Protein Science (1998), 7:1564–1574. Cambridge University Press. Printed in the USA.

Transition state in the folding of α-lactalbumin probed by the 6–120 disulfide bond

MASAMICHI IKEGUCHI,[1] MOTOAKI FUJINO,[2] MASAO KATO,[1] KUNIHIRO KUWAJIMA,[3] AND SHINTARO SUGAI[1]

[1]Department of Bioengineering, Faculty of Engineering, Soka University, 1-236 Tangi-cho, Hachioji, Tokyo 192, Japan
[2]Department of Polymer Science, Faculty of Science, Hokkaido University, Kita-ku, Sapporo, Hokkaido 060, Japan
[3]Department of Physics, School of Science, University of Tokyo, 7-3-1 Hongo, Bunkyou-ku, Tokyo 113, Japan

(RECEIVED November 17, 1997; ACCEPTED April 1, 1998)

Abstract

The guanidine hydrochloride concentration dependence of the folding and unfolding rate constants of a derivative of α-lactalbumin, in which the 6–120 disulfide bond is selectively reduced and S-carboxymethylated, was measured and compared with that of disulfide-intact α-lactalbumin. The concentration dependence of the folding and unfolding rate constants was analyzed on the basis of the two alternative models, the intermediate-controlled folding model and the multiple-pathway folding model, that we had proposed previously. All of the data supported the multiple-pathway folding model. Therefore, the molten globule state that accumulates at an early stage of folding of α-lactalbumin is not an obligatory intermediate. The cleavage of the 6–120 disulfide bond resulted in acceleration of unfolding without changing the refolding rate, indicating that the loop closed by the 6–120 disulfide bond is unfolded in the transition state. It is theoretically shown that the chain entropy gain on removing the cross-link from a random coil chain with helical stretches can be comparable to that from an entirely random chain. Therefore, the present result is not inconsistent with the known structure in the molten globule intermediate. Based on this result and other knowledge obtained so far, the structure in the transition state of the folding reaction of α-lactalbumin is discussed.

Keywords: chain entropy; folding kinetics; lysozyme; molten globule; stopped flow

Characterization of the transition state of a folding reaction is an essential step in understanding the mechanism of protein folding. Information about the transition state can be obtained only from the rate constant of the reaction, and any techniques to investigate the protein structure, such as X-ray crystallography, NMR, and CD, are not applicable to structural characterization of the transition state. Protein engineering techniques have been successfully applied to probe the structure in the transition state of folding of barnase, chymotrypsin inhibitor 2 (Fersht, 1995), T4 lysozyme (Chen et al., 1992), staphylococcal nuclease (Kalnin & Kuwajima, 1995), Arc repressor (Milla et al., 1995), and CheY (Lopez-Hernandez & Serrano, 1996). Although most mutations have been designed to examine if the particular interaction between side chains that is present in the final native structure is present in the transition state, the engineered disulfide bond is unique in providing another type of information on the transition state; that is, it can report whether the backbone in a loop region connected by the disulfide bond is relaxed or fixed in the transition state (Strausberg et al., 1993; Clarke & Fersht, 1993). Of course, the native disulfide bond can be used as a probe when it can be removed without disrupting the native structure (Goto & Hamaguchi, 1982; Eyles et al., 1994; Denton et al., 1994).

α-Lactalbumin is a small globular protein of which the folding mechanism, especially the molten globule intermediate, has been investigated extensively (Kuwajima, 1989, 1996). It has four disulfide bonds, Cys6–Cys120, Cys28–Cys111, Cys61–Cys77, and Cys73–Cys91, one of which (Cys6–Cys120) is reduced with extreme rapidity (Kuwajima et al., 1990; Gohda et al., 1995). Previously, we have prepared a derivative of bovine α-lactalbumin, in which the 6–120 disulfide bond is selectively reduced and S-carboxymethylated (2CM-3SS-BLA), and have shown that 2CM-3SS-BLA maintains the native-like structure (Kuwajima et al., 1990; Ikeguchi et al., 1992). Contributions of the 6–120 disulfide to the stability of the native (N) and molten globule (A) states have been estimated by comparing the guanidine hydrochloride (GdnHCl)-induced unfolding between disulfide-intact bovine

Reprint requests to: Masamichi Ikeguchi, Department of Bioengineering, Faculty of Engineering, Soka University, 1-236 Tangi-cho, Hachioji, Tokyo 192, Japan; e-mail: ikeguchi@t.soka.ac.jp.

Abbreviations: NMR, nuclear magnetic resonance; CD, circular dichroism; BLA, bovine α-lactalbumin, 2CM-3SS-BLA, a derivative of bovine α-lactalbumin, in which the 6–120 disulfide bond is selectively reduced and S-carboxymethylated; 4CM-2SS-BLA, a derivative of bovine α-lactalbumin, in which the 6–120 and 28–111 disulfide bonds are selectively reduced and S-carboxymethylated; HEWL, hen egg white lysozyme; GdnHCl, guanidine hydrochloride.

α-lactalbumin (BLA) and 2CM-3SS-BLA (Ikeguchi et al., 1992). The stabilization of the N state relative to the unfolded (U) state by the 6–120 disulfide has been explained by a combined effect of the destabilization of the U state due to the chain entropy loss and the destabilization of the N state due to a strain imposed on the 6–120 disulfide (Kuwajima et al., 1990; Ikeguchi et al., 1992). This conclusion has been supported by a recent calorimetric study (Hendrix et al., 1996). In the present study, we have investigated the transition state in the folding reaction of BLA by comparing the folding–unfolding kinetics of 2CM-3SS-BLA with that of intact BLA. The folding–unfolding kinetics of the two proteins were analyzed on the basis of two alternative models of the BLA folding, (1) a three-state model with the molten globule state as an obligatory folding intermediate (intermediate-controlled folding model) and (2) an apparently two-state model that involves noncooperative folding from the fully unfolded state to the molten globule state (multiple-pathway folding model). We show that the kinetic data for the two forms of BLA are better represented by the second model, suggesting that the molten globule state of BLA may not necessarily be an obligatory intermediate for the folding of this protein. The structural characteristics of this protein in the transition state of folding will be discussed.

Results and discussion

Two alternative models of the BLA folding

It has been known that the molten globule intermediate accumulates at an early stage of the folding of both BLA and 2CM-3SS-BLA and that the subsequent folding process is explained by the two-state process without any other stable intermediates, except for a minor slow folding phase observed in the presence of Ca^{2+} (Kuwajima et al., 1985, 1989; Ikeguchi et al., 1986, 1992; Balbach et al., 1995, 1996; Arai & Kuwajima, 1996). In this study, we ignore this minor slow phase, because its amplitude is only 10% of the total observable change during the folding and it is ascribable to a slow isomerization process in the unfolded state (Kuwajima et al., 1989). In our previous studies, the accumulation of the molten globule state of disulfide-intact BLA has been interpreted in terms of the two alternative models (Ikeguchi et al., 1986; Kuwajima et al., 1989). In the first model, the A state is assumed to be an obligatory intermediate on the pathway of folding (model 1):

$$\mathrm{U} \underset{}{\overset{K_{AU}}{\rightleftarrows}} \mathrm{A} \underset{k_-}{\overset{k_+}{\rightleftarrows}} \mathrm{N} \qquad (1)$$

where K_{AU} is the equilibrium constant of the A ⇄ U transition, and k_+ and k_- are the rate constants for folding and unfolding, respectively. In this model, the A state is a thermodynamic state and the A ⇄ U transition as well as the N ⇄ A transition is a cooperative two-state transition, so that the observed rate (k_{app}) is expressed by the following equation:

$$k_{app} = \frac{1}{K_{AU} + 1} k_+ + k_- \qquad (2)$$

Previously, we called this model the *intermediate-controlled folding model* (Ikeguchi et al., 1986).

In the alternative model (model 2), the A ⇄ U transition is represented as a noncooperative gradual transition in which various microstates are distributed as

$$\left.\begin{array}{c} \mathrm{D}_n \\ \uparrow\downarrow \\ \mathrm{D}_{n-1} \\ \uparrow\downarrow \\ \vdots \\ \uparrow\downarrow \\ \mathrm{D}_i \\ \uparrow\downarrow \\ \vdots \\ \uparrow\downarrow \\ \mathrm{D}_0 \end{array}\right\} \underset{k_u}{\overset{k_f}{\rightleftarrows}} \mathrm{N} \qquad (3)$$

where D_0 is the most unfolded microstate that has the structural characteristics of the U state, D_n is the most folded one that has the characteristics of the A state, and k_f and k_u are the rate constants in the folding and unfolding directions, respectively. In this model, D_is are microstates that belong to a single thermodynamic state D,[4] and there is no free energy barrier between D_i and D_{i+1}. In this model, the observed folding–unfolding rate constant is related to k_f and k_u as

$$k_{app} = k_f + k_u \qquad (4)$$

which is based on the apparently two-state mechanism. This model was originally proposed by Nozaka et al. (1978) and was previously called the *multiple-pathway folding model* (Ikeguchi et al., 1986). This model is also consistent with a recent energy landscape theory (Bryngelson et al., 1995; Dill & Chan, 1997).

Is the molten globule state obligatory for BLA folding?

In Figure 1, the GdnHCl concentration dependence of the apparent folding–unfolding rate constant (k_{app}) of 2CM-3SS-BLA is compared with that obtained previously for disulfide-intact BLA in the absence and presence of Ca^{2+} (Ikeguchi et al., 1986). Typical V-shape dependence of the apparent folding–unfolding rate on GdnHCl concentration is observed except for apo 2CM-3SS-BLA. Because the minimum of the apparent folding–unfolding rate is generally observed at the midpoint of the equilibrium unfolding transition and the transition midpoint of apo 2CM-3SS-BLA is 0.4 M GdnHCl (Ikeguchi et al., 1992), only a monotonous increase in the apparent folding–unfolding rate is observed for apo 2CM-3SS-BLA over the GdnHCl concentration range investigated (0.35 M to 4 M). For analysis of the denaturant concentration dependence of the folding–unfolding rate constants, it is important to determine which of the above two models should be used. Because the interconversion between the A and U states is rapid

[4] In our previous report (Ikeguchi et al., 1986), "D" was used to denote the unfolded state. However, we had changed the abbreviation to indicate the unfolded state from "D" to "U" (Ikeguchi et al., 1992). In this report, "D" is used to mean the denatured state, which includes both the A and U states.

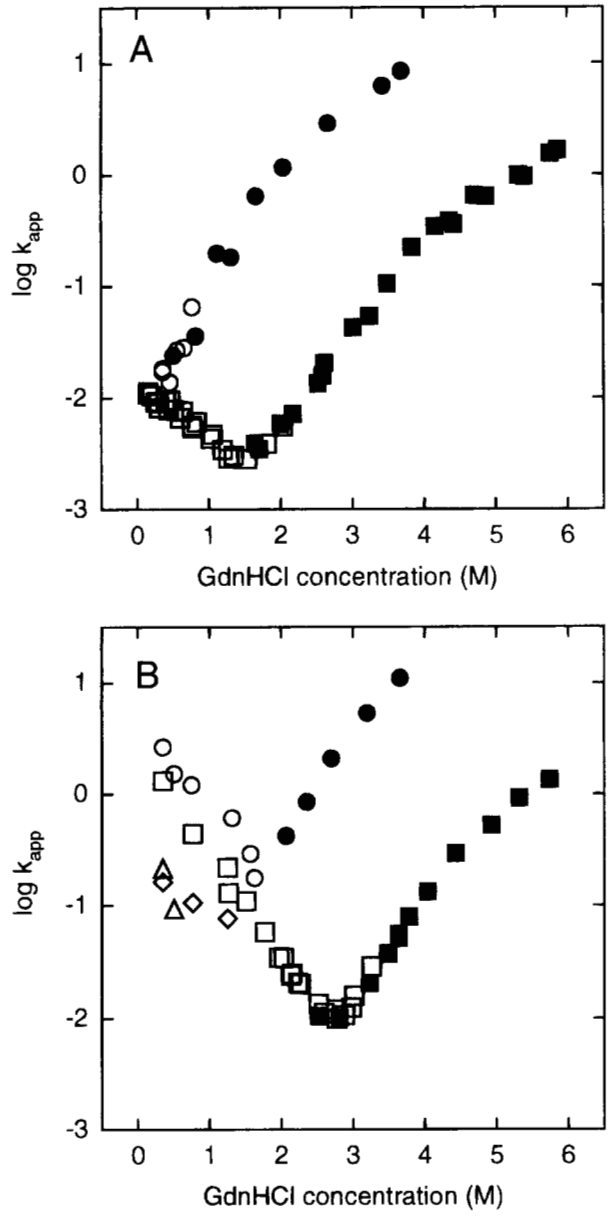

Fig. 1. GdnHCl concentration dependence of the apparent folding–unfolding rate (k_{app}) of 2CM-3SS-BLA (circles) in the absence (**A**) and presence (**B**) of 1 mM Ca^{2+}. The experimental condition, 50 mM cacodylate-50 mM NaCl (pH 7.0) and 4.5 °C. For comparison, k_{app} values of intact BLA (squares) obtained in the previous study (Ikeguchi et al., 1986) are also shown. The k_{app} values obtained from refolding experiments are indicated as open symbols, and those obtained from unfolding experiments are shown as closed symbols. The rate constants of a minor slow phase observed during the refolding of 2CM-3SS-BLA (diamonds) and intact BLA (triangles) in the 1 mM Ca^{2+} and low concentration of GdnHCl are also shown.

and kinetically separated from the subsequent folding process, it is difficult to prove which model is correct. Although the previous studies of BLA folding have assumed model 1, there is also evidence that the A $\rightleftarrows$ U transition is a noncooperative transition, being consistent with model 2 (Shimizu et al., 1993; Griko et al., 1994; Schulman & Kim, 1996; Wilson et al., 1996; Schulman et al., 1997; Pfeil, 1998). Thus, the following analysis has been made for investigating which model is more appropriate.

We have calculated k_+ and k_- on the basis of model 1 and k_f and k_u on the basis of model 2 (see Materials and methods). The results are presented in Figures 2 and 3, respectively. For many proteins that show a simple two-state folding–unfolding transition, the logarithms of the folding and unfolding rate constants are known to show linear dependence on denaturant concentration (see Fersht, 1997 and references cited therein). In the case where folding intermediates accumulate and a rollover is observed in the denaturant dependence of the logarithmic apparent rate constant, the logarithms of the folding and unfolding rate constants of an individual transition show linear dependence on denaturant concen-

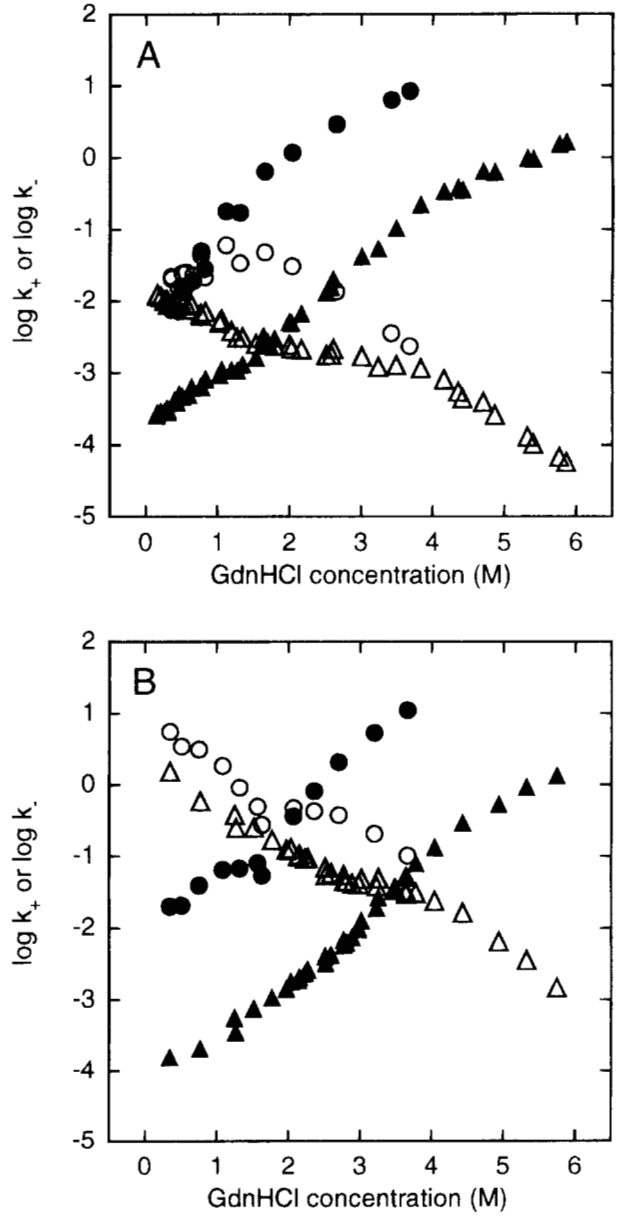

Fig. 2. GdnHCl concentration dependence of the folding (k_+) and unfolding (k_-) rates of 2CM-3SS-BLA (circles) and intact BLA (triangles) in the absence (**A**) and presence (**B**) of 1 mM Ca^{2+} at pH 7.0 and 4.5 °C. Both k_+ (open symbols) and k_- (closed symbols) were calculated assuming that the A state is an obligate intermediate (model 1) (see text).

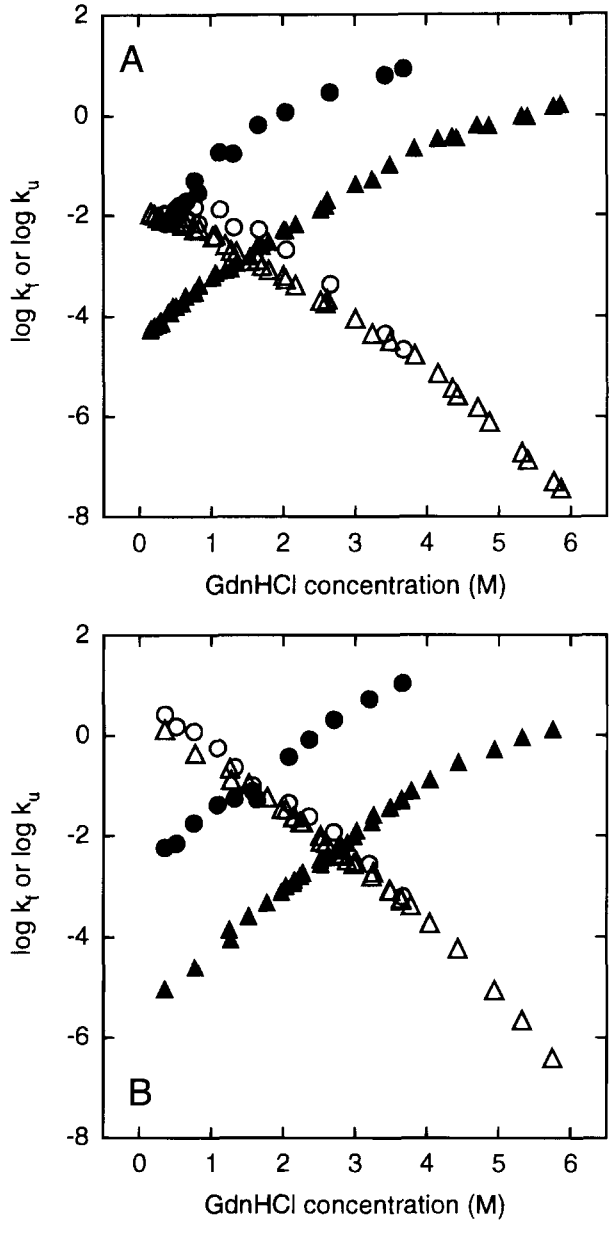

Fig. 3. GdnHCl concentration dependence of the folding (k_f) and unfolding (k_u) rates of 2CM-3SS-BLA (circles) and intact BLA (triangles) in the absence (**A**) and presence (**B**) of 1 mM Ca^{2+} at pH 7.0 and 4.5 °C. Both k_f (open symbols) and k_u (closed symbols) were calculated assuming that the two-state transition between the N and D states (model 2) (see text).

tration (Yamasaki et al., 1995; Wildegger & Kiefhaber, 1997). Therefore, if model 1 is appropriate for describing the folding kinetics of BLA, $\log k_+$ and $\log k_-$ are expected to show linear GdnHCl dependence. Similarly, if model 2 is appropriate, $\log k_f$ and $\log k_u$ may show linear GdnHCl dependence. As seen from Figure 3, $\log k_f$ and $\log k_u$ based on model 2 show linear dependence on GdnHCl concentration below 4 M. Log k_f values of 2CM-3SS-BLA coincide well with those of intact BLA over a wide concentration range of GdnHCl. Although the absolute values of k_u are different between 2CM-3SS-BLA and intact BLA, the GdnHCl dependence of $\log k_u$ is similar between 2CM-3SS-BLA and intact BLA. Therefore, the effect of disulfide bond cleavage on the folding–unfolding rate is nearly independent of GdnHCl concentration. The coincidence of k_f and the difference in k_u between 2CM-3SS-BLA and BLA are explained by a conformational entropy change in the transition state (see below). The effect of Ca^{2+} concentration is also nearly independent of GdnHCl concentration (compare $\log k_f$ between Fig. 3A,B). At least below 4 M GdnHCl, therefore, the present equilibrium and kinetic data for the folding–unfolding of BLA and 2CM-3SS-BLA are well described by model 2.

On the other hand, the plots of $\log k_+$ and $\log k_-$ vs. GdnHCl concentration (model 1) show clear bendings (Fig. 2). Log k_+ of apo 2CM-3SS-BLA increases with a GdnHCl concentration up to 1.3 M and then decreases. Similar changes in the slope of $\log k_+$ are also observed for apo BLA, holo BLA, and holo 2CM-3SS-BLA, although the sign of the slope does not change in these cases. Therefore, the GdnHCl dependence of the folding–unfolding kinetics cannot be appropriately represented by model 1. To explain the bendings in the plot of $\log k_+$ and $\log k_-$ vs. GdnHCl concentration, intermediates other than the A state must be incorporated into model 1. However, there is no evidence for the presence of such intermediates, according to the kinetic folding studies of apo BLA investigated by various techniques including UV absorption, fluorescence, CD, pulse hydrogen exchange (Arai & Kuwajima, 1996), real-time observation of appearance of the NMR signal specific to the N state (Balbach et al., 1995), and 2D NMR observation of folding at individual residue levels (Balbach et al., 1996). All of these experiments show that the observable folding kinetics is represented by a single exponential process with the same rate constant independent of the probe used. This fact rules out the possibility that an intermediate that is not included in model 1 brings about the bendings in the plot of $\log k_+$ and $\log k_-$ vs. GdnHCl concentration.

The appropriateness of model 2 is further supported by the previous hydrogen exchange studies of slowly exchanging tryptophan indole NH protons of apo BLA. Harushima et al. (1988) have shown that the exchange rates of two protected tryptophan (Trp26 and Trp104) indole NH protons of apo BLA are independent of pH. This indicates that the indole NH protons are exchanged with solvent deuteron by the so-called EX1 mechanism, in which the observed exchange rate corresponds to the rate of unfolding to yield an unprotected conformation. If model 1 is valid, the exchange rate k_{ex} should correspond to k_-. Although slight protection of the indole NH proton from exchange has been observed in the A state (unpublished results), the rate limiting step for exchange is N → A conversion because the rate of A → U conversion is very fast. However, k_- at 0 M GdnHCl calculated on the basis of model 1 is 10 times larger than the observed hydrogen exchange rate of apo BLA ($\log k_{ex} = -4.8$). Therefore, model 1 is inconsistent with the hydrogen exchange data. On the other hand, the unfolding rate k_u based on model 2 agrees well with the hydrogen exchange rate (Fig. 3), supporting the validity of model 2. In the following section, therefore, we will discuss the transition state on the basis of model 2.

Model 2 means that the molten globule structure is not obligatory for the folding of α-lactalbumin. Because the unfolding of the molten globule state cannot be represented as a cooperative two-state transition, there is a protein molecule, a part of which assumes the molten globule structure and the other part is unfolded, in the transition zone of the A ⇄ U transition. Such a molecule can directly fold to the native structure without acquiring the complete molten globule structure. However, this does not mean that the molten globule state is not a folding intermediate but a dead-end

species. The fitting of the present data with a model (A ⇄ U ⇄ N), in which the A state is a dead-end species, was found to be even worse than the fitting with model 1 (not shown). Thus, under a physiological condition, the molten globule structure is predominantly formed at an early stage in the folding, and the folding reaction proceeds from it. That is, the molten globule state is phenomenologically an on-pathway folding intermediate. The rate-limiting step of the folding reaction should be a conformational search to attain the transition state structure. If the A state contains at least a part of the transition state structure, the formation of the A state accelerates the subsequent folding process. On the other hand, if the structured part of the protein molecule in the transition state differs from and does not interact with the structured part in the A state, the rate of the subsequent folding process is independent of the formation of the A state. As discussed below, the structured part of the BLA molecule in the transition state seems to be not organized in the molten globule state. Even though the molten globule state is an on-pathway intermediate, therefore, it may neither accelerate nor decelerate the subsequent folding process. Recently, Creighton (1997) also pointed out that the molten globule is not a key to rapid folding of α-lactalbumin.

Effects of the 6–120 disulfide bond cleavage on the transition state

Previously, we have investigated the effect of the 6–120 disulfide on the stability of the native and molten globule states (Ikeguchi et al., 1992). The free energy difference between the N and U states ($\Delta G_{NU}^{H_2O}$) of 2CM-3SS-BLA is smaller by 2.4 and 3.1 kcal/mol than that of intact BLA in the absence and presence of 1 mM Ca^{2+}, respectively. We have shown that these changes in $\Delta G_{NU}^{H_2O}$ caused by the disulfide bond cleavage is attributed to two opposite effects, stabilization of the U state by an increase in chain entropy and stabilization of the N state by relaxing strain imposed on the 6–120 disulfide bond (Kuwajima et al., 1990; Ikeguchi et al., 1992). The chain-entropy increase in the U state has theoretically been estimated to be 17 cal/mol/K, which corresponds to a free energy change of 4.8 kcal/mol at 4.5 °C (Poland & Scheraga, 1965; Lin et al., 1984). The strain energy imposed on the 6–120 disulfide bond has been estimated to be 2.7 kcal/mol at 4.5 °C from the ratio of the reduction rate of the 6–120 disulfide to the reduction rate of disulfides in the urea-unfolded BLA (Kuwajima et al., 1990). In the present study, the free energy change of the N ⇄ D transition in the absence of GdnHCl ($\Delta G_{ND}^{H_2O}$) and its dependence on GdnHCl concentration (m_{ND}) have been obtained from the apparent two-state analysis of our previous data (see Materials and methods). Results are shown in Table 1, and the $\Delta G_{ND}^{H_2O}$ values are essentially the same as the $\Delta G_{NU}^{H_2O}$ values previously reported (Ikeguchi et al., 1992). Recently, Hendrix et al. (1996) have calorimetrically measured the enthalpy (ΔH), entropy (ΔS), and heat capacity (ΔC_p) changes upon thermal unfolding of 2CM-3SS-BLA and intact BLA at pH 8.0 and 2 mM $CaCl_2$. The ΔH, ΔS, and the free energy change (ΔG) for the thermal unfolding at 4.5 °C were calculated from their data and are included in Table 1. The free energy change for the thermal unfolding (ΔG) is consistent with our estimates of $\Delta G_{ND}^{H_2O}$. The difference in the ΔS of the thermal unfolding between 2CM-3SS-BLA and intact BLA is 20 cal/mol/K (Table 1), which is close to the theoretically calculated chain-entropy increase in the random-coil state.

When the transition state structure is probed by the effect of the disulfide-bond cleavage on the folding and unfolding rates, one must first address whether the position of the transition state on the reaction coordinate, i.e., the degree of structural organization in the transition state, is altered by removing the disulfide bond. The position of the transition state on the reaction coordinate is reflected in the β_T value, which is given by

$$\beta_T = \frac{-m_f^{\ddagger}}{m_u^{\ddagger} - m_f^{\ddagger}} = \frac{-m_f^{\ddagger}}{m_{ND}} \tag{5}$$

where $m_f^{\ddagger}$ and $m_u^{\ddagger}$ are constants representing the GdnHCl-concentration dependence of the activation free energies, $\Delta G_f^{\ddagger}$ and $\Delta G_u^{\ddagger}$, for folding and unfolding, respectively (Tanford, 1968; Kuwajima et al., 1989; Matouschek & Fersht, 1993), and these activation free energies are expressed by

$$\Delta G_f^{\ddagger} = \Gamma - RT \ln k_f = \Delta G_f^{\ddagger}(H_2O) - m_f^{\ddagger}[\text{GdnHCl}] \tag{6}$$

$$\Delta G_u^{\ddagger} = \Gamma - RT \ln k_u = \Delta G_u^{\ddagger}(H_2O) - m_u^{\ddagger}[\text{GdnHCl}] \tag{7}$$

Table 1. *Energetic parameters of the folding–unfolding reactions of BLA and 2CM-3SS-BLA*[a]

	apo			holo		
	2CM-3SS-BLA	BLA	Difference	2CM-3SS-BLA	BLA	Difference
$\Delta G_f^{\ddagger}(H_2O) - \Gamma$ (kcal/mol)[b]	1.9 (0.16)	2.2 (0.03)	−0.3 (0.16)	−1.0 (0.11)	−0.7 (0.04)	−0.3 (0.12)
$m_f^{\ddagger}$ (kcal/mol/M)	−0.98 (0.09)	−0.98 (0.02)		−1.34 (0.05)	−1.31 (0.02)	
$\Delta G_u^{\ddagger}(H_2O) - \Gamma$ (kcal/mol)[b]	2.9 (0.16)	5.5 (0.03)	−2.6 (0.16)	3.3 (0.11)	6.8 (0.04)	−3.5 (0.12)
$m_u^{\ddagger}$(kcal/mol/M)	1.23 (0.09)	1.24 (0.02)		1.31 (0.05)	1.45 (0.02)	
β_T	0.45 (0.05)	0.44 (0.01)		0.51 (0.04)	0.48 (0.01)	
$\Delta G_{ND}^{H_2O}$ (kcal/mol)	1.0 (0.19)	3.3 (0.19)	−2.3 (0.27)	4.3 (0.81)	7.5 (0.22)	−3.2 (0.84)
m_{ND} (kcal/mol/M)	2.21 (0.41)	2.22 (0.12)		2.65 (0.47)	2.75 (0.08)	
ΔG (kcal/mol)[c]				4.2	7.2	−3.0
ΔH (kcal/mol)[c]				1.2	9.8	−8.6
ΔS (cal/mol/K)[c]				−10.5	9.5	−20.0

[a]Values in the parentheses are standard errors.
[b]This value is equal to $-RT \ln k_f$ or $-RT \ln k_u$ at 0 M GdnHCl, and is not the absolute activation energy.
[c]Calculated from calorimetric data at pH 8.0 and 2 mM $CaCl_2$ (Hendrix et al., 1996).

where Γ is a constant and $\Delta G_f^{\ddagger}(H_2O)$ and $\Delta G_u^{\ddagger}(H_2O)$ are the activation free energies of folding and unfolding in water. Because the *m* value is proportional to the solvent exposure of groups buried in the native structure (Schellman, 1978; Alonso & Dill, 1991; Myers et al., 1995), β_T gives the relative degree of solvent exposure in the transition state. The *m* values were calculated from the slopes in Figure 3 over a GdnHCl concentration range from 0 to 4 M (Table 1). The β_T values calculated from these *m* values are included in Table 1. 2CM-3SS-BLA and intact BLA show similar β_T values, indicating that the structure in the transition state is not substantially affected by removal of the 6–120 disulfide bond, at least, with regard to the solvent exposure of hydrophobic groups.

Previously, we have estimated the Ca^{2+} binding constant of intact BLA in the transition state from the Ca^{2+} concentration dependence of the folding rate at 0.35 M GdnHCl (Kuwajima et al., 1989). The estimated Ca^{2+} binding constant is 3.1×10^5 M^{-1}, which is one or two orders of magnitude smaller than that in the native state. Although the Ca^{2+} concentration dependence of the folding rate of 2CM-3SS-BLA has not been investigated in detail, the folding rate of 2CM-3SS-BLA coincides with that of intact BLA both in the absence and in the presence of 1 mM Ca^{2+}, indicating that the Ca^{2+} binding constants of 2CM-3SS-BLA and intact BLA in the transition state are similar. It is thus concluded that the position of the transition state in the reaction coordinate, i.e., the structure in the transition state, is not altered by removing the 6–120 disulfide bond.

As seen in Figure 3, the folding rate of 2CM-3SS-BLA is only twice that of intact BLA, whereas the unfolding rate of 2CM-3SS-BLA is 100–500 times that of intact BLA. Therefore, the activation free energy of folding (the free energy difference between the transition and D states) is not changed, and the activation free energy of unfolding (the free energy difference between the transition and N states) is significantly reduced by removal of the 6–120 disulfide. The decrease in $\Delta G_u^{\ddagger}(H_2O)$ is estimated to be 2.6 kcal/mol in the absence of Ca^{2+} and 3.5 kcal/mol in the presence of 1 mM Ca^{2+} (Table 1). Based on these data, a free energy diagram for the folding reaction of 2CM-3SS-BLA and intact BLA in water was constructed (Fig. 4). The following changes in the free energy levels of the N, D, and transition states may occur when the 6–120 disulfide is removed. First, the free energy of the D state is lowered by 4.8 kcal/mol because of the chain-entropy increase. Although the D state in the native condition is the molten globule state and cannot be regarded as a random coil, we have calculated the chain-entropy increment in the D state on the assumption that the BLA molecule in the D state is approximated to be a random coil when calculating the chain-entropy increment. As shown below, this approximation is valid for the 6–120 disulfide bond of BLA. Second, because $\Delta G_f^{\ddagger}(H_2O)$ of 2CM-3SS-BLA is only 0.3 kcal/mol smaller than that of intact BLA, the free energy of the transition state must be lowered by 5.1 kcal/mol. On the other hand, the free energy of the N state is lowered by 2.7 kcal/mol because of elimination of the strain imposed on the 6–120 disulfide bond. Because $\Delta G_u^{\ddagger}(H_2O)$ of 2CM-3SS-BLA is 2.6 kcal/mol smaller than that of intact BLA, the free energy of the transition state must be lowered by 5.3 kcal/mol. Thus, the change in the free energy level of the transition state evaluated from the unfolded side (D state) is consistent with that evaluated from the native side (N state). It is also notable that the decrease in the free energy of the transition state is approximately equal to that in the D state, so that it may arise from the increase in the chain entropy. This means that, in the transition state, the loop closed by the 6–120 disulfide (i.e., the polypeptide chains from Cys6 to Cys28 and from Cys111 to Cys120 connected by the disulfide bond Cys28–Cys111) is unfolded to the extent that the chain can be regarded as a random coil in terms of the chain entropy.

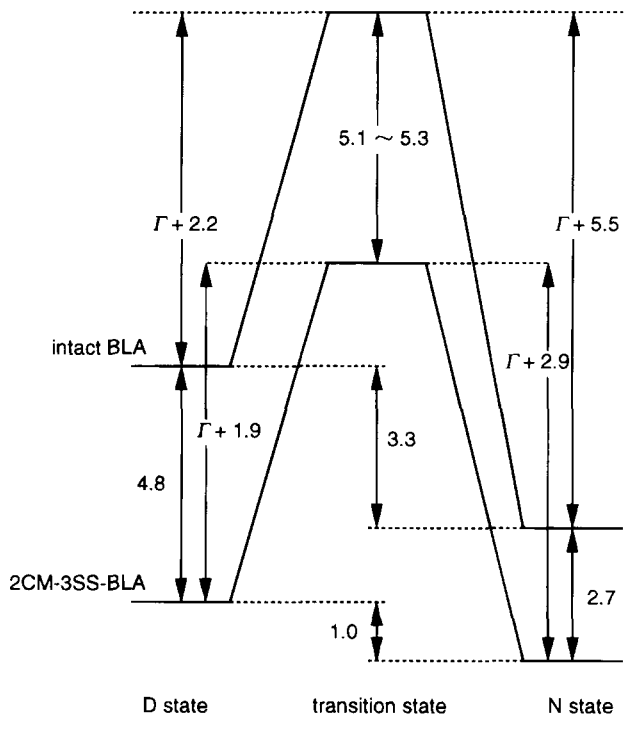

Fig. 4. Free energy profile of the folding and unfolding reactions of 2CM-3SS-BLA and intact BLA in the absence of Ca^{2+} and GdnHCl (pH 7.0, 4.5 °C). Free energy differences between states are given in kcal/mol (see Table 1).

Structure in the transition state

Figure 5 shows a schematic representation of the three-dimensional structure of native α-lactalbumin. The α-lactalbumin molecule is composed of two subdomains, an α-domain and a β-domain (Pike et al., 1996). The Ca^{2+} binding site of the protein is located at a hinge region that connects the two subdomains. Previously, we have shown that the native-like tertiary structure is formed around the Ca^{2+} binding site in the transition state of folding (Kuwajima et al., 1989). The loop closed by the 6–120 disulfide is located in the α-domain. The free energy profile in Figure 4 indicates that, in the transition state, this loop region is unfolded to the extent that the conformational freedom of the region is nearly the same as that in the U state, because the increase in chain entropy upon the 6–120 disulfide cleavage in the transition state is very close to that in the U state. Apparently, this seems to be inconsistent with the known structure of α-lactalbumin in the molten globule (A) state. It has been shown by hydrogen exchange (Baum et al., 1989; Alexandrescu et al., 1993; Chyan et al., 1993; Schulman et al., 1995) and proline scanning mutagenesis (Schulman & Kim, 1996) that the A, B, D, and 3_{10} helices in the α-domain are formed in the A state. Furthermore, disulfide exchange experiments have shown that the α-domain has a native-like backbone topology in the A state (Wu et al., 1995). These results suggest that the conforma-

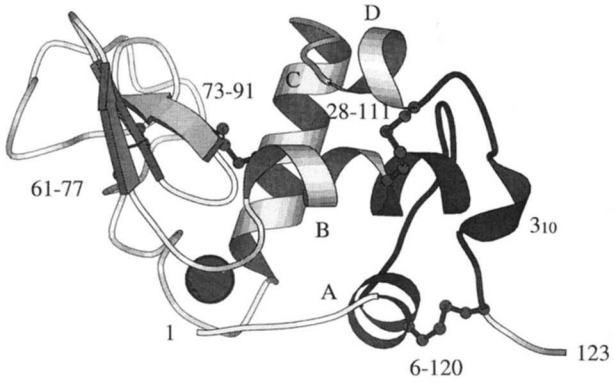

Fig. 5. Schematic representation of the three-dimensional structure of α-lactalbumin based on the coordinate of the molecule D in a PDB entry "1hfz" (Pike et al., 1996). Four disulfide bonds are shown as ball and stick models. Bound Ca^{2+} is shown as a large sphere. The loop region closed by the 6–120 disulfide (sequences from Cys6 to Cys28 and from Cys111 to Cys120) is shaded. Prepared with MOLSCRIPT (Kraulis, 1991).

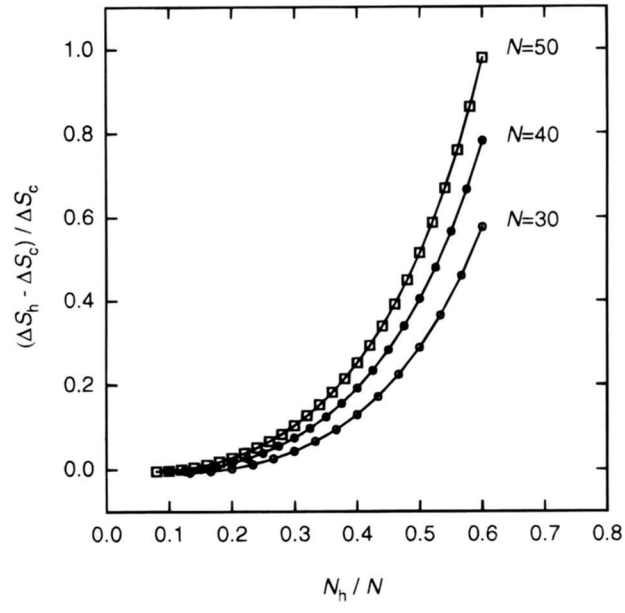

Fig. 6. Effect of a helix on the entropy loss upon cross-linking. The effect of a helical stretch on the entropy loss upon cross-linking (ΔS_h-ΔS_c) was calculated with Equation 25 and represented as a relative value to the entropy loss upon introducing a cross-link into the random coil chain (ΔS_c), which was calculated according to the method of Poland and Scheraga (1965). It was calculated as a function of the fraction of helical units in the loop formed by the cross-link for the cases that the number of the total units in the loop is 30, 40, and 50.

tional freedom in the α-domain is considerably restricted in the A state. Is the helical structure formed in the A state disrupted in the transition state? Our previous study has shown that the stability of the A state relative to the U state is reduced only by 1.1 kcal/mol upon cleaving the 6–120 disulfide bond (Ikeguchi et al., 1992). Because the free energy level of the U state is lowered by 4.8 kcal/mol when the 6–120 disulfide is cleaved, the free energy level of the A state must be lowered by 3.7 kcal/mol due to the cleavage. We have thus previously concluded that this drop in the free energy level of the A state is due to a chain entropy increase in the A state, although its magnitude is slightly smaller than that in the U state (Ikeguchi et al., 1992). That is, the polypeptide chain in the loop closed by the 6–120 disulfide bond has a considerable freedom not only in the transition state but also in the A state. This is apparently inconsistent with the observations that some helices are formed in the loop closed by the 6–120 disulfide bond in the A state. In order to solve this apparent inconsistency, we have theoretically investigated the effect of a helix on the chain entropy increment upon removing a cross-link (see Materials and methods). A simple theoretical model indicates that the maximum effect of helices on the chain entropy increment upon removing a cross-link from a random chain with helical stretches is given by the following equation:

$$-\Delta S_h - (-\Delta S_c) = -\frac{3}{2} R \left\{ \ln \frac{N}{N_c} - \frac{(N_h h)^2}{b^2 N_c} \right\} \qquad (8)$$

where ΔS_h and ΔS_c are the entropy loss upon introducing a cross-link into a random chain with helical streches and into an entirely random chain, respectively. N_h, N_c, and N are the numbers of helical, coil, and total units, respectively, in the loop closed by the cross-link ($N_h + N_c = N$), b is a length of a statistical unit, and h is a helical pitch along the helical axis. The qualitative conclusion drawn from this equation is that the chain entropy increment upon breaking the cross-link is larger for the random chain with a helical stretch rather than for the entirely random chain. The longer the helix is, the larger the effect on the chain entropy increment is (Fig. 6). If the A (5–11), B (23–34), D (105–110) helices, and a 3_{10}-helix (115–118) are assumed to be formed in the A state of α-lactalbumin (Schulman et al., 1995; Schulman & Kim, 1996), there are 13 helical units and 20 coil units in the loop closed by the 6–120 disulfide (6–28 and 111–120). By substituting these values into Equation 8, we obtain $-\Delta S_h - (-\Delta S_c) = 2.4$ cal/mol/K. This is only 15% of the entropy increment upon removing the 6–120 disulfide from the random coil ($-\Delta S_c = 17.3$ cal/mol/K). Therefore, the amount of helices present in the loop closed by the 6–120 disulfide bond does not significantly affect the chain entropy increase calculated by assuming a random coil. Similarly, even if the structure of the α-domain in the transition state is similar to that in the A state, the chain entropy increase occurring in the transition state may be comparable to that in the U state. The free energy profile shown in Figure 4 is thus not inconsistent with the known structure of α-lactalbumin in the A state. More studies, however, may be required for further characterizing the structure of α-lactalbumin in the transition state.

The structural features of the transition state in the folding of BLA obtained so far may be similar to those of a two-disulfide derivative of α-lactalbumin in which both Cys6–Cys120 and Cys28–Cys111 are removed (Ewbank & Creighton, 1993a, 1993b; Hendrix et al., 1996; Ikeguchi et al., 1996; Wu et al., 1996). The two-disulfide species (2SS-BLA) and its S-carboxymethylated derivative of BLA (4CM-2SS-BLA) have been known to bind Ca^{2+} (Ewbank & Creighton, 1993a; Hendrix et al., 1996; Ikeguchi et al., 1996). The apparent Ca^{2+} binding constant has been estimated to be 6×10^4 M^{-1} (Ewbank & Creighton, 1993a; Ikeguchi et al., unpubl. results), and it is close to the Ca^{2+} binding constant of intact BLA in the transition state (Kuwajima et al., 1989). Hendrix et al. (1996) have investigated the thermal unfolding of 4CM-2SS-BLA and have shown that 4CM-2SS-BLA unfolds cooperatively

with significantly lower values of enthalpy and entropy changes than those of intact BLA. They have thus concluded that, in Ca^{2+}-bound 4CM-2SS-BLA, the β-domain maintains a native-like fold, whereas the α-domain is unfolded. Ikeguchi et al. (1996) have investigated the structure of Ca^{2+}-bound 4CM-2SS-BLA by ^1H NMR and have shown that the residues within the β-domain and the interface between the α- and β-domains (including the C helix) show chemical shifts similar to those in the native protein. Wu et al. (1996) have also shown that the mutant of human α-lactalbumin lacking Cys6, Cys28, Cys111, and Cys120 (α-LA(β)) preferentially forms the native disulfide bond in the presence of Ca^{2+}, whereas it forms all of the possible disulfide bonds in the absence of Ca^{2+}. All of these results indicate that the β-domain of BLA can fold independently when it binds Ca^{2+}, suggesting that the β-domain may be folded in the transition state of intact BLA and that 4CM-2SS-BLA may be a good analog of the transition state in the folding reaction of BLA.

Comparison with a homologous protein, lysozyme

The folding–unfolding kinetics of hen egg white lysozyme (HEWL) has been extensively investigated by many researchers (Kato et al., 1980, 1981; Segawa & Sugihara, 1984a, 1984b; Kuwajima et al., 1985; Ikeguchi et al., 1986; Chaffotte et al., 1992; Radford et al., 1992; Miranker et al., 1993; Eyles et al., 1994; Denton et al., 1994; Itzhaki et al., 1994; Kiefhaber, 1995; Parker et al., 1995; Motoshima et al., 1996; Wildegger & Kiefhaber, 1997, Matagne et al., 1997). At acid pH, the folding kinetics of HEWL is similar to that of BLA, that is, the molten globule-like intermediate is formed in the burst phase, and the subsequent formation of the native conformation is a two-state process without additional intermediates. Under such conditions, the transition state structure has been probed by effects of the solvents, chemical modifications, inhibitor binding, and site-directed mutagenesis (Segawa & Sugihara, 1984a, 1984b; Motoshima et al., 1996). Segawa and Sugihara (1984b) have shown that the cross-link between Glu35 and Trp108 significantly accelerates the folding rate of HEWL without a large change in the unfolding rate. This result indicates that the loops closed by the cross-link (a large loop from Glu35 to Trp108 and a small loop composed by Cys30–Glu35 and Trp108–Cys115) are folded in the transition state. This loop region contains all of the β-domain and the C helix in the α-domain. Motoshima et al. (1996) have inferred the transition state structure from the analysis of unfolding kinetics of Gly-Pro and Pro-Gly mutants. They have shown that the 101–102 region, which is located at the C terminal of the C helix, and the 117–118 region, which is located between the D and 3_{10} helices in the α-domain, are unfolded, whereas the residue 47 at the β-sheet and the residues 121–122 at the 3_{10} helix in the α-domain are partially native-like in the transition state. Because the presence of inhibitor does not affect the folding rate and reduces the unfolding rate, the inhibitor binding site at the cleft between the α- and β-domains does not have a complete structure to bind the inhibitor (Segawa & Sugihara, 1984b).

Under the condition where the native state is highly stable, the folding kinetics of HEWL is complex and there are parallel folding channels (Radford et al., 1992; Miranker et al., 1993; Itzhaki et al., 1994; Kiefhaber, 1995; Matagne et al., 1997; Wildegger & Kiefhaber, 1997). On the major folding pathway, there is an intermediate in which amide protons in the α-domain are highly protected and tryptophan fluorescence is highly quenched, although the corresponding intermediate has not been observed in the folding of BLA. Under this condition, Eyles et al. (1994) and Denton et al. (1994) have investigated the effect of removal of the 6–127 disulfide, which corresponds to the 6–120 disulfide of BLA, on the folding kinetics. In the folding of a derivative of HEWL, in which the 6–127 disulfide is reduced and S-carboxymethylated, the intermediate with persistent structure in the α-domain is no longer observed, although the molten globule-like intermediate is formed in the burst phase as in the intact protein. Therefore, the folding kinetics of the derivative lacking the 6–127 disulfide is similar to that of intact HEWL at acid pH and that of BLA. In spite of this change in the folding kinetics, the rate of the formation of the native structure of the HEWL derivative is nearly identical with that of intact HEWL. This result indicates that the activation free energy of the final folding step is not changed by removing the 6–127 disulfide. Therefore, the loop closed by the 6–127 disulfide is mobile in the transition state, even though the α-helices are formed in this region. These results indicate that the transition state in the folding of HEWL is similar to that of BLA, consistent with the idea that the folding mechanism is conserved in the homologous proteins.

Materials and methods

Experimental procedures

Preparation of 2CM-3SS-BLA and stopped-flow CD experiments were performed as described previously (Ikeguchi et al., 1992). The unfolded 2CM-3SS-BLA in 4 M GdnHCl for refolding experiments or the native 2CM-3SS-BLA in the buffer for unfolding experiments was mixed with the buffer containing various concentrations of GdnHCl to give various final GdnHCl concentrations.

Analyses of the folding–unfolding kinetics

In model 1, k_+ and k_- can be calculated from k_{app} (Equation 2) and the equilibrium constants of the N $\rightleftarrows$ A transition,

$$K_{NA} = \frac{k_-}{k_+} = \frac{K_{NU}}{K_{AU}} \tag{9}$$

and the A $\rightleftarrows$ U transition (K_{AU}). From Equations 2 and 9, we obtain

$$k_+ = \frac{K_{AU}(K_{AU}+1)}{K_{NU}(K_{AU}+1)+K_{AU}} k_{app} \tag{10}$$

$$k_- = \frac{K_{NU}(K_{AU}+1)}{K_{NU}(K_{AU}+1)+K_{AU}} k_{app} \tag{11}$$

By assuming the linear dependence of ΔG_{NU} and ΔG_{AU} on GdnHCl concentration, we can obtain K_{NU} and K_{AU} at any GdnHCl concentration with the following equations:

$$\Delta G_{NU} = -RT \ln K_{NU} = \Delta G_{NU}^{H_2O} - m_{NU}[\text{GdnHCl}] \tag{12}$$

$$\Delta G_{AU} = -RT \ln K_{AU} = \Delta G_{AU}^{H_2O} - m_{AU}[\text{GdnHCl}] \tag{13}$$

where $\Delta G_{NU}^{H_2O}$ and $\Delta G_{AU}^{H_2O}$ are the free energy changes of the N $\rightleftarrows$ U and A $\rightleftarrows$ U transitions at 0 M GdnHCl, respectively. The

m_{NU} and m_{AU} are constants of proportionality for the dependence of free energy changes on denaturant concentration. These parameters have been evaluated from equilibrium unfolding experiments (Ikeguchi et al., 1986, 1992). The unfolding transition of the kinetic burst phase intermediate coincides well with the A $\rightleftarrows$ U transition calculated from $\Delta G_{AU}^{H_2O}$ and m_{AU} (Ikeguchi et al., 1986, 1992; Arai & Kuwajima, 1996).

In model 2, k_f and k_u can be calculated from k_{app} (Equation 4) and the equilibrium constant between the N and D states,

$$K_{ND} = \frac{k_u}{k_f} \tag{14}$$

that is obtained by reanalyzing the equilibrium unfolding curve previously measured by the aromatic CD at 270 nm as a two-state transition between the N and D states. It is assumed that $\ln K_{ND}$ is a linear function of GdnHCl concentration:

$$\Delta G_{ND} = -RT \ln K_{ND} = \Delta G_{ND}^{H_2O} - m_{ND}[\text{GdnHCl}] \tag{15}$$

where ΔG_{ND} and $\Delta G_{ND}^{H_2O}$ are the free energy changes between the N and D states at a given concentration of GdnHCl and 0 M GdnHCl, respectively. The m_{ND} is a constant of proportionality for the dependence of free energy changes on denaturant concentration. These parameters for BLA and 2CM-3SS-BLA are listed in Table 1. From Equations 4 and 14, k_f and k_u are calculated as

$$k_f = \frac{1}{1 + K_{ND}} k_{app} \tag{16}$$

$$k_u = \frac{K_{ND}}{1 + K_{ND}} k_{app} \tag{17}$$

Entropic effect of a cross-link on the random chain with α-helical stretch

It is well known that the introduction of cross-links in a statistical coil reduces its conformational entropy (Jacobson & Stockmayer, 1950; Schellman, 1955; Flory, 1956; Poland & Scheraga, 1965; Lin et al., 1984; Chan & Dill, 1989, 1990). This entropic effect is known as a major origin of the stabilization of the native conformation of proteins by disulfide bonds (Betz, 1993). Here, we consider the entropy change that occurs in a random coil chain with a helical stretch upon introduction of a cross-link. As a starting point, we regard the polypeptide chain that consists of $N + 1$ residues as a statistical chain of N units, where the statistical unit is a vector connecting adjacent C_α atoms and its length is represented as b. For a sufficiently long chain, the distribution of the end-to-end distance of the chain, $P(r)$, is known to obey a Gaussian function:

$$P(r) = \left(\frac{3}{2\pi b^2 N}\right)^{3/2} \exp\left(-\frac{3r^2}{2b^2 N}\right) \tag{18}$$

where r is the end-to-end distance, and the N-terminal of the chain is fixed at the origin. When we use the Cartesian coordinate, Equation 18 is expressed by

$$P(x,y,z) = \left(\frac{3}{2\pi b^2 N}\right)^{3/2} \exp\left\{-\frac{3(x^2 + y^2 + z^2)}{2b^2 N}\right\} \tag{19}$$

where x, y, and z are x-, y-, and z-axis components of r, and $x^2 + y^2 + z^2 = r^2$. For the chain with a helical stretch, we consider the chain of which the first N_h units are in the helical conformation and the remaining N_c units are in the coil conformation ($N_h + N_c = N$). If the x-axis is taken to coincide with the helix axis and the origin is set to the N-terminal of the helix, the distribution function is

$$P(x,y,z) = \left(\frac{3}{2\pi b^2 N_c}\right)^{3/2} \exp\left\{-\frac{3[(x - N_h h)^2 + y^2 + z^2]}{2b^2 N_c}\right\} \tag{20}$$

where h is a helical pitch along the helical axis.

The entropy loss is calculated from the possibility that the C-terminal end of a chain will occur in the space where the cross-link is possible. Previously, the space where the C-terminal end should be in is assumed to be a shell (Schellman, 1955; Poland & Scheraga, 1965; Lin et al., 1984) or a sphere (Pace et al., 1988; Harrison & Sternberg, 1994). Here, for convenience, we assume that this space is a cube $\Delta x \Delta y \Delta z$ of which body center is the origin. The entropy change ΔS is then given by

$$\Delta S = R \ln \int_{-\Delta x/2}^{\Delta x/2} \int_{-\Delta y/2}^{\Delta y/2} \int_{-\Delta z/2}^{\Delta z/2} P(x,y,z)\, dx\, dy\, dz \tag{21}$$

Using an approximation that the integration in Equation 21 is equal to the product of the mean probability $P(0,0,0)$ and the volume element $\Delta x \Delta y \Delta z$, we have

$$\Delta S = R \ln\{P(0,0,0)\Delta x \Delta y \Delta z\} \tag{22}$$

By substituting Equation 19 into Equation 22, we obtain the entropy loss upon introducing a cross-link in an entirely random chain

$$\Delta S_c = \frac{3}{2} R \ln \frac{3}{2\pi b^2} - \frac{3}{2} R \ln N + \frac{3}{2} R \ln \Delta x \Delta y \Delta z \tag{23}$$

By substituting Equation 20 into Equation 22, we obtain the entropy loss upon introducing a cross-link in a chain with a helix

$$\Delta S_h = \frac{3}{2} R \ln \frac{3}{2\pi b^2} - \frac{3}{2} R \ln N_c - \frac{3R(N_h h)^2}{2b^2 N_c} + \frac{3}{2} R \ln \Delta x \Delta y \Delta z \tag{24}$$

When we consider the effect of a helical stretch on the chain entropy loss upon introducing a cross-link, a debatable term, $\Delta x \Delta y \Delta z$, is eliminated, and we have

$$\Delta S_h - \Delta S_c = \frac{3}{2} R \left\{\ln \frac{N}{N_c} - \frac{(N_h h)^2}{b^2 N_c}\right\} \tag{25}$$

A simulation with $b = 3.8$ Å and $h = 1.5$ Å has revealed that the sign of Equation 25 is negative except for the case where a very short helical section and a long coil section exist (for example, if

$N_h = 6$, then N_c must be larger than 43 to obtain a positive value of $\Delta S_h - \Delta S_c$). The negative value of $\Delta S_h - \Delta S_c$ means that the entropy loss upon cross-linking is larger for the chain having a helical stretch than for the entirely random chain. In other words, cross-linking is more difficult for the chain with a helical stretch than for the entirely flexible chain. In Figure 6, the effect of a helical stretch on the entropy loss upon cross-linking, $(\Delta S_h - \Delta S_c)/\Delta S_c$, is calculated as a function of the fraction of helical residues. It is clear that the longer the helix in the loop, the larger the entropy loss is. This means that the cross-linking destabilizes the long helix in the loop. In the case where there are several helical stretches in the loop formed by the cross-link, the second term of Equation 25 depends on the relative orientation of the helices. Clearly, its magnitude is maximum when all helices have the same direction. Therefore, Equation 25 can be used to estimate the maximum of the influence of several helical stretches on the entropy loss when a cross-link is introduced.

References

Alexandrescu AT, Evans PA, Pitkeathly M, Baum J, Dobson CM. 1993. Structure and dynamics of the acid-denatured molten globule state of α-lactalbumin: A 2-dimensional NMR study. *Biochemistry 32*:1707–1718.

Alonso DOV, Dill KA. 1991. Solvent denaturation and stabilization of globular proteins. *Biochemistry 30*:5974–5985.

Arai M, Kuwajima K. 1996. Rapid formation of a molten globule intermediate in refolding of α-lactalbumin. *Folding Design 1*:275–287.

Balbach J, Forge V, Lau WS, van Nuland NAJ, Brew K, Dobson CM. 1996. Protein folding monitored at individual residues during a two-dimensional NMR experiment. *Science 274*:1161–1163.

Baum J, Dobson CM, Evans PA, Hanley C. 1989. Characterization of a partly folded protein by NMR methods: Studies on the molten globule state of guinea pig α-lactalbumin. *Biochemistry 28*:7–13.

Balbach J, Forge V, van Nuland NAJ, Winder SL, Hore PJ, Dobson CM. 1995. Following protein folding in real time using NMR spectroscopy. *Nature Struct Biol 2*:865–870.

Betz SF. 1993. Disulfide bonds and the stability of globular proteins. *Protein Sci 2*:1551–1558.

Bryngelson JD, Onuchic JN, Socci ND, Wolynes PG. 1995. Funnels, pathways, and the energy landscape of protein folding: A synthesis. *Proteins Struct Funct Genet 21*:167–195.

Chaffotte AF, Guillou Y, Goldberg ME. 1992. Kinetic resolution of peptide bond and side chain far-UV circular dichroism during the folding of hen egg white lysozyme. *Biochemistry 31*:9694–9702.

Chan HS, Dill KA. 1989. Intrachain loops in polymers: Effects of excluded volume. *J Chem Phys 90*:492–509.

Chan HS, Dill KA. 1990. The effects of internal constraints on the configurations of chain molecules. *J Chem Phys 92*:3118–3135.

Chen BL, Baase WA, Nicholson H, Schellman JA. 1992. Folding kinetics of T4 lysozyme and nine mutants at 12 °C. *Biochemistry 31*:1464–1476.

Chyan CL, Wormald C, Dobson CM, Evans PA, Baum J. 1993. Structure and stability of the molten globule state of guinea-pig α-lactalbumin: A hydrogen exchange study. *Biochemistry 32*:5681–5691.

Clarke J, Fersht AR. 1993. Engineered disulfide bonds as probes of the folding pathway of barnase: Increasing the stability of proteins against the rate of denaturation. *Biochemistry 32*:4322–4329.

Creighton TE. 1997. How important is the molten globule for correct protein folding? *Trends Biochem Sci 22*:6–10.

Denton ME, Rothwarf DM, Scheraga HA. 1994. Kinetics of folding of guanidine-denatured hen egg white lysozyme and carboxymethyl(Cys(6),Cys(l27))-lysozyme: A stopped-flow absorbance and fluorescence study. *Biochemistry 33*:11225–11236.

Dill KA, Chan HS. 1997. From Levinthal to pathway to funnels. *Nature Struct Biol 4*:10–19.

Ewbank JJ, Creighton TE. 1993a. Pathway of disulfide-coupled unfolding and refolding of bovine α-lactalbumin. *Biochemistry 32*:3677–3693.

Ewbank JJ, Creighton TE. 1993b. Structural characterization of the disulfide folding intermediates of bovine α-lactalbumin. *Biochemistry 32*:3694–3707.

Eyles SJ, Radford SE, Robinson CV, Dobson CM. 1994. Kinetic consequences of the removal of a disulfide bridge on the folding of hen lysozyme. *Biochemistry 33*:13038–13048.

Fersht AR. 1995. Characterizing transition states in protein folding: An essential step in the puzzle. *Curr Opin Struct Biol 5*:79–84.

Fersht AR. 1997. Nucleation mechanisms in protein folding. *Curr Opin Struct Biol 7*:3–9.

Flory PJ. 1956. Theory of elastic mechanisms in fibrous proteins. *J Am Chem Soc 78*:5222–5235.

Gohda S, Shimizu A, Ikeguchi M, Sugai S. 1995. The superreactive disulfide bonds in α-lactalbumin and lysozyme. *J Protein Chem 14*:731–737.

Goto Y, Hamaguchi K. 1982. Unfolding and refolding of the reduced constant fragment of the immunoglobulin light chain. Kinetic role of the intrachain disulfide bond. *J Mol Biol 156*:911–926.

Griko YV, Freire E, Privalov PL. 1994. Energetics of the α-lactalbumin states: A calorimetric and statistical thermodynamic study. *Biochemistry 33*:1889–1899.

Harrison PM, Sternberg MJE. 1994. Analysis and classification of disulfide connectivity in proteins: The entropic effect of cross-linkage. *J Mol Biol 244*:448–463.

Harushima Y, Kuwajima K, Sugai S. 1988. Hydrogen exchange of the tryptophan residues in bovine α-lactalbumin studied by UV spectroscopy. *Biopolymers 27*:629–644.

Hendrix TM, Griko Y, Privalov P. 1996. Energetics of structural domains in α-lactalbumin. *Protein Sci 5*:923–931.

Ikeguchi M, Hoshino N, Hirata T, Shimizu A, Sugai S. 1996. Structure of the two disulfide folding intermediate of bovine α-lactalbumin. *Prog Biophys Mol Biol 65S1*:58.

Ikeguchi M, Kuwajima K, Mitani M, Sugai S. 1986. Evidence for identity between the equilibrium unfolding intermediate and a transient folding intermediate: A comparative study of the folding reactions of α-lactalbumin and lysozyme. *Biochemistry 25*:6965–6972.

Ikeguchi M, Sugai S, Fujino M, Sugawara T, Kuwajima K. 1992. Contribution of the 6–120-disulfide bond of α-lactalbumin to the stabilities of its native and molten globule states. *Biochemistry 31*:12695–12700.

Itzhaki LS, Evans PA, Dobson CM, Radford SE. 1994. Tertiary interactions in the folding pathway of hen lysozyme: Kinetic studies using fluorescent probes. *Biochemistry 33*:5212–5220.

Jacobson H, Stockmayer WH. 1950. Intramolecular reaction in polycondensations. I. The theory of linear systems. *J Chem Phys 18*:1600–1606.

Kalnin NN, Kuwajima K. 1995. Kinetic folding and unfolding of staphylococcal nuclease and its six mutants studied by stopped-flow circular dichroism. *Proteins Struct Funct Genet 23*:163–176.

Kato S, Okamura M, Shimamoto N, Utiyama H. 1980. Spectral evidence for a rapidly formed structural intermediate in the refolding kinetics of hen egg-white lysozyme. *Biochemistry 20*:1080–1085.

Kato S, Shimamoto N, Utiyama H. 1981. Identification and characterization of the direct folding process of hen egg-white lysozyme. *Biochemistry 21*:38–43.

Kiefhaber T. 1995. Kinetic traps in lysozyme folding. *Proc Natl Acad Sci USA 92*:9029–9033.

Kraulis PJ. 1991. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J Appl Cryst 24*:946–950.

Kuwajima K. 1989. The molten globule state as a clue for understanding the folding and cooperativity of globular-protein structure. *Proteins Struct Funct Genet 6*:87–103.

Kuwajima K. 1996. The molten globule state of α-lactalbumin. *FASEB J 10*:102–109.

Kuwajima K, Hiraoka Y, Ikeguchi M, Sugai S. 1985. Comparison of the transient folding intermediates in lysozyme and α-lactalbumin. *Biochemistry 24*:847–881.

Kuwajima K, Ikeguchi M, Sugawara T, Hiraoka Y, Sugai S. 1990. Kinetics of disulfide bond reduction in α-lactalbumin by dithiothreitol and molecular basis of superreactivity of the Cys6–Cys120 disulfide bond. *Biochemistry 29*:8240–8249.

Kuwajima K, Mitani M, Sugai S. 1989. Characterization of the critical state in protein folding. Effects of guanidine hydrochloride and specific Ca^{2+} binding on the folding kinetics of α-lactalbumin. *J Mol Biol 206*:574–561.

Lin SH, Konishi Y, Denton ME, Scheraga HA. 1984. Influence of an extrinsic cross-link on the folding pathway of ribonuclease A. Conformational and thermodynamic analysis of cross-linked (lysine^7-lysine^{41})-ribonuclease A. *Biochemistry 23*:5504–5512.

Lopez-Hernandez E, Serrano L. 1996. Structure of the transition state for folding of the 129 aa protein CheY resembles that of a smaller protein, CI-2. *Folding Design 1*:43–55.

Matagne A, Radford SE, Dobson CM. 1997. Fast and slow tracks in lysozyme folding: Insight into the role of domains in the folding process. *J Mol Biol 267*:1068–1074.

Matouschek A, Fersht AR. 1993. Application of physical organic chemistry to engineered mutants of proteins: Hammond postulate behavior in the transition state of protein folding. *Proc Natl Acad Sci USA 90*:7814–7818.

Milla ME, Brown BM, Waldburger CD, Sauer RT. 1995. P22 arc repressor: Transition state properties inferred from mutational effects on the rates of protein unfolding and refolding. *Biochemistry 34*:13914–13919.

Miranker A, Robinson CV, Radford SE, Aplin RT, Dobson CM. 1993. Detection of transient protein folding populations by mass spectrometry. *Science 262*:896–900.

Motoshima H, Ueda T, Imoto T. 1996. Analysis of the transition state in the unfolding of hen lysozyme by introduction of Gly-Pro and Pro-Gly sequences at the same site. *J Biochem 119*:1019–1023.

Myers JK, Pace CN, Scholtz JM. 1995. Denaturant *m* values and heat capacity changes: Relation to changes in accessible surface areas of protein unfolding. *Protein Sci 4*:2138–2148.

Nozaka M, Kuwajima K, Nitta K, Sugai S. 1978. Detection and characterization of the folding pathway of human α-lactalbumin. *Biochemistry 17*:3753–3758.

Pace CN, Grimsley GR, Thomson JA, Barnett B. 1988. Conformational stability and activity of ribonuclease T_1 with zero, one, and two intact disulfide bonds. *J Biol Chem 263*:11820–11825.

Parker MJ, Spencer J, Clarke AR. 1995. An integrated kinetic analysis of intermediates and transition states in protein folding reactions. *J Mol Biol 253*:771–786.

Pfeil W. 1998. Is the molten globule a third thermodynamic state of protein? The example of alpha-lactalbumin. *Proteins Struct Funct Genet 30*:43–48.

Pike ACW, Brew K, Acharya KR. 1996. Crystal structures of guinea-pig, goat and bovine α-lactalbumin highlight the enhanced conformational flexibility of regions that are significant for its action in lactose synthase. *Structure 4*:691–703.

Poland DC, Scheraga HA. 1965. Statistical mechanics of noncovalent bonds in polyamino acids. VIII. Covalent loops in proteins. *Biopolymers 3*:379–399.

Radford SE, Dobson CM, Evans PA. 1992. The folding of hen lysozyme involves partially structured intermediates and multiple pathways. *Nature 358*:302–307.

Schellman JA. 1955. The stability of hydrogen-bonded peptide structures in aqueous solution. *Compt Rend Lab Carlsberg Ser Chim 29*:230–259.

Schellman JA. 1978. Solvent denaturation. *Biopolymers 17*:1305–1322.

Schulman BA, Kim PS. 1996. Proline scanning mutagenesis of a molten globule reveals non-cooperative formation of a protein's overall topology. *Nature Struct Biol 3*:682–687.

Schulman BA, Kim PS, Dobson CM, Redfield C. 1997. A residue-specific NMR view of the non-cooperative unfolding of a molten globule. *Nature Struct Biol 4*:630–634.

Schulman BA, Redfield C, Peng ZY, Dobson CM, Kim PS. 1995. Different subdomains are most protected from hydrogen exchange in the molten globule and native states of human α-lactalbumin. *J Mol Biol 253*:651–657.

Segawa S, Sugihara M. 1984a. Characterization of the transition state of lysozyme unfolding. I. Effect of protein-solvent interactions on the transition state. *Biopolymers 23*:2473–2488.

Segawa S, Sugihara M. 1984b. Characterization of the transition state of lysozyme unfolding. II. Effects of the intrachain crosslinking and the inhibitor binding on the transition state. *Biopolymers 23*:2489–2498.

Shimizu A, Ikeguchi M, Sugai S. 1993. Unfolding of the molten globule state of α-lactalbumin studied by ^{1}H NMR. *Biochemistry 32*:13198–13203.

Strausberg S, Alexander P, Wang L, Gallagher T, Gilliland G, Bryan P. 1993. An engineered disulfide cross-link accelerates the refolding rate of calcium-free subtilisin by 850-fold. *Biochemistry 32*:10371–10377.

Tanford C. 1968. Protein denaturation. *Adv Protein Chem 23*:121–282.

Wildegger G, Kiefhaber T. 1997. Three-state model for lysozyme folding: Triangular folding mechanism with an energetically trapped intermediate. *J Mol Biol 270*:294–304.

Wilson G, Hecht L, Barron LD. 1996. The native-like tertiary fold in the molten globule α-lactalbumin appears to be controlled by a continuous phase transition. *J Mol Biol 261*:341–347.

Wu LC, Peng ZY, Kim PS. 1995. Bipartite structure of the α-lactalbumin molten globule. *Nature Struct Biol 2*:281–286.

Wu LC, Schulman BA, Peng ZY, Kim PS. 1996. Disulfide determinants of calcium-induced packing in α-lactalbumin. *Biochemistry 35*:859–863.

Yamasaki K, Ogasahara K, Yutani K, Oobatake M, Kanaya S. 1995. Folding pathway of *Escherichia coli* ribonuclease HI: A circular dichroism, fluorescence, and NMR study. *Biochemistry 34*:16552–16562.