N. (1986) *Methods Enzymol. 131*, 266−280.
23. Mizuguchi, M., Arai, M., Ke, Y., Nitta, K., and Kuwajima, K. (1998) *J. Mol. Biol. 283*, 265−277.
24. Redfield, C., Schulman, B. A., Milhollen, M. A., Kim, P. S., and Dobson, C. M. (1999) *Nat. Struct. Biol. 6*, 948−952.
25. Hiraoka, Y., Segawa, T., Kuwajima, K., Sugai, S., and Murai, N. (1980) *Biochem. Biophys. Res. Commun. 95*, 1098−1104.
26. Mizuguchi, M., Masaki, K., Demura, M., and Nitta, K. (2000) *J. Mol. Biol. 298*, 985−995.
27. Imoto, T., Yamada, H., Yasukochi, T., Yamada, E., Ito, Y., Ueda, T., Nagatani, H., Miki, T., and Horiuchi, T. (1987) *Protein Eng. 1*, 333−338.
28. Mine, S., Ueda, T., Hashimoto, Y., and Imoto, T. (1997) *Protein Eng. 10*, 1333−1338.
29. Takano, K., Tsuchimori, K., Yamagata, Y., and Yutani, K. (1999) *Eur. J. Biochem. 266*, 675−682.
30. Hargrove, M. S., Krzywda, S., Wilkinson, A. J., Dou, Y., Ikeda-Saito, M., and Olson, J. S. (1994) *Biochemistry 33*, 11767−11775.
31. Ohage, E., and Steipe, B. (1999) *J. Mol. Biol. 291*, 1119−1128.
32. Ganesh, C., Banerjee, A., Shah, A., and Varadarajan, R. (1999) *FEBS Lett. 454*, 307−311.
33. Suciu, D., and Inouye, M. (1996) *Mol. Microbiol. 21*, 181−195.
34. Duverger, N., Murry-Brelier, A., Latta, M., Reboul, S., Castro, G., Mayaux, J. F., Fruchart, J. C., Taylor, J. M., Steinmetz, A., and Denefle, P. (1991) *Eur. J. Biochem. 201*, 373−383.
35. Schultz, D. A., and Baldwin, R. L. (1992) *Protein Sci. 1*, 910−916.
36. Uversky, V. N., Abdullaev, Z. K., Arseniev, A. S., Bocharov, E. V., Dolgikh, D. A., Latypov, R. F., Melnik, T. N., Vassilenko, S., and Kirpichnikov, M. P. (1999) *Biochim. Biophys. Acta 1432*, 324−332.
37. Yoda, T., Saito, M., Arai, M., Horii, K., Tsumoto, K., Matsushima, M., Kumagai, I., and Kuwajima, K. (2000) *Proteins* (in press).
38. Matouschek, A., Kellis, J. T., Jr., Serrano, L., and Fersht, A. R. (1989) *Nature 340*, 122−126.
39. Hall, L., Craig, R. K., Edbrooke, M. R., and Campbell, P. N. (1982) *Nucleic Acids Res. 10*, 3503−3515.
40. Wang, M., Scott, W. A., Rao, K. R., Udey, J., Conner, G. E., and Brew, K. (1989) *J. Biol. Chem. 264*, 21116−21121.
41. Hurley, W. L., and Schuler, L. A. (1987) *Gene 61*, 119−122.
42. Kumagai, I., Tamaki, E., Kakinuma, S., and Miura, K. (1987) *J. Biochem.* (*Tokyo*) *101*, 511−517.
43. Chandra, N., Brew, K., and Acharya, K. R. (1998) *Biochemistry 37*, 4767−4772.
44. Koradi, R., Billeter, M., and Wüthrich, K. (1996) *J. Mol. Graphics 14*, 51−55.

BI001735J