Pathway of Oxidative Folding of α-Lactalbumin: A Model for Illustrating the Diversity of Disulfide Folding Pathways†

Jui-Yoa Chang* and Li Li

Research Center for Protein Chemistry, Institute of Molecular Medicine and Department of Biochemistry and Molecular Biology, The University of Texas, Houston, Texas 77030

Received February 26, 2002; Revised Manuscript Received April 19, 2002

ABSTRACT: The pathway of oxidative folding of α-lactalbumin (αLA) (four disulfide bonds) has been characterized by structural and kinetic analysis of the acid-trapped folding intermediates. In the absence of calcium, oxidative folding of αLA proceeds through highly heterogeneous species of one-, two-, three-, and four-disulfide (scrambled) intermediates to reach the native structure. In the presence of calcium, the folding intermediates of αLA comprise two predominant isomers (αLA-**IIA** and αLA-**IIIA**) adopting exclusively native disulfide bonds, including the two disulfide bonds ($Cys^{61}-Cys^{77}$ and $Cys^{73}-Cys^{91}$) located within the β-sheet calcium binding domain. αLA-**IIA** is a two-disulfide species consisting of $Cys^{61}-Cys^{77}$ and $Cys^{73}-Cys^{91}$ disulfide bonds. αLA-**IIIA** contains $Cys^{61}-Cys^{77}$, $Cys^{73}-Cys^{91}$, and $Cys^{28}-Cys^{111}$ disulfide bonds. The underlying mechanism of the contrasting folding pathways of calcium-bound and calcium-depleted αLA is congruent with the cause of diversity of disulfide folding pathways observed among many well-characterized three-disulfide proteins, including bovine pancreatic trypsin inhibitor and hirudin. Our study also reveals novel aspects of the folding mechanism of αLA that have not been described previously.

One of the well-established methods for analysis of protein folding pathways is oxidative folding of disulfide-containing proteins (*1*−*8*). Using this method, proteins are first reduced and denatured and then allowed to refold in the presence of redox buffers that promote disulfide bond formation and shuffling. The pathway of refolding is subsequently tracked and analyzed by the mechanism of formation of native disulfide bonds. For instance, a protein that contains three disulfide bonds can potentially adopt 75 different disulfide isomers (15 one-disulfide, 45 two-disulfide, and 15 three-disulfide isomers). Among them are seven isomers that contain exclusively native disulfide bonds (three one-disulfide, three two-disulfide, and one three-disulfide isomers). All these isomers could potentially exist during oxidative folding. The disulfide folding pathway is characterized and defined by the heterogeneity and structures of these isomers that accumulate during the process of oxidative folding that leads to the formation of the species containing three native disulfide bonds.

Application of the technique of oxidative folding has allowed elucidation of folding pathways of several three-disulfide proteins, including the most extensively investigated model of bovine pancreatic trypsin inhibitor (BPTI) (*1*, *2*, *7*−*9*) as well as hirudin (*10*), tick anticoagulant peptide (TAP) (*11*, *12*), potato carboxypeptidase inhibitor (PCI) (*13*, *14*), epidermal growth factor (EGF) (*15*, *16*), insulin-like growth factor (*17*, *18*), etc. The results have shown that even among small three-disulfide proteins, the folding mechanisms may vary substantially. These differences are illustrated by (a) the extent of the heterogeneity of folding intermediates, (b) the predominance of intermediates containing native disulfide bonds, and (c) the accumulation of three-disulfide scrambled isomers as intermediates. The underlying cause of this diversity was subsequently shown to associate with the fashion of how local structural elements of disulfide proteins are being stabilized and can be predicted from the mechanism of reductive unfolding of these proteins (*19*). Those with their native disulfide bonds reduced in a collective and simultaneous manner typically exhibit both a high degree of heterogeneity of folding intermediates and the accumulation of scrambled isomers along the folding pathway (e.g., hirudin and PCI) (*10*, *13*, *14*). A sequential reduction of native disulfide bonds is indicative of the existence of stable structural domains. In these cases (e.g., BPTI and EGF), limited types of intermediates adopting nativelike structures tend to dominate the folding pathway (*7*, *8*, *16*).

Analysis of the folding pathway of four-disulfide proteins presents yet another level of technical challenge due to the exponential increase in the number of disulfide isomers. There are 28, 210, 420, and 104 species of one-, two-, three-, and four-disulfide isomers, respectively, that may potentially serve as folding intermediates. Even if only a fraction of these isomers exist along the pathway, their separation and identification will still amount to a daunting task. One of the most extensively investigated models is bovine ribonuclease A. Scheraga and co-workers (*20*−*23*) showed that the folding pathway of ribonuclease A comprises both hetero-

† This work was supported by the endowment of the Robert Welch Foundation.

* To whom correspondence should be addressed: Institute of Molecular Medicine, 2121 W. Holcombe Blvd., Houston, TX 77030. Telephone: (713) 500-2458. Fax: (713) 500-2424. E-mail: Rowen.Chang@uth.tmc.edu.

10.1021/bi020169k CCC: $22.00
Published on Web 06/05/2002

geneous intermediates and three-disulfide species adopting native disulfide bonds. Another four-disulfide protein that has been characterized in detail is α-lactalbumin (αLA),[1] which also consists of complex folding intermediates and predominant species containing three native disulfide bonds (*24*−*26*). However, in the previous studies, folding conditions were limited to defined sets of redox buffer and folding intermediates of αLA were mainly analyzed by gel electrophoresis. Here we reexamine the pathway of oxidative folding of αLA in varying redox buffers. We also analyze the folding intermediates of αLA using reversed-phase high-pressure liquid chromatography (RP-HPLC), which has been shown to resolve more than 50 different isoforms of scrambled αLA (*27*, *28*). Our study reveals novel aspects of the folding mechanism of αLA that have not been observed previously.

αLA represents an attractive four-disulfide model of oxidative folding for another important reason. αLA comprises an α-helical domain and a β-sheet calcium binding domain. The β-sheet domain is known to be considerably stabilized upon binding to calcium (*24*, *25*, *29*) during the folding. Analysis of the folding pathway(s) of αLA both in the absence and in the presence of calcium thus may provide valuable information for corroborating the underlying mechanism that has been proposed to account for the diversity of folding pathways observed among many three-disulfide proteins (*19*).

EXPERIMENTAL PROCEDURES

Materials. Calcium-depleted bovine αLA (L-6010) was used throughout this study and was obtained from Sigma. The protein was further purified by HPLC and was shown to be more than 97% pure. Thermolysin (P-1512), dithiothreitol, reduced and oxidized glutathione, 2-mercaptoethanol, and GdmCl were also purchased from Sigma with purities of greater than 99%. All buffers that were used for folding experiments were degassed with a water pump for 30 min and kept in a closed system whenever they were not being used.

Nomenclature of Four-Disulfide Scrambled Isomers of αLA. Scrambled species of αLA are designated by the following formula: X-αLA-(species assigned on HPLC), where X stands for scrambled. For instance, X-αLA-a represents species a of scrambled αLA.

Oxidative Folding of Fully Reduced αLA. The native αLA (2 mg/mL) was first reduced and denatured in Tris-HCl buffer (0.1 M, pH 8.4) containing 5 M GdmCl and 30 mM dithiothreitol. Reduction and denaturation were carried out at 22 °C for 90 min. To initiate folding, the sample was passed through a PD-10 column (Sephadex-25, Pharmacia) equilibrated in 0.1 M Tris-HCl buffer (pH 8.4). Reduced and denatured αLA was recovered in a volume of 1.1 mL, which was immediately diluted with the same Tris-HCl buffer to a final protein concentration of 0.5 mg/mL (42 μM) and with selected concentrations of redox agents. Folding experiments were conducted both in the absence and in the presence $CaCl_2$ (5 mM). Folding intermediates of αLA were trapped in a time course manner using two different methods. (a) Aliquots of the folding sample were mixed with an equal volume of 4% trifluoroacetic acid in water. Trapped folding intermediates were analyzed by HPLC or stored at −20 °C. (b) Aliquots of the folding sample were mixed with an equal volume of Tris-HCl buffer (0.1 M, pH 8.4) containing vinylpyridine (0.26 M). The reaction was carried out at 23 °C for 45 min. Vinylpyridine-derivatized samples were purified by gel filtration (PD-10 column) equilibrated in 0.5% aqueous trifluoroacetic acid.

Analysis of the Folding Intermediates by Reversed-Phase HPLC. Analysis and isolation of folding intermediates of αLA were achieved by reversed-phase HPLC using the following conditions. Solvent A for HPLC was water containing 0.1% trifluoroacetic acid. Solvent B was acetonitrile and water (9:1, v/v) containing 0.086% trifluoroacetic acid. The gradient was from 22 to 37% solvent B in 15 min and from 37 to 56% solvent B from 15 to 60 min. The flow rate was 0.5 mL/min. The column was a Zorbax 300SB C-18 column for peptides and proteins (4.6 mm, 5 μm). The column temperature was 23 °C.

*Structural Analysis of Folding Intermediates αLA-**IIA** and αLA-**IIIA**.* The HPLC fractions of αLA-**IIA** and αLA-**IIIA** (20 μg) were isolated, freeze-dried, and derivatized with 50 μL of vinylpyridine (0.13 M) in Tris-HCl bufer (0.1 M, pH 7.5) at 23 °C for 45 min. Vinylpyridine-derivatized samples were further purified by HPLC, freeze-dried, and treated with 2 μg of thermolysin (Sigma, P-1512) in 65 μL of *N*-ethylmorpholine/acetate buffer (50 mM, pH 6.4). Digestion was carried out at 37 °C for 16 h. Peptides were then isolated by HPLC and analyzed by amino acid sequencing and mass spectrometry to identify the disulfide-containing peptides.

Reductive Unfolding. The native αLA (L-6010) (0.5 mg/mL) was dissolved in Tris-HCl buffer (0.1 M, pH 8.4) containing varying concentrations of dithiothreitol (0.01−30 mM) both in the absence and in the presence of $CaCl_2$ (5 mM). Reduction was carried out at 23 °C for 10 min or in a time course manner, quenched with an equal volume of 4% aqueous trifluoroacetic acid, and analyzed by HPLC. The samples were stored at −20 °C.

Amino Acid Sequencing and Mass Spectrometry. The amino acid sequences of disulfide-containing peptides were analyzed by automatic Edman degradation using a Perkin-Elmer Procise sequencer (model 494) equipped with an on-line PTH-amino acid analyzer. The molecular masses of disulfide-containing peptides were determined with a MALDI-TOF mass spectrometer (Perkin-Elmer Voyager-DE STR).

RESULTS

Oxidative Folding of αLA in the Absence of Redox Buffer. The reduced and denatured αLA was first allowed to refold in Tris-HCl buffer in the absence and presence 2-mercaptoethanol (0.2 mM), which represent the simplest folding conditions. Under these conditions, air oxidation by residual oxygen accounts for the formation of disulfide bonds. Folding intermediates were trapped in a time course manner either by acid quenching or by reaction with vinylpyridine. They were then characterized for their heterogeneity and chromatographic behavior by HPLC, and distributions of disulfide species by MALDI mass spectrometry.

The HPLC patterns of acid-trapped intermediates remain similar regardless of whether 2-mercaptoethanol is present.

[1] Abbreviations: αLA, α-lactalbumin; X-αLA, scrambled α-lactalbumin; HPLC, high-pressure liquid chromatography; GSH, reduced glutathione; GSSG, oxidized glutathione.

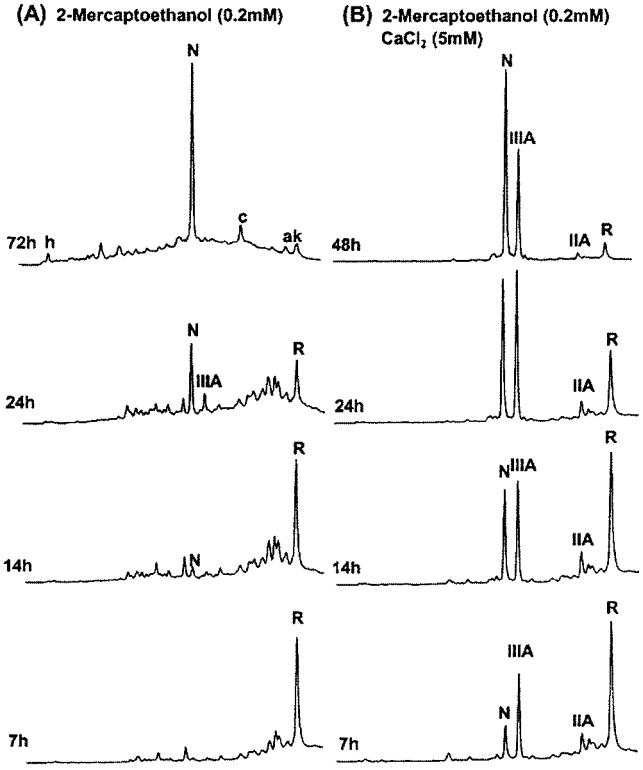

FIGURE 1: HPLC chromatograms of acid-trapped intermediates of αLA folding performed in the presence of 2-mercaptoethanol. Folding was carried out at 22 °C in Tris-HCl buffer (pH 8.4) containing (A) 2-mercaptoethanol (0.2 mM) or (B) 2-mercaptoethanol (0.2 mM) and $CaCl_2$ (5 mM). The protein concentration was 0.5 mg/mL. Intermediates of folding were trapped by acidification (4% trifluoroacetic acid) and analyzed by HPLC using the conditions described in Experimental Procedures. N indicates the elution position of the native αLA. R indicates the elution position of the starting material, fully reduced αLA. a, c, h, and k indicate the elution positions of four major four-disulfide scrambled intermediates observed in the absence of calcium. IIIA and IIA indicate the elution positions of a predominant three-disulfide intermediate and a major two-disulfide intermediate, respectively. Their disulfide structures are shown in Figure 6.

Results of HPLC runs performed in the presence of 2-mercaptoethanol (0.2 mM) are presented in Figure 1A. Folding of αLA was shown to undergo a sequential conversion through one-, two-, three-, and four-disulfide intermediates to reach the native structure. This was judged from the analysis of disulfide species (Figure 3, left column). However, the flow of intermediates is sluggish (Figure 4A), and the efficiency of renaturation is extremely low in both cases. Approximately 10% of the native αLA was recovered after folding had proceeded for 24 h (Figure 5). Inclusion of EDTA (1 mM) in the presence of 2-mercaptoethanol (0.2 mM) decreased the rate of oxidative folding of αLA by a factor of 4-fold. In this case, the 24 h trapped sample is indistinguishable from the 7 h sample shown in Figure 1A. The folding intermediates are heterogeneous (Figure 1A). Among them, one of the three-disulfide isomers (αLA-**IIIA**) and four of the four-disulfide isomers (X-αLA-a, X-αLA-c, X-αLA-h, and X-αLA-k) were recognized and characterized. Their disulfide structures are given in Figure 6.

Folding experiments with αLA were also carried out in the presence of 2-mercaptoethanol (0.2 mM) and $CaCl_2$ (5 mM). Acid-trapped folding intermediates were similarly analyzed by HPLC (Figure 1B). Inclusion of $CaCl_2$ alters the folding pathway of αLA dramatically, though this is not unexpected since αLA is a calcium binding protein. It reduces the complexity of folding intermediates and promotes a predominant buildup of one two-disulfide intermediate (αLA-**IIA**) and explicitly a three-disulfide intermediate (αLA-**IIIA**) (Figure 1B). Both αLA-**IIA** and αLA-**IIIA** contain two native disulfide bonds located within the β-sheet calcium binding domain (Figure 6). In addition, αLA-**IIIA** also comprises a third native disulfide (Cys^{28}−Cys^{111}). Conversion of αLA-**IIIA** to the native αLA requires only the formation of the fourth native disulfide bond (Cys^{6}−Cys^{120}), but this reaction is very slow even in the presence of calcium. It is apparent from the folding kinetics and the predominance of αLA-**IIIA** that, under such conditions, conversion of αLA-**IIIA** to the native structure represents a major rate-limiting step of αLA folding. The folding mechanism of αLA displayed in Figure 1B bears a striking resemblance to that of BPTI and EGF folding (*7*, *8*, *16*).

Oxidative Folding of α-LA in the Presence of Redox Buffer. The influence of redox agents (GSH and GSSG) on the folding of αLA was evaluated systematically. Their effects were first analyzed in the absence of calcium. Using GSH alone (0.2−1 mM), the folding mechanism of αLA is indistinguishable from that in the presence of 2-mercaptoethanol (Figure 1A). This was concluded from HPLC analysis of acid-trapped folding intermediates. On the other hand, GSSG (0.3−10 mM) plays a unique role in promoting the folding of αLA. It accelerates the formation of disulfide bonds and the flow of folding intermediates. For example, in the presence of 1 and 5 mM GSSG, the rate of disulfide bond formation increases by an astounding 300- and 1500-fold, respectively (compare panels A−C of Figure 4). For instance, the composition of disulfide species of the 2 min trapped sample determined in the presence of 5 mM GSSG (Figure 3, bottom right panel) is nearly identical to that of the 48 h trapped sampled determined in the absence of GSSG (Figure 3, top left panel). The 1500-fold acceleration is derived from the ratio of 2880 min (48 h) to 2 min. Under such conditions, folding intermediates rapidly accumulate as fully oxidized four-disulfide species. This is best demonstrated by the patterns of folding intermediates presented in Figure 2C and the composition of disulfide species illustrated in Figure 3 (right column). When reduced αLA was allowed to refold in the presence of 5 mM GSSG, more than 90% of the protein was found as four-disulfide species after folding had proceeded for 10 min. Among them, less than 10% was recovered as the native structure, and the remaining protein was shown to be non-native scrambled isomers. However, non-native scrambled isomers become trapped in part due to the absence of free thiols for promotion of their conversion to the native structure. The efficiency of αLA folding thus can be further improved by including both GSSG and GSH in the folding buffer (Figure 2A). In the presence of GSSG and GSH (1 mM each), ∼65% of the protein was recovered as the native αLA after folding had proceeded for 24 h. The composition of folding intermediates is exceedingly heterogeneous in both cases (Figure 2A,C). Among them are five distinguishable four-disulfide species, X-αLA-a, X-αLA-b, X-αLA-c, X-αLA-h, and X-αLA-k. Their disulfide structures are shown in Figure 6. It is important to mention that a negligible amount of folding intermediates exists with

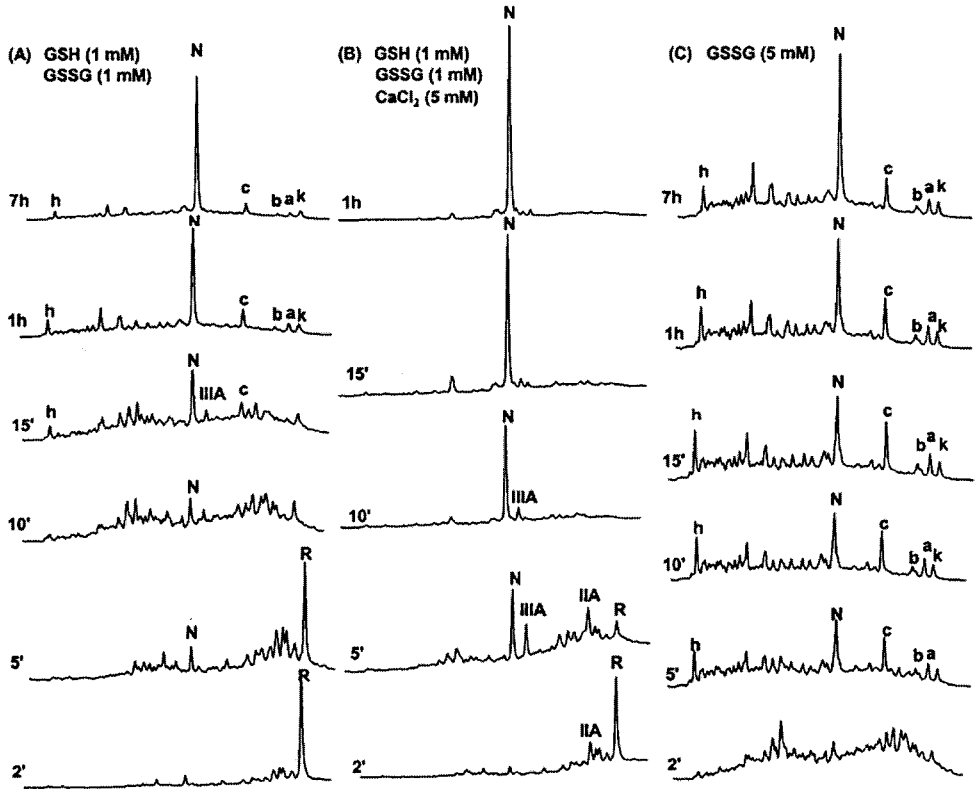

FIGURE 2: HPLC chromatograms of acid-trapped intermediates of αLA folding performed in the presence of redox agents. Folding was carried out at 22 °C in Tris-HCl buffer (pH 8.4) containing (A) reduced and oxidized glutathione (1 mM GSH and 1 mM GSSG), (B) reduced and oxidized glutathione (1 mM GSH and 1 mM GSSG) and $CaCl_2$ (5 mM), and (C) oxidized glutathione (5 mM GSSG). The protein concentration was 0.5 mg/mL. Intermediates of folding were trapped at different time points by acidification (4% trifluoroacetic acid) and analyzed by HPLC using the conditions described in Experimental Procedures. N indicates the elution position of native αLA. R indicates the elution position of the starting material, fully reduced αLA. IIIA and IIA indicate the elution positions of a major three-disulfide intermediate and a major two-disulfide intermediate, respectively. a−c, h, and k indicate the elution positions of five major four-disulfide scrambled intermediates. Their disulfide structures are shown in Figure 6.

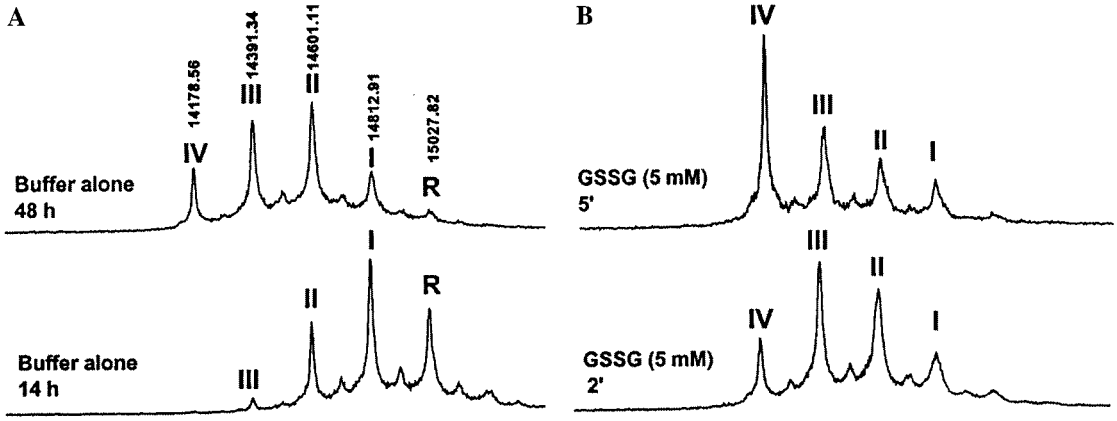

FIGURE 3: Molecular mass of the vinylpyridine-trapped folding intermediates of αLA. Folding was carried out at 22 °C in the following buffers: (A) Tris-HCl buffer (pH 8.4) alone and (B) Tris-HCl buffer (pH 8.4) containing GSSG (5 mM). The folding intermediates of α-LA contain various concentrations of zero-disulfide (R), one-disulfide (I), two-disulfide (II), three-disulfide (III), and four-disulfide (IV) species. One of the four-disulfide species is native α-LA (N). As a result of the reaction with vinylpyridine, these five populations of disulfide species can be distinguished by their relative molecular mass and identified by MALDI mass spectrometry. Each additional pair of vinylpyridine modifications increases the molecular mass by 212 Da.

glutathione-mixed disulfide, as shown by mass analysis (Figure 3).

Finally, oxidative folding of αLA was carried out in buffer containing both redox agents (1 mM GSSG and 1 mM GSH) and $CaCl_2$ (5 mM). The presence of the indicated concentration of calcium improves the rate of renaturation of native αLA by 2-fold (Figure 5). The folding pathway is characterized by both heterogeneous intermediates and the predominance of two nativelike species (αLA-**IIA** and αLA-**IIIA**) (Figure 2B).

Reductive Unfolding of Native αLA. Reductive unfolding of native αLA was analyzed both in the absence and in the presence of $CaCl_2$ (5 mM). The protein was reduced either with varying concentrations of dithiothreitol (from 0.01 to

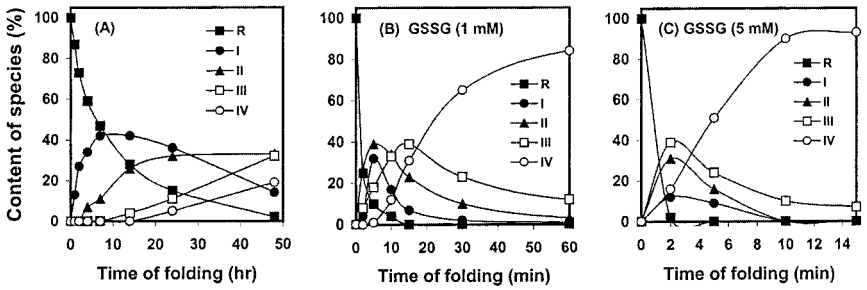

FIGURE 4: Quantitative analysis of various disulfide species along the pathway of oxidative folding of αLA. Folding was carried out at 22 °C in the following buffers: (A) Tris-HCl buffer (pH 8.4) alone, (B) Tris-HCl buffer (pH 8.4) containing GSSG (1 mM), and (C) Tris-HCl buffer (pH 8.4) containing GSSG (5 mM). R and I−IV indicate zero-disulfide (starting material), one-disulfide, two-disulfide, three-disulfide, and four-disulfide species, respectively. Quantitative analysis of various disulfide species of α-LA was based on the peak response of MALDI mass spectra (Figure 3). In the presence of GSSG (5 mM), the flow rate of folding intermediates increases by 1500-fold.

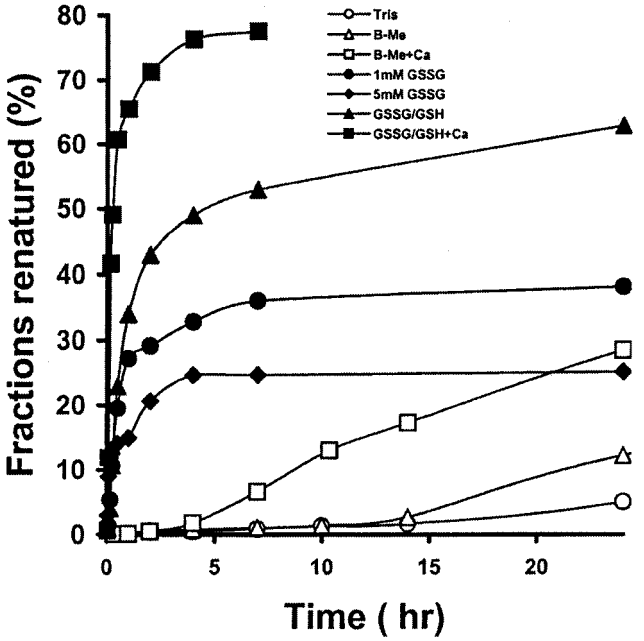

FIGURE 5: Renaturation curves of oxidative folding of αLA under different folding conditions. Folding was carried out at 22 °C in Tris-HCl buffer (pH 8.4) alone (○) or containing (△) 2-mercaptoethanol (0.2 mM), (□) 2-mercaptoethanol (0.2 mM) and $CaCl_2$ (5 mM), (◆) GSSG (5 mM), (●) GSSG (1 mM), (▲) GSSG and GSH (both at 1 mM), or (■) GSSG and GSH (both at 1 mM) with $CaCl_2$ (5 mM). The renaturation curves were determined by the rate of recovery of native αLA.

30 mM) for a fixed period of time (10 min) or with a fixed concentration of dithiothreitol in a time course manner. The results are shown in Figure 7. One of the native disulfide bonds of αLA (Cys^6−Cys^{120}) is similarly sensitive to reduction regardless of whether the protein is calcium-bound. Selective reduction of the Cys^6−Cys^{120} disulfide bond and conversion of the native αLA to αLA-**IIIA** occur rapidly at concentrations of dithiothreitol ranging between 30 and 300 μM. However, the pathway of reduction of the remaining three native disulfide bonds of αLA-**IIIA** is affected by calcium. In the absence of $CaCl_2$, reduction of αLA-**IIIA** proceeds in a nearly all-or-none manner, with direct conversion to the fully reduced αLA. In the presence of $CaCl_2$, conversion of αLA-**IIIA** to the reduced αLA involves a major two-disulfide intermediate (αLA-**IIA**) with preferential reduction of the Cys^{28}−Cys^{111} disulfide bond. Our data of reductive unfolding of αLA are fully consistent with those reported previously by Ewbank and Creighton (*25*, *26*).

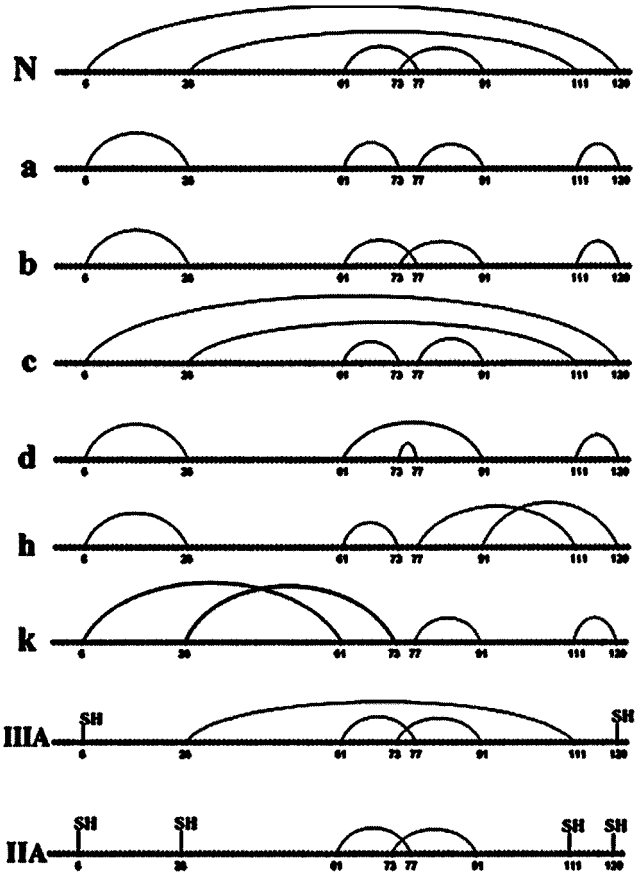

FIGURE 6: Disulfide structures of folding intermediates identified along the pathway of oxidative folding of αLA. The four disulfide bonds of native α-LA (indicated by N) are the Cys^6−Cys^{120}, Cys^{28}−Cys^{111}, Cys^{61}−Cys^{77}, and Cys^{73}−Cys^{91} bonds. The structures of isomers a−d, h, and k (all four-disulfide scrambled species) were determined previously (*27*, *28*). **IIIA** is a predominant three-disulfide intermediate. **IIA** is a major two-disulfide intermediate. The disulfide structures of **IIA** and **IIIA** were derived from Edman sequencing and MALDI mass analysis of disulfide-containing peptides of thermolysin-digested samples.

*Structural Analysis of a Predominant Two-Disulfide Intermediate (αLA-**IIA**) and a Three-Disulfide Kinetic Trap (αLA-**IIIA**).* Two predominant folding intermediates, designated as αLA-**IIA** and αLA-**IIIA**, were observed when folding of αLA was carried out in the presence of calcium. These two intermediates were treated with vinylpyridine, were further purified by HPLC, and were both shown to be a single species. Analysis by MALDI mass spectrometry revealed molecular masses of 14 603.5 Da for αLA-**IIA**

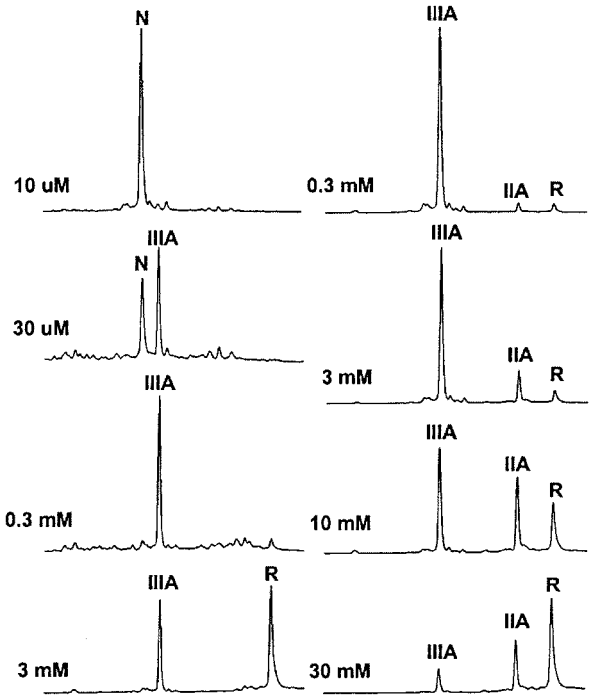

FIGURE 7: Reductive unfolding of αLA. The reactions were carried out both in the absence (left) and in the presence (right) of calcium (5 mM). α-LA (0.5 mg/mL) was reduced with the indicated concentrations of dithiothreitol at 22 °C for 10 min; the reaction was then quenched by acidification (4% trifluoroacetic acid), and the mixture was analyzed by HPLC using the conditions described in Experimental Procedures.

and 14 391.3 Da for αLA-**IIIA**. These molecular masses are 424 and 212 Da higher, respectively, than that of the native αLA (14 179 Da), which can be ascribed to 4 and 2 mol of conjugated vinylpyridine, respectively (MW = 105). The results thus confirm that αLA-**IIA** contains four cysteines and αLA-**IIIA** contains two free cysteines. The vinylpyridine-treated αLA-**IIA** and αLA-**IIIA** were digested with thermolysin. Thermolytic peptides were isolated by HPLC and analyzed by Edman sequencing and MALDI mass spectrometry to identify the structures of the disulfide-containing peptides. Data that were obtained (not shown) verify that αLA-**IIIA** forms there native disulfide bonds (Cys^{28}−Cys^{111}, Cys^{61}−Cys^{77}, and Cys^{73}−Cys^{91}), whereas αLA-**IIA** contains two native disulfide bonds (Cys^{61}−Cys^{77} and Cys^{73}−Cys^{91}). Their structures are illustrated in Figure 6.

Structures of Well-Populated Four-Disulfide (Scrambled) Intermediates of αLA. Five major four-disulfide intermediates were detected as folding of αLA was performed in the absence of calcium. They were designated as X-αLA-a, X-αLA-b, X-αLA-c, X-αLA-h, and X-αLA-k. The disulfide structures of these scrambled species (shown in Figure 6) were further confirmed by comparing the peptide maps of the thermolysin-digested samples with those reported previously by our laboratory (*27*, *28*) (data not shown).

DISCUSSION

αLA is among the most popular and extensively investigated models for understanding the mechanism of protein folding. Aside form the technique of oxidative folding (*24*, *25*), the folding pathway of αLA has also been elucidated by the method of disulfide scrambling (*27*, *28*) and, most commonly, by the conventional approach without disruption of its four native disulfide bonds (*30*−*40*). Here, we assess the pathway of oxidative folding of αLA and compare our results with the documented data obtained using both oxidative folding (disulfide formation) (*24*−*26*) and disulfide scrambling (*28*). The pathway of oxidative folding of αLA is also compared to that of several three-disulfide proteins (*7*−*19*).

Pathway of Oxidative Folding of αLA. Oxidative folding of αLA can be achieved using a wide range of conditions. However, the efficiency of folding (rate of renaturation) is dependent upon and enhanced by the cumulative effect of three distinct factors: (a) the kinetics of formation of disulfide bonds, which is accelerated by the presence of GSSG, (b) the kinetics of disulfide shuffling (rearrangement), which requires free thiol, and (c) the rate of formation (folding) of the β-sheet domain, which is facilitated by the presence of calcium. At the same time, the folding pathway of αLA and the complexity of folding intermediates are also determined by the same three factors as competing parameters. In summary, the folding pathway of αLA can be basically distinguished by the absence or presence of calcium.

In the absence of calcium, oxidative folding of αLA proceeds through heterogeneous one-, two-, three-, and four-disulfide intermediates to reach the native structure. This is the case regardless of whether redox agents are present. Without GSSG, the formation of disulfide bonds and the flow of folding intermediates are extremely sluggish (Figure 1A). The presence of GSSG accelerates the flow of folding intermediates. For instance, the rate of disulfide formation of reduced αLA increases by exactly 5-fold as the concentration of GSSG increases from 1 to 5 mM. Under such conditions, folding intermediates of αLA accumulate rapidly as four-diuslfide scrambled species (Figure 2C), and their conversion to the native structure represents a major rate-limiting step of folding. Inclusion of additional GSH in the folding buffer facilitates the disulfide shuffling (*28*, *41*) and the conversion of scrambled αLA to the native structure (Figure 2A).

Among the heterogeneous intermediates, one three-disulfide species (αLA-**IIIA**) and five four-disulfide scrambled species (X-αLA-a, X-αLA-b, X-αLA-c, X-αLA-h, and X-αLA-k) were identified and structurally characterized (Figure 6). αLA-**IIIA** is noticeable only with slow folding in the absence of GSSG (Figure 1A). It contains three native disulfide bonds, including the two (Cys^{61}−Cys^{77} and Cys^{73}−Cys^{91}) that stabilize the β-sheet domain. On the other hand, four-disulfide intermediates predominate with fast folding in the presence of GSSG (Figure 2C). One of the major four-disulfide intermediates (X-αLA-c) includes two native disulfide bonds (Cys^{6}−Cys^{120} and Cys^{28}−Cys^{111}) located within the α-helical domain. X-αLA-c has been shown to bear essentially all of the structural characteristics of the widely observed molten globule of αLA (*32*, *35*, *42*). The pattern and structure of four-disulfide folding intermediates are indistinguishable form those observed along the folding pathway of αLA using the technique of disulfide scrambling (*28*). It is important to mention here that calcium-depleted αLA is unable to fold quantitatively to form the native structure. A fraction of the non-native αLA always exists in a state of equilibrium with the native protein under physi-

ological pH and nondenaturing conditions. These non-native αLAs are distributed among more than 40 fractions of four-disulfide scrambled isomers and may constitute more than 25−30% of the total protein (see the 7 h sample of Figure 2A).

In the presence of calcium, the folding pathway of αLA differs significantly. It favors thermodynamically the formation of the β-sheet domain, increases the yield of the two nativelike intermediates (αLA-**IIA** and αLA-**IIIA**), and diminishes the complexity of the folding intermediates altogether. With slow folding in the absence of GSSG, αLA-**IIA** and αLA-**IIIA** represent the predominant and practically the only folding intermediates (Figure 1B). It is apparent that under these conditions, formation of the fourth disulfide bond (Cys^6-Cys^{120}) of αLA-**IIIA** accounts for the rate-limiting step of the oxidative folding of αLA. With fast folding in the presence of redox agents, heterogeneous species of folding intermediates resurface, but αLA-**IIA** and αLA-**IIIA** remain as two major species of folding intermediates (Figure 2B).

Comparison of Our Results with the Existing Data of Oxidative Folding of α*-LA.* The pathway of oxidative folding of αLA was also analyzed extensively by other laboratories. Rao and Brew (*24*) showed that folding of αLA was Ca^{2+}-dependent. Whereas in the absence of Ca^{2+} only ∼2% of the native protein was generated, in the presence of Ca^{2+}, near-quantitative renaturation of αLA was found. The folding mechanism and intermediates of αLA were further analyzed by Ewbank and Creighton (*25*, *26*) using gel electrophoresis. A similar effect of calcium was observed, and the folding intermediates were found to be highly heterogeneous. Specifically, native αLA was generated only when Ca^{2+} bound to and stabilized a three-disulfide intermediate lacking the Cys^6-Cys^{120} disulfide bond (i.e., αLA-**IIIA**), which then rapidly forms the fourth disulfide bond. The rate-limiting steps were shown to be disulfide rearrangement of the two- and three-disulfide species (*25*). Our data, including the extent of heterogeneity of folding intermediates and the presence of αLA-**IIIA** as a major folding intermediate, are largely consistent with these previous findings. Nonetheless, there are a few important discrepancies upon which we need to elaborate. Some of these inconsistencies are attributed to the different conditions employed for αLA folding.

First, we showed that native αLA can be generated not only in the presence but also in the absence of calcium. This is true either with or without the assistance of redox agents. For instance, ∼55% of the native αLA was recovered as folding was conducted for 7 h in the presence of GSH and GSSG (both at 1 mM) without calcium (Figure 3A). These data appear to contradict the earlier findings (*24*, *25*). However, our folding conditions are devoid of EDTA as an ion chelator. When the same experiment was performed in the presence of EDTA (1 mM), the yield of native αLA decreased from 55% to <8% (data not shown). This raises the possibility that a trace amount of Ca^{2+} contaminant may account for our high recovery of native αLA. Nonetheless, any Ca^{2+} contaminant should be removed by gel filtration following reduction and denaturation of αLA and before folding was initiated. The Tris base used for folding buffer preparation also has a purity of greater than 99.9%, with less than 0.0001% potential contamination of any metal ion. Such a trace amount of contaminant, if it exists at all, will have a concentration in the nanomolar range, which still could not account for the 55−65% yield of native αLA if equal molar binding of Ca^{2+} and αLA was required, since the concentration of αLA was ∼35 μM.

Second, we observed that scrambled (four-disulfide) species of αLA represent prevalent folding intermediates when folding is carried out in calcium-depleted buffer in the presence of redox agents (Figure 2A,C). Similar four-disulfide intermediates have not been described or identified in the previous studies (*24*, *25*). The origin of this discrepancy is not apparent to us. If anything, Ewbank and Creighton (*25*, *26*) used reduced and oxidized dithiothreitol instead of reduced and oxidized glutathione as a redox agent.

Third, we demonstrated that the rate-limiting step of αLA folding is dependent upon folding conditions. Two remarkable examples are shown here with αLA folding. In one case (Figure 1B), conversion of a nativelike three-disulfide intermediate (αLA-**IIIA**) to the native structure accounts for the rate-limiting step. In another instance (Figure 2C), conversion of four-disulfide intermediates to native αLA represents the rate-limiting step. Indeed, when factors that stabilize nativelike intermediates (e.g., BPTI) (*7*−*9*) and kinetic traps (e.g., EGF) (*16*) are excluded, the kinetics of the disulfide folding pathway is basically determined by the relative rate constant of disulfide formation (K_{DF}) and disulfide shuffling (rearrangement) (K_{DS}), which in turn can be regulated and controlled by adjusting the composition of the redox buffer (*41*). Specifically, the rate of disulfide formation is accelerated by the presence of GSSG or cystine, whereas the efficiency of disulfide rearrangement is promoted by GSH, cysteine, 2-mercaptoethanol, or protein disulfide isomerase. For αLA, these effects are conspicuous as oxidative folding was carried out under calcium-depleted conditions. When the rate of disulfide formation far exceeds that of disulfide reshuffling (e.g., in the presence of a high concentration of GSSG), fully oxidized scrambled isomers will accumulate and their conversion to the native structure, which requires disulfide shuffling, would account for the rate-limiting step of oxidative folding. A systematic study of such effects on the folding kinetics of hirudin has been reported (*41*). Similar phenomena have also been observed with several three-disulfide proteins (*11*, *13*), as well as αLA folding in the absence of calcium demonstrated here.

Comparison of the Folding Pathway(s) of α*LA Elucidated by the Techniques of Disulfide Formation (Oxidative Folding) and Disulfide Scrambling.* The folding pathway of αLA was recently analyzed using the technique of disulfide scrambling (*27*, *28*), which permits reversible conversion between the native and scrambled isoforms. One of the scrambled isomers of αLA (X-αLA-a) that possesses the highest free energy was used as the starting material of folding (*28*). X-αLA-a was recovered under strong denaturing conditions (6 M GdmCl) (*27*) and was allowed to refold through disulfide scrambling to form native αLA. Folding intermediates were trapped kinetically by acid quenching and analyzed quantitatively by reversed-phase HPLC. The results revealed two major on-pathway productive intermediates, two major off-pathway kinetic traps, and at least 30 additional minor transient intermediates. Of the two major on-pathway intermediates, one takes on a nativelike α-helical domain (X-αLA-c) and the other comprises a structured β-sheet, calcium binding domain (X-αLA-b). The two major kinetic

traps (X-αLA-h and X-αLA-k) are stabilized by locally formed non-nativelike structures.

The folding pathway of αLA revealed by disulfide scrambling (*28*) is strikingly similar to that elucidated here by disulfide formation (oxidative folding). Without calcium, the pathway of disulfide scrambling (see panels A and B of Figure 2 of ref *28*) is indistinguishable from that of oxidative folding performed in the buffer containing a high concentration of GSSG (Figure 2C). The folding intermediates of both pathways exhibit comparable heterogeneity and comprise the same prevalent species of four-disulfide isoforms (X-αLA-b, X-αLA-c, X-αLA-h, and X-αLA-k). In the presence of calcium, the pathway of disulfide scrambling consists of one predominant four-disulfide intermediate (X-αLA-b) (*28*), whereas the pathway of disulfide formation includes two major isomers (αLA-**IIA** and αLA-**IIIA**). It is imperative to mention that X-αLA-b, αLA-**IIA**, and αLA-**IIIA** all share the two native disulfide bonds (Cys^{61}−Cys^{77} and Cys^{73}−Cys^{91}) positioned within the *β*-sheet calcium binding domain (Figure 6).

Comparison of the Disulfide Folding Pathways of αLA with That of Three-Disulfide Proteins. The pathways of oxidative folding of several three-disulfide proteins have been characterized, including the most extensively documented case of bovine pancreatic trypsin inhibitor (BPTI) (*7*−*9*) as well as hirudin (*10*), tick anticoagulant peptide (TAP) (*11*, *12*), potato carboxypeptidase inhibitor (PCI) (*13*, *14*), epidermal growth factor (EGF) (*15*, *16*), insulin-like growth factor (*17*, *18*), etc. The major characteristics of their disulfide folding pathways, including the extent of heterogeneity of folding intermediates, their disulfide structures, and the accumulation of fully oxidized scrambled isomers, vary considerably among different proteins. The folding intermediates of BPTI comprise only one- and two-disulfide species. Of the 60 species of possible one- and two-disulfide isomers, only five or six species were shown to predominate along the folding pathway of BPTI, and most of them contain exclusively native disulfide bonds (*8*). The folding pathways of hirudin and PCI are nearly indistinguishable. However, it differs from that of BPTI by (a) a much higher degree of heterogeneity of one- and two-disulfide intermediates and (b) the presence of three-disulfide scrambled isomers as folding intermediates. The folding mechanism of EGF is further different from those of both BPTI and hirudin. It bears resemblance as well as dissimilarity to that of BPTI and hirudin. Folding intermediates of EGF consist of limited numbers of one- and two-disulfide species like those observed in the case of BPTI, but they also comprise scrambled three-disulfide isomers found in the cases of hirudin and other proteins. All these data, taken together, clearly magnify the extent of diversity of disulfide folding pathways. A systematic study of reductive unfolding of these proteins subsequently reveals that these characteristics are closely correlated to the fashion of how native disulfide bonds are stabilized (*19*). Those with their native disulfide bonds reduced by dithiothreitol in a collective (all-or-none) manner generally exhibit both a high degree of heterogeneity of folding intermediates and the accumulation of scrambled (fully oxidized) isomers along the folding pathway. Hirudin and potato carboxypeptidase inhibitor are but two examples (*10*, *13*, *14*). A sequential reduction of the native disulfide bonds is associated with the presence of predominant intermediates possessing nativelike structures. This mechanism is exemplified by the cases of BPTI and EGF (*7*−*9*, *16*).

The data of three-disulfide proteins and the behaviors of reductive unfolding of native αLA (Figure 7) provide the groundwork for interpreting the diversity of folding pathway(s) of αLA observed here. In the absence of calcium, reduction of αLA-**IIIA** proceeds through a nearly all-or-none mechanism; therefore, the folding pathway of αLA resembles that of hirudin and potato carboxypeptidase inhibitor, which is depicted by heterogeneous folding intermediates and the accumulation of four-disulfide scrambled isomers along the folding pathway. Conversely, the binding of calcium serves to generate a stable *β*-sheet domain. The native disulfide bonds of calcium-bound αLA were shown to be reduced in a sequential manner, with selective and sequential reduction of the two native disulfide bonds (Cys^{6}−Cys^{120} and Cys^{28}−Cys^{111}) located within the α-helical domain (*25*) (Figure 7). Consequently and consistently, the folding pathway of αLA in the presence of calcium consists of only two predominant intermediates adopting nativelike structures and bears a resemblance to that observed in the case of BPTI (*7*−*9*).

ACKNOWLEDGMENT

J.-Y.C. acknowledges the constructive discussion with Dr. Keith Brew.

REFERENCES

1. Creighton, T. E. (1978) *Prog. Biophys. Mol. Biol. 33*, 231−297.
2. Creighton, T. E., and Goldenberg, D. P. (1984) *J. Mol. Biol. 179*, 497−526.
3. Wedemeyer, W. J., Welker, E., Narayan, M., and Scheraga, H. A. (2000) *Biochemistry 39*, 4207−4216.
4. Frand, A. R., Cuozzo, J. W., and Kaiser, C. A. (2000) *Trends Cell Biol. 10*, 203−210.
5. Woycechowsky, K. J., and Raines, R. T. (2000) *Curr. Opin. Chem. Biol. 4*, 533−539.
6. Creighton, T. E. (1986) *Methods Enzymol. 131*, 83−106.
7. Creighton, T. E. (1990) *Biochem. J. 270*, 1−16.
8. Weissman, J. S., and Kim, P. S. (1991) *Science 253*, 1386−1393.
9. Goldengerg, D. P. (1992) *Trends Biochem. Sci. 17*, 247−261.
10. Chatrenet, B., and Chang, J.-Y. (1993) *J. Biol. Chem. 268*, 20988−20996.
11. Chang, J.-Y. (1996) *Biochemistry 35*, 11702−11709.
12. Chang, J.-Y., and Ballatore, A. (2000) *J. Protein Chem. 19*, 299−310.
13. Chang, J.-Y., Cannals, F., Schindler, P., Querol, E., and Aviles, F. X. (1994) *J. Biol. Chem. 269*, 22087−22094.
14. Venhudova, G., Canals, F., Querol, E., and Aviles, F. X. (2001) *J. Biol. Chem. 276*, 11683−11690.
15. Wu, J., Yang, Y., and Watson, J. T. (1998) *Protein Sci. 7*, 1017−1028.
16. Chang, J.-Y., Li, L., and Lai, P.-S. (2001) *J. Biol. Chem. 276*, 4845−4852.
17. Miller, J. A., Narhi, L. O., Hua, Q. X., Rosenfeld, R., Arakawa, T., Rohde, M., Prestrelski, S., Lauren, S., Stoney, K. S., Tsai, L., and Weiss, M. A. (1993) *Biochemistry 32*, 5203−5213.
18. Yang, Y., Wu, J., and Watson, J. T. (1999) *J. Biol. Chem. 274*, 37598−37604.
19. Chang, J.-Y., Li, L., and Bulychev, A. (2000) *J. Biol. Chem. 275*, 8287−8289.
20. Scheraga, H. A., Konishi, Y., and Ooi, T. (1984) *Adv. Biophys. 18*, 21−41.
21. Scheraga, H. A., Wedemeyer, W. J., and Welker, E. (2001) *Methods Enzymol. 341*, 189−221.
22. Welker, E., Narayan, M., Volles, M. J., and Scheraga, H. A. (1999) *FEBS Lett. 460*, 477−479.
23. Rothwarf, D. M., Li, Y.-J., and Scheraga, H. A. (1998) *Biochemistry 37*, 3760−3766.

24. Rao, K. R., and Brew, K. (1989) *Biochem. Biophys. Res. Commun. 163*, 1390−1396.
25. Ewbank, J. J., and Creighton, T. E. (1993) *Biochemistry 32*, 3677−3693.
26. Ewbank, J. J., and Creighton, T. E. (1993) *Biochemistry 32*, 3694−3707.
27. Chang, J.-Y., and Li, L. (2001) *J. Biol. Chem. 276*, 9705−9712.
28. Chang, J.-Y. (2002) *J. Biol. Chem. 277*, 120−126.
29. Permyakov, E. A., and Berliner, L. J. (2000) *FEBS Lett. 473*, 269−274.
30. Sugai, S., and Ikeguchi, M. (1994) *Adv. Biophys. 30*, 37−84.
31. Arai, M., and Kuwajima, K. (2000) *Adv. Protein Chem. 53*, 209−282.
32. Kuwajima, K. (1996) *FASEB J. 10*, 102−109.
33. Creighton, T. E., and Ewbank, J. J. (1994) *Biochemistry 33*, 1534−1538.
34. Peng, Z. Y., Wu, L. C., and Kim, P. S. (1995) *Biochemistry 34*, 3248−3252.
35. Wu, L. C., Peng, Z. Y., and Kim, P. S. (1995) *Nat. Struct. Biol. 2*, 281−286.
36. Wu, L. C., and Kim, P. S. (1998) *J. Mol. Biol. 280*, 175−182.
37. Pardon, E., Haezebrouck, P., De Baetselier, A., Hooke, S. D., Fancourt, K. T., Desmet, J., Dobson, C. M., Van Dael, H., and Joniau, M. (1995) *J. Biol. Chem. 270*, 10514−10524.
38. Redfield, C., Schulman, B. A., Milhollen, M. A., Kim, P. S., and Dobson, C. M. (1999) *Nat. Struct. Biol. 6*, 948−952.
39. Wu, L. C., Schulman, B. A., Peng, Z. Y., and Kim, P. S. (1996) *Biochemistry 35*, 859−863.
40. Schulman, B. A., Redfield, C., Peng, Z. Y., Dobson, C. M., and Kim, P. S. (1995) *J. Mol. Biol. 253*, 651−657.
41. Chang, J.-Y. (1994) *Biochem. J. 300*, 643−650.
42. Chang, J.-Y., Bulychev, A., and Li, L. (2001) *FEBS Lett. 487*, 298−300.

BI020169K