BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

A Native Disulfide Stabilizes Non-Native Helical Structures in Partially Folded States of Equine β-Lactoglobulin

Mio Yamamoto,[†] Kanako Nakagawa,[†] Kazuo Fujiwara,[†] Akio Shimizu,[‡] Mitsunori Ikeguchi,[§] and Masamichi Ikeguchi*,[†]

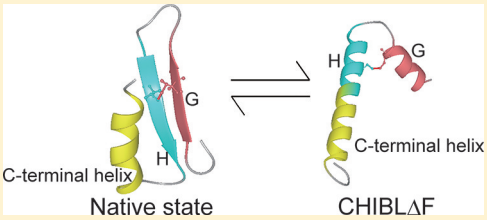
[†]Department of Bioinformatics, Soka University, 1-236 Tangi-cho, Hachioji, Tokyo 192-8577, Japan

[‡]Department of Environmental Engineering for Symbiosis, Soka University, 1-236 Tangi-cho, Hachioji, Tokyo 192-8577, Japan

[§]Department of Supramolecular Biology, Graduate School of Nanobioscience, Yokohama City University, 1-7-29 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan

Supporting Information

ABSTRACT: Equine β-lactoglobulin (ELG) assumes non-native helices during refolding and in partially folded states. Previously, circular dichroism (CD) combined with site-directed mutagenesis identified helical regions in the acid- and cold-denatured states of ELG. It is also known that a fragment of ELG, CHIBL (residues 88−142), has a structure similar to that of the cold-denatured state. For the study reported herein, the structure of a shorter fragment, CHIBLΔF (residues 97−142), was investigated by CD and nuclear magnetic resonance spectroscopy. The secondary chemical shifts clearly showed that non-native α-helices are present in two different regions, residues 98−107 and 114−135, and are connected by a native disulfide bond. The CD spectra of two peptides that correspond to the helical regions are characterized by weak helical signatures, and the sum of their CD spectra is nearly the same as the spectrum of disulfide-reduced CHIBLΔF. Therefore, the non-native helices are stabilized by the disulfide, and non-native helix formation may occur only during the refolding of the disulfide-intact protein. Supporting this conclusion is the observation that tear lipocalin, a homologue of ELG that lacks the disulfide, does not form non-native helices during folding.



Protein structures are marginally stable because various stabilizing and destabilizing factors offset each other. Many attempts have been made to understand how individual interactions, e.g., hydrogen bonds, electrostatic interactions, and hydrophobic interactions, contribute to stability.[1−4] However, it is difficult to evaluate the contribution of each interaction because the native protein structure is a cooperative system. Conversely, often the unfolding transition of folding intermediates and their equilibrium analogues, i.e., partially folded states, is less cooperative, which makes the energetic dissection of their structures possible. Characterization of the contributions made by various interactions to the stabilities of folding intermediates is important if we are to understand the interactions responsible for folding events. Detailed structural characterization for folding intermediates and equilibrium analogues has been done for classical molten globule states of apomyoglobin, cytochrome *c*, α-lactalbumin, and ribonuclease H.[5,6] A recently developed method utilizing the NMR relaxation dispersion has made it possible to determine the three-dimensional structure of "invisible" folding intermediates of small proteins such as a SH3 domain[7] and a FF domain[8] at atomic resolution. Using hydrogen-exchange-directed protein engineering, "hidden" intermediates that exist after the rate-limiting step were populated to determine the high-resolution structure.[9,10]

We previously characterized the acid-denatured state of equine β-lactoglobulin (ELG), which is an equilibrium analogue of folding intermediate of this protein.[11−16] ELG is a major whey protein of 162 residues and has two disulfide bonds (Cys66−Cys160 and Cys106−Cys119). Although the three-dimensional structure of ELG has not been obtained, it seems to be an eight-stranded (labeled A−H) up-and-down β-barrel fold with an α-helix succeeding to the last strand in the β-barrel (C-terminal helix) and the off-barrel β-strand I from the known three-dimensional structure of the homologous protein[17] and NMR analyses.[13] The acid-denatured (A) state of ELG is a molten globule. It has a compact globular structure with native-like and non-native secondary structures,[11,16] and its structure appears to be the same as, or nearly the same as, the ELG burst-phase folding intermediate.[12] ELG also has a partially folded state (C state) that manifests below 0 °C in 2 M urea,[14] which has an expanded chainlike conformation and a helical content greater than that of the A state. Interestingly, the A state is converted to a conformation similar to that of the C state when the temperature is <0 °C, suggesting that the A state is stabilized by hydrophobic interactions.

A mutant ELG (C66A/C160A), in which Cys66 and Cys160, which forms a native disulfide, were replaced with alanines, has also been characterized and is structurally similar to that of

Received: August 19, 2011
Revised: November 10, 2011
Published: November 10, 2011

wild-type ELG under native conditions.[18] C66A/C160A also assumes a conformation similar to the A state of ELG at acid pH in the presence of salt, while it assumes a conformation similar to the C state of ELG at acid pH in the absence of salt and at 25 °C.[18,19] In the absence of salt, with a decrease in temperature, the helical content of C66A/C160A, as monitored by CD spectroscopy, increases monotonically, whereas the helical content increases cooperatively as the temperature decreases in the presence of 0.1 M KCl and reaches the same final value as that found for the mutant in the absence of salt at low temperature.[15] These results suggest that electrostatic repulsions at low ionic strength or attenuation of hydrophobic interactions at low temperature expands the molecule and alters its secondary structure.

The secondary structure for each of the A and C states has been characterized by proline-scanning mutagenesis,[16] which suggested that, in the A state, the region of H-strand assumed a non-native helix in addition to the native-like C-terminal helix and that the F- and G-strands assumed a β-hairpin. In the C state, the region of F-, G-, and H-strands were shown to assume non-native helices in addition to the native-like C-terminal helix.

To identify the interactions responsible for secondary structure formation in the A and C states, the structures of several peptides have been investigated.[15] Although short peptides that corresponded to the H-strand and the C-terminal helix formed little helical structure, the fragment CHIBL (core of a helical intermediate of beta-lactoglobulin), which corresponds to the residues 88−142 that contains the F-, G-, and H-strands, and the C-terminal helix in the native state, formed a remarkable helical structure. The proline substitution of CHIBL showed CD spectral changes similar to the corresponding mutants of the full-length protein in the C state. The temperature dependence of the CHIBL secondary structures, as reported by CD spectroscopy, was linear, as was that of C66A/C160A in the C state. Therefore, CHIBL seemed to assume a structure similar to its corresponding region in the C state of C66A/C160A. Furthermore, the interactions responsible for helix formation in the C state seemed to be present in the CHIBL sequence, although the helices in CHIBL were not stable in isolation.

To understand how the helices in CHIBL are stabilized, it is necessary to know the more detailed structure of CHIBL. For the work reported herein, we constructed a shorter fragment of ELG (CHIBLΔF (residues 97−142), Figure 1), in which the F-strand region was absent, and characterized the structure of CHIBLΔF using CD and multidimensional heteronuclear nuclear magnetic resonance (NMR) spectroscopy. It was expected that CHIBLΔF assumed a structure similar to that of CHIBL because the L95P substitution little reduced the CD intensity of CHIBL.[15] CHIBLΔF was favorable for NMR spectroscopy because its expression level in *E. coli* is higher than that of CHIBL. Additionally, the conformations of shorter peptides were characterized by CD.

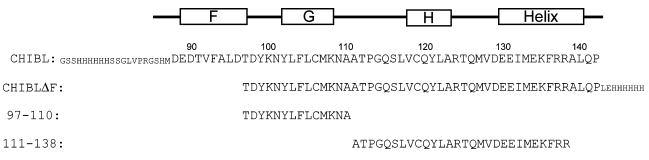


Figure 1. Amino acid sequences for CHIBL, CHIBLΔF, 97−110, and 111−138. Residue numbers correspond to those of full-length ELG. The native F, G, and H β-strands and the helix are boxed. The N-terminal hexahistidine sequence of CHIBL and the C-terminal hexahistidine sequence of CHIBLΔF are presented in a smaller font.

MATERIALS AND METHODS

Construction of a Gene Encoding CHIBLΔF and Expression Vectors. The gene encoding CHIBLΔF was constructed using PCR, suitable primers, and the C66A/C160A DNA construct as the template. Its sequence was confirmed by DNA sequencing using an ABI PRISM 3100-Avant sequencer. The gene was inserted into pET21a, and the expression vector was transformed into *Escherichia coli* BL21(DE3).

Expression, Refolding, and Purification of CHIBLΔF. To obtain nonisotopically labeled CHIBLΔF, *E. coli* harboring the expression plasmid was grown in Luria−Bertani medium. CHIBLΔF was found in the insoluble fraction of the cell lysate and was solubilized in 50 mM Tris-HCl, pH 8.0, containing 8 M urea, 0.5 mM hydroxyethyl disulfide, and 5 mM β-mercaptoethanol. Using its C-terminal hexahistidine tag, CHIBLΔF was purified by Ni^{2+}-chelating chromatography (GE Healthcare). After the protein had been applied to the resin, it was washed with 50 mM Tris-HCl, pH 8.0, containing 10 mM imidazole, 0.5 M NaCl, and 4 M urea, and then CHIBLΔF was eluted with a linear gradient of 100−600 mM imidazole in an otherwise identical solution. Fractions containing CHIBLΔF were dialyzed against water, in which CHIBLΔF precipitated. The precipitate was collected by centrifugation, washed with distilled water, and lyophilized. The purity of CHIBLΔF was confirmed by SDS-PAGE and analytical reverse-phase HPLC, and its molecular mass was confirmed by matrix-assisted laser desorption ionization time-of-flight mass spectrometry using an Autoflex II TOF/TOF mass spectrometer (Bruker Daltonics).

Uniformly ^{15}N-labeled and $^{15}N/^{13}C$-labeled CHIBLΔF were obtained by culturing *E. coli* that harbored the expression plasmid in M9 minimal medium and $^{15}NH_4Cl$ or $^{15}NH_4Cl/[^{13}C]$ glucose, respectively, as the sole nitrogen and carbon sources. The labeled proteins were purified as described above.

Reduction of CHIBLΔF was carried out at room temperature for 15 min in 50 mM Tris-HCl, pH 8.0, containing 20 mM dithiothreitol, 6.0 M guanidine HCl, and 2 mM ethylenediaminetetraacetic acid. After reduction, the reagents were removed by Sephadex G-25 chromatography (GE Healthcare) in 0.1 M H_3PO_4 at pH 1.6. Fractions containing CHIBLΔF were used for CD experiments. 5,5′-Dithiobis(2-nitrobenzoic acid) was used to quantify the number of thiol groups in a sample.[20]

Peptide Synthesis. Peptides were synthesized using a PS3 peptide synthesizer (Protein Technologies), fluorenylmethoxycarbonyl chemistry, and amide resin so that each C terminus was amidated. After cleaving the peptides from the resin, each N terminus was acetylated with *N*-acetylimidazole. Peptide 97-110 was purified by reverse-phase HPLC through an Inertsil ODS-3 column (20 × 150 mm, GL Science). Peptide 111-138 was purified by two rounds of HPLC through a Cosmosil C8 column (4.6 × 250 mm, Nacalai Tesque, Inc.). Peptide purities were confirmed by analytical reverse-phase HPLC, and their molecular masses were confirmed by matrix-assisted laser desorption ionization time-of-flight mass spectrometry.

CD Spectroscopy. CD spectra were measured using a Jasco J-720 spectropolarimeter or an Applied Photophysics Chirascan CD spectrometer. Unless otherwise indicated, the

peptide concentration was 0.2 mg/mL as measured by ultraviolet spectroscopy with extinction coefficients derived from the number of tyrosines in the peptides and a tyrosine extinction coefficient of 1280 M^{-1} cm^{-1}.[21] Cuvettes with path lengths between 1 and 10 mm were used depending on the peptide concentration.

Sedimentation Equilibrium Ultracentrifugation. Experiments were performed using a Beckman XL-A analytical ultracentrifuge at 12 °C, a rotor speed of 38 500 rpm, an An-50 Ti rotor, and a charcoal-filled Epon centerpiece. Data were collected using continuous radial scanning. Equilibrium was attained within 20 h.

NMR Resonance Assignment. The labeled CHIBLΔF samples (~1 mM, 5 mg/mL) were prepared in 90% H_2O/10% D_2O, 0.1 M phosphoric acid, pH 1.8. Sodium 3-trimethylsilyl-2,2,3,3-^{2}H propionate served as the internal ^{1}H reference (0.003 ppm), and ^{13}C and ^{15}N chemical shifts were indirectly referenced.[22] NMR spectra were recorded at 12 °C using an ECA-500 NMR spectrometer (JEOL) and the automation software GORIN provided by the manufacturer. For backbone resonance assignments, ^{1}H−^{15}N HSQC, HNCACB, CBCA(CO)NH, HNCO, HN(CA)CO, C(CO)NH, and ^{15}N-edited TOCSY spectra were acquired. To obtain NOEs, an ^{15}N-edited NOESY spectrum was acquired with a 100 ms mixing time. NMR data were processed using NMRPipe,[23] and resonances were assigned using KUJIRA.[24]

Pulse-Field Gradient NMR Spectroscopy. The apparent translational diffusion coefficient D for CHIBLΔF in 0.1 M deuterated phosphoric acid, pD 1.5, or 8 M deuterated urea, pD 1.5 (adjusted with DCl), was measured at 12 °C using pulse-field gradient NMR spectroscopy with the stimulated echo pulse sequence. Dioxane and 2,2-dimethyl-2-silapentane-5-sulfonic acid were the internal references. Signal-intensity attenuations for CHIBLΔF and dioxane were measured as a function of field gradient strength. Plots of signal intensities were fit to eq 1 using nonlinear least-squares regression analysis to obtain values for D.

$$I = \exp[-(\gamma\delta G)^2(\Delta - \delta/3)D] \quad (1)$$

with I, the signal intensity; γ, the ^{1}H gyromagnetic ratio; G and δ, the magnitude and duration of the magnetic field gradient pulses, respectively; and Δ, the time between field gradient pulses, with Δ = 100 ms, δ = 8 or 10 ms, and G = 10−330 mT/m. The dependence of CHIBLΔF concentration on D was measured for samples that contained between 0.5 and 2.3 mM CHIBLΔF. The diffusion coefficient at infinite dilution, $D°$, was estimated by linear extrapolation of the data. The hydrodynamic radius of CHIBLΔF was calculated using eq 2

$$R_{\mathrm{CHIBL\Delta F}} = (D°_{\mathrm{dioxane}}/D°_{\mathrm{CHIBL\Delta F}})R_{\mathrm{dioxane}} \quad (2)$$

with R_{dioxane}, the hydrodynamic radius of dioxane, 2.12 Å.[25]

RESULTS

CD Spectra of CHIBLΔF. The CD spectrum of CHIBLΔF in 0.1 M phosphoric acid, pH 1.7 (Figure 2A), was independent of the CHIBLΔF concentration between 0.02 and 0.4 mg/mL and was identical to that for CHIBLΔF in 0.1 M HCl/KCl, pH 1.7 (data not shown), which indicated that the conformation of CHIBLΔF did not depend on the concentration of chloride. The CD spectrum of CHIBLΔF was similar to that of CHIBL; both spectra indicated that the proteins had a substantial amount of α-helix. The mean residue ellipticity ($[\theta]_{\mathrm{MRE}}$) for

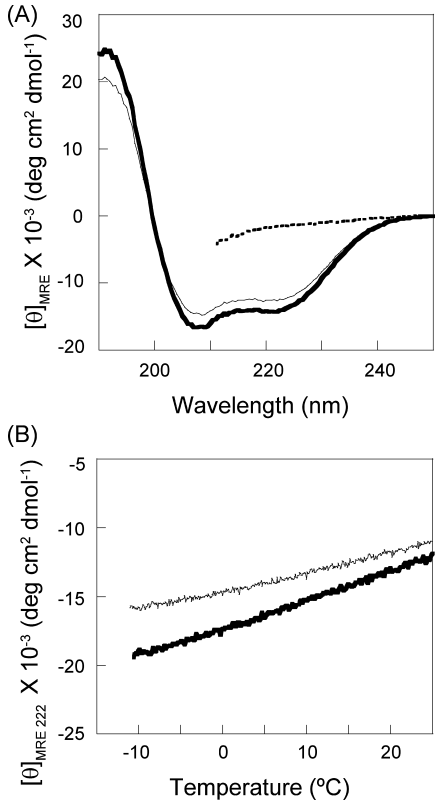


Figure 2. (A) CD spectra of CHIBL (thin, solid line) and CHIBLΔF (thick, solid line) in 0.1 M phosphoric acid, pH 1.7, at 12 °C. The broken line is the spectrum of CHIBLΔF in 8 M urea, pH 1.7. (B) Temperature dependence of $[\theta]_{\mathrm{MRE}}$ at 222 nm for CHIBL (thin, solid line) and CHIBLΔF (thick, solid line).

CHIBLΔF at 222 nm was slightly more negative than that of CHIBL. The dependence on temperature by $[\theta]_{\mathrm{MRE}}$ for the two proteins at 222 nm is shown in Figure 2B. The $[\theta]_{\mathrm{MRE}}$ at 222 nm for CHIBLΔF decreased gradually with decreasing temperature, as did that for CHIBL and C66A/C160A in the C state,[15] suggesting that the conformation of CHIBLΔF was in equilibrium between helical and disordered structures. Therefore, it was difficult to estimate the number of residues that assumed a helical conformation. To assess the conformational properties of each residue in CHIBLΔF, we therefore used heteronuclear multidimensional NMR spectroscopy.

NMR Spectroscopy of CHIBLΔF. The ^{1}H−^{15}N HSQC spectrum of CHIBLΔF at 12 °C is shown in Figure 3. HNCACB, CBCA(CO)NH, HNCO, HN(CA)CO, and C(CO)NH spectra were acquired for ^{1}HN, ^{15}N, ^{13}Cα, ^{13}Cβ, and ^{13}CO resonance assignments. An ^{15}N-edited TOCSY spectrum was used to obtain ^{1}Hα assignments. Backbone ^{1}H, ^{13}C, ^{15}N, and ^{13}Cβ assignments were obtained for all residues including the residues in the hexahistidine tag (Table S1 of the Supporting Information).

Chemical shifts are sensitive to protein structure.[26,27] Deviations of the observed chemical shifts from those for random coils, i.e., the so-called secondary chemical shifts, are valuable indicators of secondary structure propensities. When a residue is located in an α-helix, usually the positions of its ^{1}Hα and ^{13}Cβ resonances shift upfield relative to their random-coil values, whereas the ^{13}Cα and ^{13}CO resonances shift downfield. When a residue is located in a β-sheet or an extended conformation, the relative positions of the aforementioned

dx.doi.org/10.1021/bi2013239 | Biochemistry 2011, 50, 10590−10597

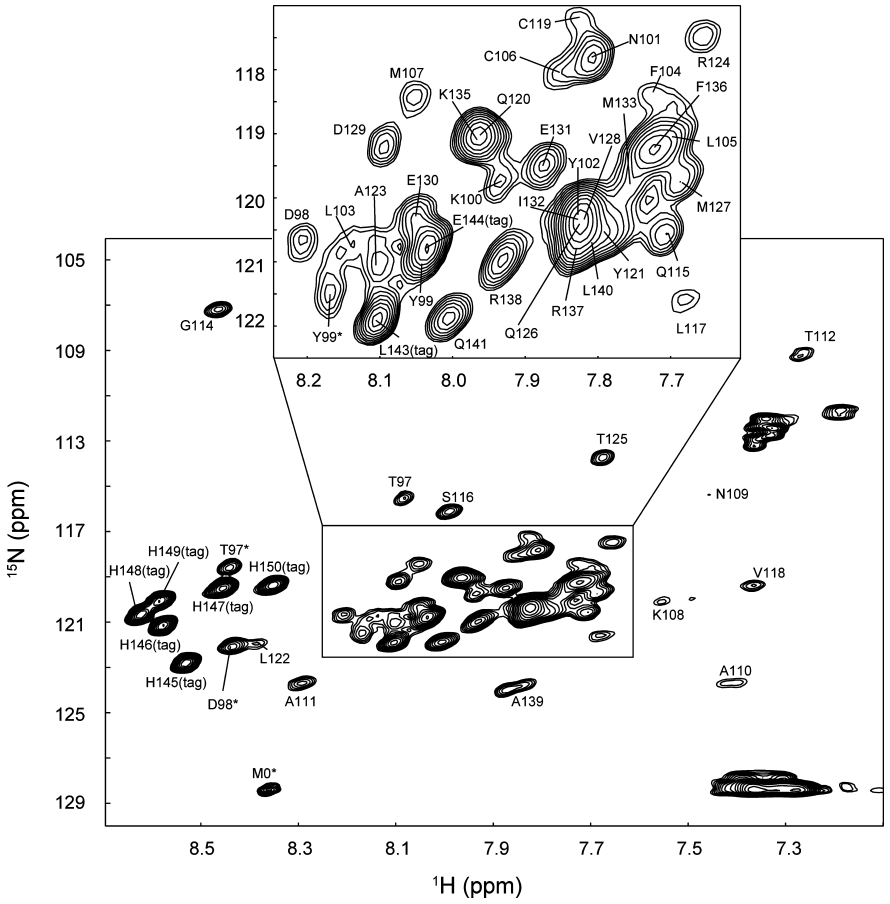


Figure 3. ^{1}H−^{15}N HSQC spectrum of CHIBLΔF in 0.1 M phosphoric acid (90% H_2O/10% D_2O), pH 1.8, at 12 °C. Assigned peaks are labeled. The sample contained two forms of CHIBLΔF: one with a formylated N-terminal methionine (labeled M0*) and the other with a free N-terminal methionine. Because of the heterogeneity, two cross-peaks, labeled with asterisks, are present for certain residues. Cross-peaks arising from the hexahistidine sequence are labeled as (tag).

resonances are usually reversed. We calculated secondary chemical shifts for the ^{13}CO signals, and for the $^{13}C\alpha$, $^{13}C\beta$, and $^{1}H\alpha$ signals, which are shown in Figure 4 and Figure S1, respectively, using values for random coil-type resonances obtained from Wishart et al.[28] Downfield secondary $^{13}C\alpha$ and ^{13}CO shifts were observed for 98-107 (which includes most of the G-strand) and for 114-135 (which includes the H-strand and part of the C-terminal helix), which indicated their $\phi-\psi$ were helical. The upfield shifts observed for the $^{1}H\alpha$ resonances of the same residues provided further evidence for the presence of helical conformations. Marsh and colleagues[29] developed a method to evaluate the fraction of α- or β-structures using a combination of the secondary shifts of different nuclei, denoted the secondary structure propensity (SSP) score. For a given residue, an SSP score equal to 1 or −1 reflects a fully formed α-helical or β-strand conformation, respectively. We calculated SSP scores for CHIBLΔF using the $^{1}H\alpha$, $^{13}C\alpha$, and $^{13}C\beta$ chemical shifts for each residue (Figure 4B), and the scores suggested that 98-107 and 114-135 assumed α-helical-type structures, although the populations of the helical conformations were ∼50%, which is consistent with the observation that the $[\theta]_{MRE}$ at 222 nm of CHIBLΔF depends on temperature (Figure 2B). The SSP score averaged over all residues in a polypeptide chain multiplied by $[\theta]_{MRE}$ at 222 nm of a fully helical polypeptide of infinite length (−42 500 deg cm^2 dmol^{-1} [30]) should give the observed $[\theta]_{MRE}$, if no β-sheet structure is present. When we performed this calculation with zeros instead of negative SSP scores, the average SSP score was 0.30, and the calculated value was −12 900 deg cm^2 dmol^{-1}, which is nearly that of the observed $[\theta]_{222}$, −14 300 deg cm^2 dmol^{-1}. Thus, the fraction of helix indicated by the averaged SSP score was consistent with the $[\theta]_{MRE}$ value at 222 nm.

NOEs are powerful constraints used for protein structure determination. NOEs provide long-range distance constraints on the global topology and short-range constraints on the local structures (secondary structures). In general, however, few NOEs in an unfolded or partially folded protein are observed because the polypeptide chain fluctuates, and conformations with defined secondary structures are not substantially populated. Conversely, sequential $d_{\alpha N}(i,i+1)$ NOEs are observed in the spectra of virtually all unstructured states.[27] Furthermore, sequential $d_{NN}(i,i+1)$ NOEs are detected for residues in helices even in unfolded or partially folded proteins.[31] Although $d_{NN}(i,i+1)$ NOEs are also found for β-sheet and turn residues, they are much weaker (for β-sheets) or not contiguous (for turns) and therefore can be used to discriminate between helical and nonhelical conformations.[31,32]

In the ^{15}N-edited NOESY spectrum of CHIBLΔF, we did not find long-range NOEs but observed only sequential $d_{NN}(i,i+1)$ and $d_{\alpha N}(i,i+1)$ NOEs. Sequential $d_{\alpha N}(i,i+1)$ NOEs were observed for the complete sequence, and sequential $d_{NN}(i,i+1)$ NOEs were observed for the residues of 100−108 and 114−135, except when the resonances of adjacent residues overlapped (data not shown). Therefore, these two regions

dx.doi.org/10.1021/bi2013239 | Biochemistry 2011, 50, 10590−10597

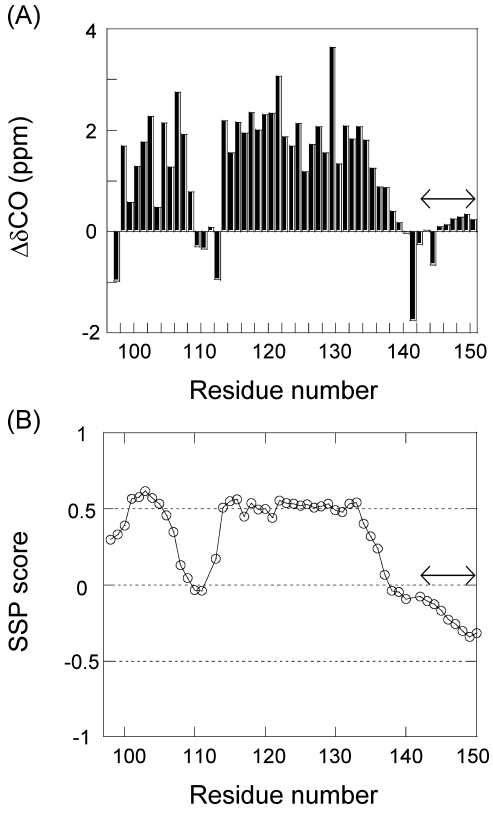


Figure 4. (A) Chemical shift deviations from random coil values for ^{13}CO CHIBLΔF resonances. (B) SSP scores. Double arrows indicate values for the hexahistidine sequence.

assumed helical structures, which is consistent with the positions of the chemical shifts and that the two helices do not pack together tightly.

Global Shape of CHIBLΔF. To additionally characterize the CHIBLΔF conformation, we measured its apparent translational diffusion coefficient D as a function of its concentration using pulse-field gradient NMR spectroscopy. As a control, D was measured for CHIBLΔF in 8 M urea. The CD spectrum of CHIBLΔF in 8 M urea indicated that CHIBLΔF was disordered (Figure 2A). The value for $D°$ at infinite dilution was estimated by linearly extrapolating the concentration dependences (Figure 5) and was $(9.93 \pm 0.17) \times 10^{-11}$ and $(6.01 \pm 0.22) \times 10^{-11}$ m^2 s^{-1}, respectively, for CHIBLΔF in the absence or presence of 8 M urea. The hydrodynamic radius of CHIBLΔF in the absence or presence of 8 M urea, calculated using eq 2, was 17.9 ± 0.6 and 21.8 ± 0.4 Å, respectively, the latter of which was similar to the value expected for a 54-residue unfolded protein (21.5 Å).[25] The hydrodynamic radius of aqueous CHIBLΔF was smaller than that found for CHIBLΔF in 8 M urea and also indicated that CHIBLΔF was monomeric at a concentration of ∼1 mM, which was the concentration used to obtain the CHIBLΔF NMR spectra. Up to a concentration of ∼2 mM, furthermore, CHIBLΔF was monomeric in 0.1 M phosphoric acid at 12 °C and pH 1.7, as assessed by analytical ultracentrifugation (data not shown).

Shorter Peptides Structures. Previously, we synthesized peptides that contained the H-strand (111−128 and 112−127) and the C-terminal helix (124−138 (F136Y)).[15] Although their CD spectra indicated that 124−138 was partially helical, 111−128 and 112−127 were not helical. We also concluded that the

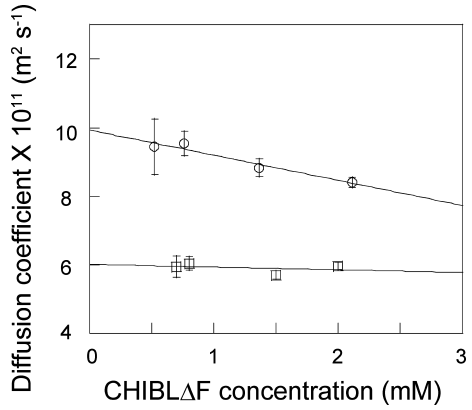


Figure 5. Apparent diffusion coefficients for CHIBLΔF as a function of concentration. Circles and squares indicate the values for CHIBLΔF in 0.1 M phosphoric acid/D_2O and 8 M deuterated urea/D_2O, respectively, at 12 °C. Uncorrected pD = 1.5.

non-native helix that had been identified by substituting a proline for A123 in C66A/C160A in the C state could not form in the absence of long-range interactions. However, the present NMR study indicated that the non-native helix in the H-strand (originally identified by the A123P substitution) and the native-like C-terminal helix (originally identified by the I132P and/or R137P substitutions) were present as one long helix, which suggested that the non-native α-helix in the H-strand may be stable in the longer peptide. To examine this possibility, we synthesized the peptide 111−138. Its CD spectrum was similar to that of 124−138 (F136Y) (Figure 6A), suggesting that the C-terminal α-helix propagated toward the N-terminus and all residues of the peptide 111-138 showed the helicity similar to that of 124−138 (F136Y) (30% and 38% at 25 and −2 °C, respectively[15]). Alternatively, the presence of the N-terminal region (111−123) might have caused a 2-fold increase in the helical propensity of 124−138 without assuming a helical conformation itself. In this case, the $[\theta]_{MRE}$ at 222 nm of 111−138 would be the same as that of 124−138. Although limitation in the available peptide sample did not allow the NMR analysis, the latter possibility is unlikely because the SSP scores of residues 113−132 in CHIBLΔF were approximately the same (Figure 4B).

We also synthesized the peptide 97−110, which according to its CD spectrum has little helical structure (Figure 6A). Figure 6B presents the CD spectra of 97−110, 111−138, and reduced and oxidized CHIBLΔF using the molar ellipticity ($[\theta]_M$) scale, which allows for a direct comparison of the sum of the fragments' spectra with those of CHIBLΔF. The CD spectrum of oxidized CHIBLΔF had a much deeper trough at 222 nm than did the sum of the fragments. However, the $[\theta]_M$ of reduced CHIBLΔF was less negative (Figure 6B), indicating that Cys106−Cys119 stabilized the helices in 98−107 and 114−135. Intriguingly, the CD spectrum of reduced CHIBLΔF was nearly the sum of the fragments' spectra. Therefore, the helices in reduced CHIBLΔF were mainly stabilized by local interactions, i.e., by the intrinsic helical propensity of the sequence. Furthermore, we measured the CD spectrum of the peptide mixture, in which 97−110 and 111−138 are included at 48 μM each. The spectrum of the mixture was identical to the sum of the fragments' spectra (data not shown), indicating that the two peptides do not interact with each other at least at the concentration used.

dx.doi.org/10.1021/bi2013239 | *Biochemistry* 2011, 50, 10590−10597

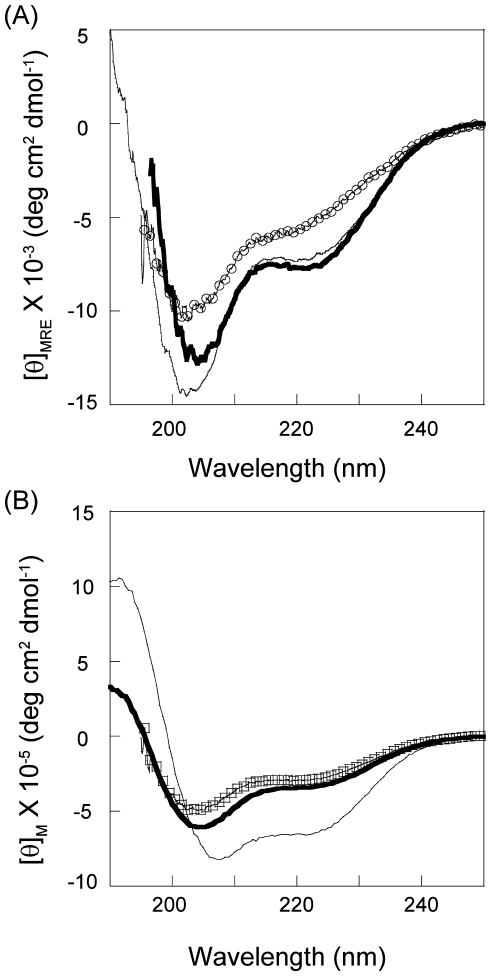


Figure 6. (A) CD spectra of 111−138 (thick, solid line), 124−138 (thin, solid line), and 97−110 (circles) in 0.1 M phosphoric acid, pH 1.7, at 25 °C. The vertical axis is presented by $[\theta]_{MRE}$. (B) CD spectra of reduced (thick, solid line) and oxidized (thin, solid line) CHIBLΔF. Squares show the sum of the 97−110 and 111−138 spectra. Conditions are as in (A). The vertical axis is presented by $[\theta]_M$.

DISCUSSION

Structure of CHIBLΔF. For the study reported herein, we characterized the structure of an ELG fragment, CHIBLΔF, by CD and NMR spectroscopy. The CD spectrum of CHIBLΔF indicates that the peptide is substantially α-helical, and the values of the secondary chemical shifts for its NMR resonances indicate that the helices are found in residues 98−107 and 114−135. However, the SSP values and the temperature dependence of its CD spectrum indicate that, at 12 °C, ∼50% of CHIBLΔF is helical on average. Do the helices in 98−107 and 114−135 form independently? Is the simultaneous formation of the two helices compatible with the presence of Cys106−Cys119? To address these questions, we built structures for CHIBLΔF using MODELER.[33] A total of 100 structures were constructed with the constraint that helical regions were only found for residues 98−107 and 114−135. The conformational soundness of the structures were assessed using PROCHECK.[34] Six models had torsion angles preferred by disulfide bonds (96.8 ± 14.8° and −85.8 ± 10.7°), which suggested that the two non-native helices and the disulfide could exist simultaneously (Figure 7). In the six models, the two helices had different orientations, indicating that the 14-residue loop formed by the disulfide does not restrict the relative positions of the helices, which is a conclusion consistent with the lack of long-range NOEs found for CHIBLΔF. The PROCHECK-validated structures have gyration radii between 14.3 and 16.3 Å. Taking into account the coexistence of disordered conformation, these values are consistent with the measured hydrodynamic radius of 17.9 Å.

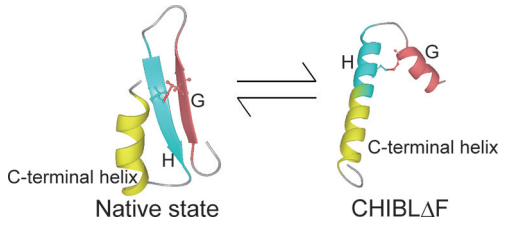


Figure 7. Schematic representation of one of structural models of CHIBLΔF and the corresponding region of ELG in the native conformation. G and H strands and C-terminal helix regions in the native conformation are colored magenta, cyan, and yellow, respectively.

Proline Substitutions. Previously, we probed the CHIBL structure at 25 and −2 °C by substituting prolines.[15] In the following discussion, we use averaged values for $[\theta]_{MRE}$ at 222 nm obtained at 25 and −2 °C because they should be identical to the values at 12 °C, which was the temperature used for NMR spectroscopy. This assumption is valid because the temperature dependence of $[\theta]_{MRE}$ at 222 nm is nearly linear for CHIBL (Figure 2B) and its proline mutants (data not shown). The L103P substitution in CHIBL increased $[\theta]_{MRE}$ at 222 nm by 5900 deg cm^2 dmol^{-1}. If the L103P substitution completely disrupted the helix formed by residues 98−107, average SSP score would be 0.22, from which $\Delta[\theta]_{MRE}$ at 222 nm calculated to be 3700 deg cm^2 dmol^{-1}. Therefore, the L103P substitution may destabilize not only the helix in residues 98−107 but also the other helix. Another likely possibility is that the helical fraction in residues 98−107 is greater for CHIBL than for CHIBLΔF. The A123P and I132P substitutions increased the ellipticity by 6400 and 5500 deg cm^2 dmol^{-1}, respectively. Assuming that the A123P or I132P substitution disrupts the entire helix in residues 114−135, the expected $\Delta[\theta]_{MRE}$ at 222 nm is 8700 deg cm^2 dmol^{-1}, indicating that the A123P and I132P substitutions may only partially disrupt the helix. The change induced by the A123P/I132P double mutation (8700 deg cm^2 dmol^{-1}) is similar to that expected for complete disruption of the helix in residues 114−135. In conclusion, the results of the proline substitution study are consistent with the results presented herein, although apparently prolines do not always fully disrupt a helix.

Helix Stabilization by Cys106−Cys119. Cys106−Cys119 significantly enhances the stabilities of the G- and H-strands non-native α-helices. How does Cys106−Cys119 stabilize these helices? One possibility is that the side chains of the two helices interact to stabilize the helical conformations. The presence of Cys106−Cys119 would therefore increase the effective concentration of the interacting helical side chains, which, in turn, would increase stability. In the CHIBLΔF NOESY spectrum, however, long-range NOEs associated with such side chain interactions were not observed. A second possibility is that the entropy effect associated with helix nucleation is modulated by the disulfide. The helix−coil transition is approximated by Zimm−Bragg theory[35] or Lifson−Roig theory.[36] Both theories define two parameters: a

dx.doi.org/10.1021/bi2013239 | Biochemistry 2011, 50, 10590−10597

nucleation parameter and a propagation parameter. For nucleation, three consecutive residues must adopt helical $\phi-\psi$ angles so that an $i,i+4$ hydrogen bond can form, whereas for propagation only one residue is required to form an additional hydrogen bond. We expect that the loop formed by Cys106−Cys119, which contains only 14 residues, should reduce the number of possible conformations allowed for these 14 residues such that the entropic cost of nucleation is decreased in comparison with that required for nucleation of an "open" chain. In CHIBLΔF, the C-terminus of the helix 98−107 and the N-terminus of the helix 114−135 are located within the loop formed by Cys106−Cys119. Therefore, both helices may be stabilized by the decreased unfavorable entropic requirement for nucleation.

Implication for the Folding of ELG. The structural characteristics of CHIBLΔF are consistent with those identified for the longer fragment, CHIBL, based on a proline substitution study,[15] which suggests that the CHIBLΔF structure is similar to the corresponding region in CHIBL. We also have shown that the structure of CHIBL is similar to that of the corresponding region in the C state of ELG.[15] Therefore, characterization of the CHIBLΔF structure and interactions has increased our understanding of the local structures in partially folded states of the full-length protein and interactions responsible for its formation. As discussed above, the disulfide stabilizes the non-native helices in CHIBLΔF. In other words, the non-native helices are a consequence of the presence of the disulfide. The reduction in helicity upon reduction of the disulfide was also observed for CHIBL and the C state of C66A/C160A (data not shown). Therefore, non-native helix formation may not occur during folding if Cys106−Cys119 is absent. Consequently, a comparison of ELG and human tear lipocalin folding is instructive. ELG and human tear lipocalin are homologous proteins and members of the lipocalin family.[37] Although it has been suggested that human tear lipocalin does not form a non-native helical structure during the burst phase of folding because its sequence has a lower helical propensity compared with ELG,[37] it is also possible that the absence of non-native helical structure during the burst phase of folding may be a consequence of the absence of a disulfide corresponding to Cys106−Cys119.

Recently, Sakurai et al.[38] proposed that the non-native helices formed in A−D strands contribute to suppression of incorrect β-structure. Their discussion is based on the previous hydrogen-exchange experiments for bovine β-lactoglobulin, which suggest that the native-like structure is formed at F−H strands and the C-terminal helix and that the non-native helices are formed in A−D strands.[39] However, our previous proline scanning mutagenesis for ELG showed that the H strand region assumes a non-native helix in the equilibrium analogues of the early kinetic intermediate.[16] The present NMR result supports this conclusion. At least for ELG, therefore, the non-native helix is formed at the H strand region. It should be addressed in the future whether the non-native helix at H strand has a role to suppress incorrect β-structure formation or another role to accelerate/decelerate the succeeding folding process.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Chemical shifts for NMR resonances of CHIBLΔF (Table S1) and chemical shift deviations from random coil values for $^{13}C\alpha$, $^{13}C\beta$, and $^{1}H\alpha$ CHIBLΔF resonances (Figure S1). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Ph: +81-42-691-9444. Fax: +81-42-691-9312. E-mail: ikeguchi@soka.ac.jp.

■ REFERENCES

(1) Fersht, A. R., Matouschek, A., and Serrano, L. (1992) The folding of an enzyme. I. Theory of protein engineering analysis of stability and pathway of protein folding. *J. Mol. Biol. 224*, 771−782.
(2) Honig, B., and Yang, A. S. (1995) Free energy balance in protein folding. *Adv. Protein Chem. 46*, 27−58.
(3) Pace, C. N., Shirley, B. A., McNutt, M., and Gajiwala, K. (1996) Forces contributing to the conformational stability of proteins. *FASEB J. 10*, 75−83.
(4) Baase, W. A., Liu, L., Tronrud, D. E., and Matthews, B. W. (2010) Lessons from the lysozyme of phage T4. *Protein Sci. 19*, 631−641.
(5) Arai, M., and Kuwajima, K. (2000) Role of the molten globule state in protein folding. *Adv. Protein Chem. 53*, 209−282.
(6) Chamberlain, A. K., and Marqusee, S. (2000) Comparison of equilibrium and kinetic approaches for determining protein folding mechanisms. *Adv. Protein Chem. 53*, 283−328.
(7) Korzhnev, D. M., Salvatella, X., Vendruscolo, M., Di Nardo, A. A., Davidson, A. R., Dobson, C. M., and Kay, L. E. (2004) Low-populated folding intermediates of Fyn SH3 characterized by relaxation dispersion NMR. *Nature 430*, 586−590.
(8) Korzhnev, D. M., Religa, T. L., Banachewicz, W., Fersht, A. R., and Kay, L. E. (2010) A transient and low-populated protein-folding intermediate at atomic resolution. *Science 329*, 1312−1316.
(9) Feng, H., Zhou, Z., and Bai, Y. (2005) A protein folding pathway with multiple folding intermediates at atomic resolution. *Proc. Natl. Acad. Sci. U. S. A. 102*, 5026−5031.
(10) Kato, H., Feng, H., and Bai, Y. (2007) The folding pathway of T4 lysozyme: the high-resolution structure and folding of a hidden intermediate. *J. Mol. Biol. 365*, 870−880.
(11) Ikeguchi, M., Kato, S., Shimizu, A., and Sugai, S. (1997) Molten globule state of equine beta-lactoglobulin. *Proteins 27*, 567−575.
(12) Fujiwara, K., Arai, M., Shimizu, A., Ikeguchi, M., Kuwajima, K., and Sugai, S. (1999) Folding-unfolding equilibrium and kinetics of equine beta-lactoglobulin: equivalence between the equilibrium molten globule state and a burst-phase folding intermediate. *Biochemistry 38*, 4455−4463.
(13) Kobayashi, T., Ikeguchi, M., and Sugai, S. (2000) Molten globule structure of equine beta-lactoglobulin probed by hydrogen exchange. *J. Mol. Biol. 299*, 757−770.
(14) Yamada, Y., Yajima, T., Fujiwara, K., Arai, M., Ito, K., Shimizu, A., Kihara, H., Kuwajima, K., Amemiya, Y., and Ikeguchi, M. (2005) Helical and expanded conformation of equine beta-lactoglobulin in the cold-denatured state. *J. Mol. Biol. 350*, 338−348.
(15) Nakagawa, K., Yamada, Y., Fujiwara, K., and Ikeguchi, M. (2007) Interactions responsible for secondary structure formation during folding of equine beta-lactoglobulin. *J. Mol. Biol. 367*, 1205−1214.
(16) Nakagawa, K., Tokushima, A., Fujiwara, K., and Ikeguchi, M. (2006) Proline scanning mutagenesis reveals non-native fold in the molten globule state of equine beta-lactoglobulin. *Biochemistry 45*, 15468−15473.
(17) Brownlow, S., Morais Cabral, J. H., Cooper, R., Flower, D. R., Yewdall, S. J., Polikarpov, I., North, A. C., and Sawyer, L. (1997) Bovine beta-lactoglobulin at 1.8 A resolution--still an enigmatic lipocalin. *Structure 5*, 481−495.
(18) Yamada, Y., Nakagawa, K., Yajima, T., Saito, K., Tokushima, A., Fujiwara, K., and Ikeguchi, M. (2006) Structural and thermodynamic consequences of removal of a conserved disulfide bond from equine beta-lactoglobulin. *Proteins 63*, 595−602.
(19) Yamada, Y., Yajima, T., Tsukamoto, S., Nakagawa, K., Fujiwara, K., Kihara, H., and Ikeguchi, M. (2007) Chloride-ion concentration

dependence of molecular dimension in the acid-denatured state of equine beta-lactoglobulin. *J. Appl. Crystallogr. 40*, S213−S216.
(20) Riddles, P. W., Blakeley, R. L., and Zerner, B. (1979) Ellman's reagent: 5,5′-dithiobis(2-nitrobenzoic acid)--a reexamination. *Anal. Biochem. 94*, 75−81.
(21) Gill, S. C., and von Hippel, P. H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem. 182*, 319−326.
(22) Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L., and Sykes, B. D. (1995) 1H, 13C and 15N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR 6*, 135−140.
(23) Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR 6*, 277−293.
(24) Kobayashi, N., Iwahara, J., Koshiba, S., Tomizawa, T., Tochio, N., Guntert, P., Kigawa, T., and Yokoyama, S. (2007) KUJIRA, a package of integrated modules for systematic and interactive analysis of NMR data directed to high-throughput NMR structure studies. *J. Biomol. NMR 39*, 31−52.
(25) Wilkins, D. K., Grimshaw, S. B., Receveur, V., Dobson, C. M., Jones, J. A., and Smith, L. J. (1999) Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques. *Biochemistry 38*, 16424−16431.
(26) Wishart, D. S. (2011) Interpreting protein chemical shift data. *Prog. Nucl. Magn. Reson. Spectrosc. 58*, 62−87.
(27) Dyson, H. J., and Wright, P. E. (2002) Insights into the structure and dynamics of unfolded proteins from nuclear magnetic resonance. *Adv. Protein Chem. 62*, 311−340.
(28) Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., and Sykes, B. D. (1995) 1H, 13C and 15N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *J. Biomol. NMR 5*, 67−81.
(29) Marsh, J. A., Singh, V. K., Jia, Z., and Forman-Kay, J. D. (2006) Sensitivity of secondary structure propensities to sequence differences between alpha- and gamma-synuclein: implications for fibrillation. *Protein Sci. 15*, 2795−2804.
(30) Rohl, C. A., Chakrabartty, A., and Baldwin, R. L. (1996) Helix propagation and N-cap propensities of the amino acids measured in alanine-based peptides in 40 volume percent trifluoroethanol. *Protein Sci. 5*, 2623−2637.
(31) Dyson, H. J., and Wright, P. E. (1998) Equilibrium NMR studies of unfolded and partially folded proteins. *Nat. Struct. Biol. 5* (Suppl.), 499−503.
(32) Eliezer, D., Chung, J., Dyson, H. J., and Wright, P. E. (2000) Native and non-native secondary structure and dynamics in the pH 4 intermediate of apomyoglobin. *Biochemistry 39*, 2894−2901.
(33) Sali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol. 234*, 779−815.
(34) Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr. 26*, 283−291.
(35) Zimm, B. H., and Bragg, J. K. (1959) Theory of the Phase Transition between Helix and Random Coil in Polypeptide Chains. *J. Chem. Phys. 31*, 526−535.
(36) Lifson, S., and Roig, A. (1961) On the Theory of Helix---Coil Transition in Polypeptides. *J. Chem. Phys. 34*, 1963−1974.
(37) Tsukamoto, S., Yamashita, T., Yamada, Y., Fujiwara, K., Maki, K., Kuwajima, K., Matsumura, Y., Kihara, H., Tsuge, H., and Ikeguchi, M. (2009) Non-native alpha-helix formation is not necessary for folding of lipocalin: Comparison of burst-phase folding between tear lipocalin and beta-lactoglobulin. *Proteins 76*, 226−236.
(38) Sakurai, K., Fujioka, S., Konuma, T., Yagi, M., and Goto, Y. (2011) A Circumventing Role for the Non-Native Intermediate in the Folding of beta-Lactoglobulin. *Biochemistry 50*, 6498−6507.
(39) Kuwata, K., Shastry, R., Cheng, H., Hoshino, M., Batt, C. A., Goto, Y., and Roder, H. (2001) Structural and kinetic characterization of early folding events in beta-lactoglobulin. *Nat. Struct. Biol. 8*, 151−155.