Structural Characterization of the Molten Globule State of Apomyoglobin by Limited Proteolysis and HPLC-Mass Spectrometry†

Yeoun Jin Kim,[‡,§] Young A Kim,[§,⊥] Nokyoung Park,[§] Hyeon S. Son,[||,#] Kwang S. Kim,[||] and Jong Hoon Hahn*[,§]

Department of Chemistry, Division of Molecular and Life Sciences, and National Creative Research Initiative Center for Superfunctional Materials, Pohang University of Science and Technology, Pohang, 790-784, South Korea

Received February 18, 2005; Revised Manuscript Received March 24, 2005

ABSTRACT: A method to characterize the structural conformation of an acidic molten globule apomyoglobin (apoMb) at pH 4.2 was developed using limited proteolysis and HPLC−mass spectrometry (HPLC−MS). Endoproteinase Glu-C, which has a double maximum activity at pH 4.0 and pH 7.8 toward glutamic acid (Glu), was used as a proteolytic enzyme. Using this method enabled us to compare the proteolytic cleavages of native apoMb (at pH 8.0) and molten globule (at pH 4.2) directly. Only the first cleavage event in each molecule was considered as reflecting original structural information since the original structure of the protein can be altered after the fist cleavage. Structural changes of apoMb in various pH conditions were studied here to elucidate the local helicity of molten globule apoMb. Among 13 Glu sites, only Glu83 and Glu85 in the F-helix were cleaved at pH 8.0, which confirms that only helix F is frayed upon removal of heme group. At acidic molten globule state, rapid cleavages at Glu38, Glu52, Glu54, Glu85, and Glu148 were detected, while the remaining eight sites were protected. Glu6 and Glu18 in the A-helix, and Glu105 in the G-helix were protected due to the helicity of the secondary structures. The cleavage at Glu38 and the protection at Glu41 in the C-helix indicate that the first half of the C-helix is frayed and the second half of the C-helix is structured. Cleavage at both Glu52 and Glu54 in the D-helix proves that the D-helix is disordered. The N-terminal end of the E-helix at Glu59 was protected, and the beginning of the F-helix was protected by aid of the pH-induced C-cap of the E-helix. The cleavage at Glu148 in H suggests that the C-terminal end of the H-helix is disordered. The A-helix and the first half of the B-helix were highly stable.

Protein folding, the process by which a protein folds from a linear chain of amino acid residues into its biologically active three-dimensional structure is still not well-understood (*1*). It is now generally accepted that one or more distinct, populated intermediates are involved in many cases of the folding process. Statistical models of folding occurring through multiple microscopic pathways have been developed (*2*, *3*). Although knowing the structures of the intermediates is important to understand the hierarchic folding pathway, their structural characterization is very difficult because they exist only transiently during the folding process. It has been shown that several globular proteins exist in equilibrium molten globule states when they are placed under mildly denaturing conditions (e.g., acidic pH and low-salt concentration, or moderate denaturant concentration) (*4*). The importance of studying the equilibrium molten globule is in its greater stability that affords a better chance of determining the structure. The structural detail of the molten globule may help to explain the factors that govern the rate and mechanism of protein folding (*5*). One of the most extensively characterized molten globule states is the I_1 state of apomyoglobin (apoMb)[1] which is observed in equilibrium at pH 4.2 and low-salt concentration (*6*, *7*). A structural similarity between this acidic form of apoMb and an early folding intermediate has been reported (*8*).

Myoglobin (Mb) is composed of eight helices, and the helices are connected by seven loops. The structure of Mb is known to be partially unfolded when its heme group is removed. Eliezer et al. suggested that the A, B, C, D, E, G, and H helices (but not the F-helix) are entirely intact in native apoMb based on an NMR spectroscopic study (*9*). Then, Fontana et al. reported that F is disordered in apoMb using limited proteolysis (*10*, *11*). They concluded that removing the heme group from the holoMb results in fraying of the F-helix that provides a histidine for binding the heme. ApoMb loses more helicity at pH 4.2 as it enters the molten

† This work has been financially supported by the Center for Advanced Bioseparation Technology, Inha University, South Korea.

* To whom correspondence should be addressed: E-mail, hahn@postech.ac.kr; phone, +82-54-279-2118; fax, +82-54-279-8365.

§ Department of Chemistry, Division of Molecular and Life Sciences, Pohang University of Science and Technology.

|| National Creative Research Initiative Center for Superfunctional Materials and Department of Chemistry, Pohang University of Science and Technology.

‡ Current address: Celera Genomics, 45 West Gude Dr., Rockville, MD 20850.

⊥ Current address: Samsung Advanced Institute of Technology, P.O. Box 111, Suwon 440-600, South Korea.

Current address: Laboratory for Computational Biology and Bioinformatics, School of Public Health, Seoul National University, 28 Yongon-Dong Chongno-Gu, Seoul 110-799, South Korea.

[1] Abbreviations: Mb, myoglobin; HPLC−MS, high-performance liquid chromatography−mass spectrometry; *m*/*z*, mass-to-charge ratio; TIC, total ion chromatogram; ESI, electrospray ionization; Glu, glutamic acid.

10.1021/bi050305w CCC: $30.25
Published on Web 04/27/2005

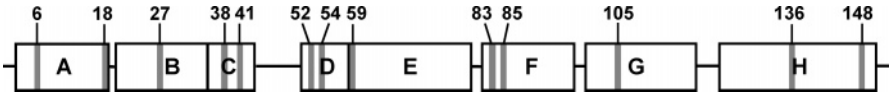

FIGURE 1: Schematic block model of a native holoMb. Eight boxes (A−H) and lines along the protein chain indicate eight helices and loops, respectively. Thirteen Glu residues are indicated with gray bars in the boxes. The size of boxes, bars, and loops is proportional to the actual number of amino acids in each group.

globule state. Circular dichroism and NMR spectroscopic studies have shown that only three of the eight α-helices (A, G, and H) in Mb are structured in the I_1 state where they form a hydrophobic A-G-H core (*7*, *12*, *13*). A subsequent study using multidimensional NMR spectroscopy showed that part of the B-helix also participates in this compact core and some parts of the C, D, and E helices are helical as well (*2*).

We have developed a mass spectrometric method to monitor the local rigidities of the apoMb molten globule to explain its structural status. Our strategy is monitoring the limited proteolysis of acidic apoMb molten globule (at pH 4.2) compared to native apoMb (at pH 8.0) using HPLC−electrospray ionization mass spectrometry (HPLC−ESI MS). Limited proteolysis has proven to be a useful and reliable method for probing the structures and dynamics of proteins in solution (*11*, *14*−*18*). The reasoning here is that proteolytic hydrolysis takes place only at the flexible regions of the protein structure since residues in rigid regions, such as helices, are not accessible to the hydrolytic enzymes. Therefore, identifying cleavage sites during the proteolysis indicates nonstructured regions in the protein. The most important question for this strategy in molten globule studies would be how to compare the two different states based on their proteolytic products in different pH settings. The same enzyme should be applied in both conditions to avoid false positive differential analysis. We used endoproteinase Glu-C (EC 3.4.21.19) which cleaves specifically at the C-terminal glutamic acid in ammonium bicarbonate buffer solution. Using Glu-C as a proteinase provided two important advantages in this study. First, Glu-C has a broad activity range between pH 3.0 and pH 9.0 (*19*). In this range, the enzyme shows a double maximum activity at pH 4.0 and 7.8, and the same specificity and reactivity can be expected at two different pH conditions (*19*). This fact allowed us to compare the molten globule (pH 4.2) and the native apoMb (pH 8.0) directly by probing identical sites in both of them. Second, Glu-C's specificity toward Glu provides reasonable numbers and locations of probing sites for apoMb. ApoMb has 13 glutamic acids evenly distributed over the eight helices, and each helical segment has at least one Glu. None of the loop regions has a Glu reside. Therefore, we can monitor the structural changes of all eight helices by probing each Glu residue without adding complexity attributed to the cleavages of loop region. A schematic block model of the secondary structures of horse heart myoglobin in Figure 1 shows the location of each of the glutamic acid residue in each helix (*20*).

MATERIALS AND METHODS

Preparation of Apomyoglobin. Horse heart apoMb was prepared by 2-butanone extraction of an acidified solution of holoMb (Sigma, St. Louis, MO). A 1% solution of myoglobin was adjusted to pH 1.5 with concentrated HCl at 4 °C. An equal volume of 2-butanone, cooled to 4 °C, was added, and these two phases were thoroughly mixed. The organic layer containing the extracted heme was discarded, and the extraction was repeated twice. The hazy, colorless aqueous layer was dialyzed exhaustively against sodium bicarbonate (50 mg/L) and then with water (*21*).

Limited Proteolysis and HPLC−Mass Spectrometry. ApoMb (1 mg/mL) was dissolved in a 15 mM ammonium bicarbonate buffer solution (pH 8.0). The reaction was initiated by mixing the apoMb solution with endoproteinase Glu-C (EC 3.4.21.19; Sigma, St. Louis, MO) at 24 °C. The enzyme/substrate (E/S) ratio was 1:100 by weight. Aliquots (200 μL each) were taken from the reaction mixture after 1, 2, 4, and 10 min, and 50 μL formic acid was added to each aliquot to stop the reaction by acidification. After evaporation with a Speed-Vac system (Labconco, Kansas City, MO), the samples were reconstituted with 200 μL of 0.1% trifluoroacetic acid (TFA; Sigma, St. Louis, MO) solution, and then subjected to HPLC−ESI MS separation and identification. HPLC system (HP1050, Hewlett-Packard, Palo Alto, CA) was directly connected to a mass spectrometer (Quattro, Micromass, Manchester, U.K.) equipped with an ESI source. A Vydac C8 column (1.0 mm × 250 mm, The Separations Group, Hesperia, CA) was used for separation. Elution was carried out at a flow rate of 40 μL/min with a linear gradient of water/acetonitrile containing 0.1% TFA from 12% to 50% in 70 min.

pH-Dependent Limited Proteolysis of Apomyoglobin. A set of 15 mM ammonium acetate buffer solutions was prepared in various pH conditions (pH 4.2, 5.0, 5.5, 6.0, 6.5, 7.0, 7.5, and 8.0). After incubation of apoMb in each buffer solution for 2 min, the samples were subjected to the same protocol as described above.

Molecular Dynamics. An MD method was employed to simulate the myoglobin segments in low-pH environment. The MD simulation consists of five different stages: (i) initialization, (ii) minimization, (iii) heating, (iv) equilibration, and (v) dynamics. The initialization stage includes the denaturation of the segment F that is disordered in low pH (pH 4.2). The segment was then heated to 5000 K while other segments were fixed at initial positions so that only the segment F could be denatured. The residues His80 and His81 were protonated to reflect the high-proton concentration. This mimics the denatured state of the myoglobin at pH 4.2. Taking this initial structure, 50 000 cycles of ABNR (adopted basis Newton Raphson) energy minimization were performed on the initial system followed by SD (steepest descent) minimization for 4000 cycles. The system temperature was gradually increased from 0 to 300 K in steps of 5 K during the heating stage. Rescaled velocities were assigned to the system every 0.1 ps during heating and every 1 ps during the equilibration period of 20 ps. Having equilibrated the system, 100 ps (1 000 000 steps of 0.000 1 time steps) of further simulation was performed while the system temperature was kept at 300 K by coupling to a temperature bath (*22*). The MD software package CHARMM V26 (*23*)

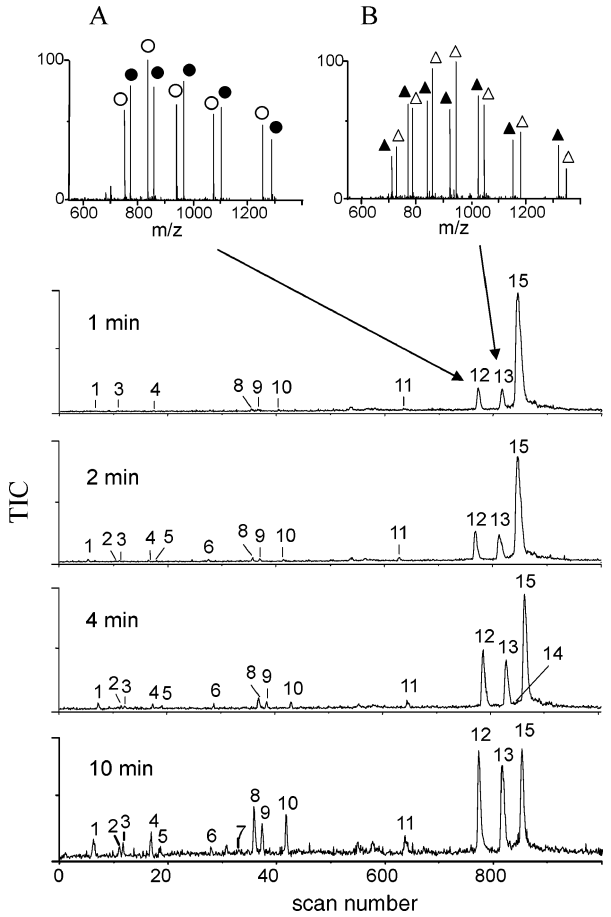

FIGURE 2: Limited proteolysis of apoMb at pH 8.0 monitored by HPLC−ESI MS. Each peak was identified with merged mass spectrum under each peak. Assigned peaks are listed in Table 1. The *x*-axis of the chromatogram is the scan number and the *y*-axis is the total ion current (TIC). (A) ESI mass spectrum of the 12th peak. Two series of the multiple charged ion peaks are deconvoluted to 7 526.5 and 7727 Da and identified as peptide 86−153 and 84−153, respectivaly. (B) ESI mass spectrum of the 13th peak deconvoluted to 9440 and 9242 Da and identified as peptide 1−85 and 1−83.

was used with CHARMM 22 parameter set. Simulations were performed on Silicon Graphics R8000 O2 workstation, and graphics analyses were done using RasMol (Version 2.6-beta-2).

RESULTS AND DISCUSSION

The native apoMb at pH 8.0 and the molten globule apoMb at pH 4.2 were digested with Glu-C. The enzymatic digests quenched at different times were analyzed using HPLC−ESI MS. Limited proteolysis of holoMb was also performed at both pH conditions as a control.

Limited Proteolysis Study of Native ApoMb at pH 8.0. Figure 2 shows total ion current (TIC) chromatograms of four digests obtained after 1, 2, 4, and 10 min incubation of native apoMb with Glu-C at pH 8.0. Peaks were identified by matching its observed molecular mass with the theoretical masses obtained by in silico digestion (Table 1).

Several cleavages were identified, but we limited our consideration only to the first cleavages in the whole protein body (the first nicking event). The first nicking event can be evaluated as having two peptides (matching partners) that cover the whole protein sequence. For example, the first nicking at Glu83 will produce peptides 1−83 and 84−153.

Table 1: Peptide Identification of Each Peak in the Digest of Native ApoMb at pH 8.0

no.	obsd mass[a]	segment	calcd mass	helices
1	1467.4 (±0.3)	137−148	1467.7	H
2	1764.1 (±0.4)	39−52	1764.1	C, D
3	1963.9 (±0.6)	39−54	1964.2	C, D
4	2315.4 (±0.3)	86−105	2315.8	F, G
5	2516.0 (±0.5)	84−105	2516.0	F, G
6	1970.8 (±0.5)	137−153	1970.3	H
7	2780 (±1)	60−85	2780.3	E, F
	2578.3 (±0.6)	60−83	2580.1	E, F
8	3125.9 (±0.8)	55−83	3126.7	D, F
	3326.2 (±0.9)	55−85	3326.9	D, F
9	3525.6 (±0.5)	53−85	3527.1	D, F
10	3276.9 (±0.5)	106−136	3276.7	G, H
11	5774.0 (±0.9)	84−136	5774.7	F, H
	5574.4 (±0.6)	86−136	5574.5	F, H
12	7526.5 (±0.9)	86−153	7526.7	F
	7727 (±1)	84−153	7726.9	F
13	9440 (±1)	1−85	9441.8	F
	9242 (±1)	1−83	9241.6	F
14	6132 (±1)	1−54	6132.9	D
	5931 (±2)	1−52	5932.7	D
15	16951 (±5)	apoMb	16950.5	

[a] Each *m*/*z* value shows the average obtained in the four chromatograms, and the uncertainty is the standard deviation of the corresponding *m*/*z* values in four chromatograms.

If the cleavage at Glu83 was not from the first nicking, you will not detect those counterparts at the same time. This rule is important to eliminate cleavages that are not related with structure. Other consecutive cleavages may occur after losing the original structural information, and thus, those additional cleavages are not directly related to the structure that we want to elucidate in this study.

The first nicking events were found only at Glu83 and Glu85 (in the F-helix) under this condition (pH 8.0). N-terminal parts of the fragments 1−83 (closed triangles in the inserted spectrum B) and 1−85 (open triangles) were found in the 13th peak. The counterparts of those fragments, 84−153 (closed circles in the inserted spectrum A) and 86−153 (open circles) were found in the 12th peak. The smaller peaks (1st ∼ 11th peak), which are increasing with the incubation time, are products of additional cleavages after the initial cleavage at Glu83 and/or Glu85. If these peptides were products of the first nicking, fragments containing intact Glu83 and/or Glu85 should be found, such as 1−136 or 1−148 for the first cleaving of Glu136 or Glu148 in H, 39−153 for Glu38 in D, 55−148 for Glu54, Glu148 in D and H, and so on. However, none of these fragments was found throughout the four sets of chromatograms. The relative intensities of the coeluted peptides, 1−85 and 1−83 in Figure 2A, 86−153 and 84−153 in Figure 2B, were nearly equal. This result shows the proteolysis rates at the Glu83 and Glu85 are very similar, which supports the fact that the F-region is fully disordered so that enzymatic cleavages on both sites are not discriminative. In conclusion, the only flexible chain in the native apoMb at pH 8.0 is the F-region. This result supports the fact that the helicity of the F-helix depends on the presence of the heme group, and it is consistent with the previous study using other proteases (*11*). The structural change upon presence of heme group features an interesting fact regarding structural−functional relationship; fraying the F-helix without changing the entire structure would be the most efficient way to allow reception of the

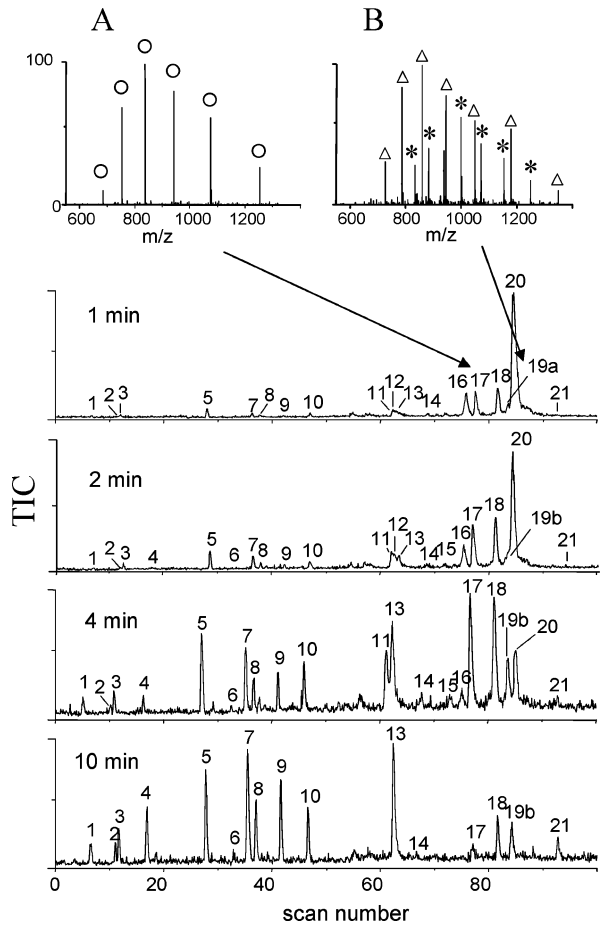

FIGURE 3: Limited proteolysis of apoMb at pH 4.2 monitored by HPLC−ESI MS. The experimentals are the same as in the method described in Figure 2. (A) ESI mass spectrum of the 17th peak deconvoluted to 7 527.1 Da and identified as peptide 86−153. (B) ESI mass spectrum of the 18th peak deconvoluted to 9 440.8 and 14 998 Da and identified as peptide 1−85 and 1−136 , respectively.

heme group back to the myoglobin so that it could secure its biological function. Once it obtains the heme, the protein changes its local structure to keep the heme safely by refolding the F-helix.

It is notable that Glu6 and Glu18 in the A-helix and Glu27 in the B-helix were not cleaved by the enzyme within 10 min of incubation even after the local structures were destructed by the cleavages at Glu83 and/or Glu85 in the F-region. The cleavages at the other Glu's were found. This result shows that deconstruction of neighboring domains does not cause unfolding of the A and B helices. The structural independence of these regions indicates the possibility that those helices have a high propensity for spontaneous helix formation. However, there is an opportunity that the peptides of this region become rapidly aggregated following cleavage of neighboring regions.

Limited Proteolysis Study of Molten Globule ApoMb at pH 4.2, Helix C and D. Figure 3 shows the TIC chromatograms of the digests acquired after 1, 2, 4, and 10 min incubation at pH 4.2. A total of 26 peptides from 21 peaks were identified as listed in Table 2. Byproducts from nonspecific cleavages were not detected at this low-pH condition. While Glu6 and Glu18 in the A-helix, Glu27 in the B-helix, and Glu59 in the E-helix are protected here as they were at pH 8.0, the other helices had altered their proteolytic susceptibility at this lower pH.

Table 2: Peptide Identification of Each Peak in the Digest of Molten Globule ApoMb at pH 4.2.

no.	obsd mass	segment	calcd mass	helices
1	1467.2 (±0.3)	137−148	1467.7	H
2	1764.1 (±0.6)	39−52	1764.1	C, D
3	1963.9 (±0.7)	39−54	1964.2	C, D
4	2315.3 (±0.8)	86−105	2315.8	F, G
5	1970.1 (±0.5)	137−153	1970.3	H
6	2779.6 (±0.8)	60−85	2780.2	E, F
7	3326.3 (±0.6)	55−85	3326.9	D, F
8	3526.6 (±0.3)	53−85	3527.1	D, F
9	3277.2 (±0.7)	106−136	3276.7	G, H
10	5271.3 (±0.8)	39−85	5273.1	C, F
11	8882.2 (±0.8)	55−136	8883.4	D, H
12	10829 (±1)	39−136	10829.6	C, H
13	5573.4 (±0.9)	86−136	5574.5	F, H
13	5228.4 (±0.9)	106−153	5229.0	G
115	7023.9 (±0.7)	86−148	7024.2	F, H
16	10834 (±2)	55−153	10835.6	D
	11035 (±2)	53−153	11035.8	D
	12781 (±3)	39−153	12781.9	C
17	7527.1 (±0.9)	6−153	7526.7	F
18	9440.8 (±1)	1−85	9441.8	F
	14998 (±3)	1−136	14998.3	H
19a	16442 (±6)	1−148	16447.9	H
19b	6132.8 (±0.5)	1−54	6132.9	D
	5931 (±1)	1−52	5932.7	D
20	16948 (±5)	apoMb	16950.5	
21	4185.9 (±0.6)	1−38	4186.7	C

The 11th, 12th, and 16th peaks contain peptides released by cleavages on helix C and D. The 16th peak contains peptides 53−153, 55−153, and 39−153. These peptides were released by cleavage of the Glu52 and Glu54 in the D-chain and Glu38 in the C-chain. Peptides 55−136 and 53−136 (in the 11th peak) and 39−136 (in the 12th peak) were also detected. These cleavages indicate that parts of C and D are disordered in the molten globule state, while they form helical structures at pH 8.0. However, no evidence supporting the first nicking of the Glu41 in the C was found, and the Glu41 was still intact after 10 min. Therefore, the fluctuating region in the C-helix can be narrowed down to only the first half of the C-helix.

Helix E and F. The E-helix has only one probe site at the very first residue, Glu59. The E-helix has been reported to be disordered in the molten globule state. However, Glu59 was protected in this study. Therefore, at least the N-terminal end of the E-helix is more structured than previously thought. We expected that the Glu83 and the Glu85 in the F would be cleaved at pH 4.2 as they were at pH 8.0, or cleaved even more easily because the entire structure is less compact. Surprisingly, however, Glu83 was protected in this case. Only peptides released by Glu85 cleavage, 86−153 (open circles in spectrum A of Figure 3) and 1−85 (open triangle in spectrum B), were found. Neither 84−153 nor 1−83 was found throughout the chromatograms. The coexisting peptide with 1−85 in spectrum B (indicated with crosses) is a peptide 1−136, which was not found in the Figure 2 at the pH 8.0 experiment. In addition, any peptides that could be released after cleavage of Glu83 such as 84−105, 60−83, 55−83, 53−83, and 84−153 were not detected.

The protection of the Glu83 may be attributed to the extended rigidity of the C-terminus E-helix rather than partial refolding of the F-helix. In an α-helix, the first four NH groups and the last four C═O groups, potential proton donors and acceptors, necessarily lack intrahelical hydrogen bonding

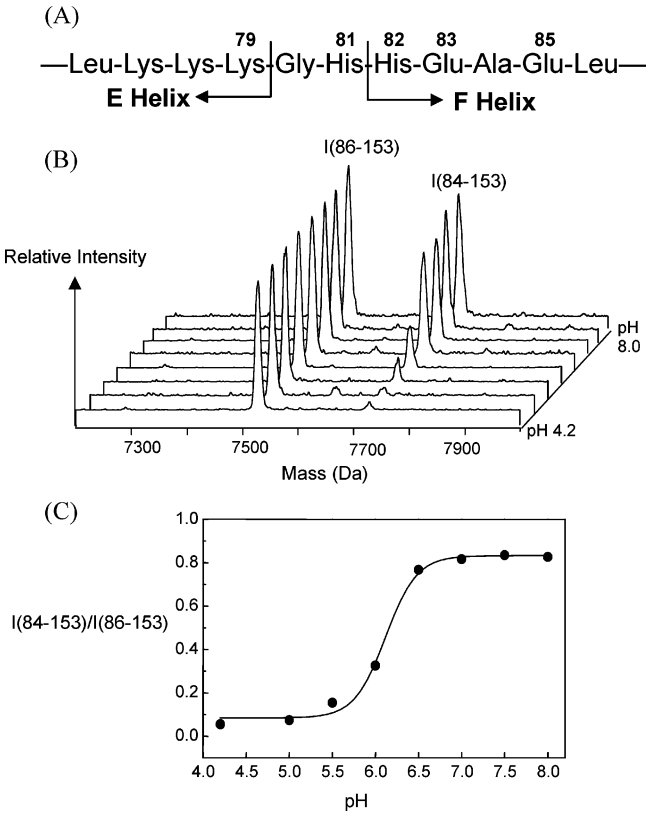

FIGURE 4: (A) Partial amino acid sequence in the probing region. (B) Transformed spectra of 84−153 and 86−153 peptide peaks obtained at various pH conditions (pH 4.2, 5.0, 5.5, 6.0, 6.5, 7.0, 7.5, and 8.0). Relative intensities of 84−153 peptide (7 726.9 Da), $I_{(84-153)}$, increasing with pH values, measured with the constant intensity of 86−153 peptide (7 526.7 Da), $I_{(86-153)}$, as internal standard. (C) Intensity ratio, $I_{(84-153)}$ /$I_{(86-153)}$ versus pH. The inflection point in this plot is at pH 6.14.

pairs. They are sometimes involved in the helix capping process via hydrogen bonding between α-helix residues and flanking residues (*24*−*28*). Some proteins are stabilized by capping motifs, and the fraying of α-helices is prevented (*24*). While there is no chance to form a C-terminal cap (C-cap) at the end of the E-helix at pH 8.0, the protonation of the imidazole side chain of His82 in the F at pH 4.2 can provide a proton donor, and thereby the His82 can form a hydrogen bond with the terminal backbone of helix E. Therefore, enzymatic accessibility to Glu83 becomes restricted by the local rigidity of the newly formed C-cap in acidic environment.

A mass spectrometric titration study was designed to identify a turning point of the proteolytic susceptibility at Glu83. Eight different pH conditions (pH 4.2, 5.0, 5.5, 6.0, 6.5, 7.0, 7.5, and 8.0) were prepared, and each apoMb sample was digested for 2 min at each condition, and then the cleavage at the Glu83 was monitored by measuring the intensities of two peptides (86−153 and 84−153) which were generated after the cleavage of Glu85 and Glu83, respectively (Figure 4A). Figure 4B shows that the *I*84−153 (relative intensity of the 84−153), decreases with lowering pH values, while *I*86−153 is constant. We calculated the normalized intensities of *I*84−153 at each point using *I*86−153 as an internal standard, and then plotted them with corresponding pH values as shown in Figure 4C.

pH 6.14 was observed as a midpoint of the rigidity of the Glu83 as shown in Figure 4C. Same analysis was performed

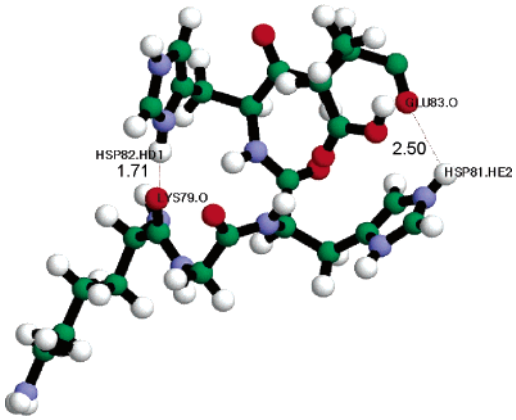

FIGURE 5: Average trajectory structure from the dynamics simulation (data shown from Lys79 to Glu83). Two hydrogen bonds were formed between His82 and Lys79 with a length of 1.71 Å and between His81 and Glu83 with 2.50 Å.

for the counter peptides by monitoring intensities of peptides 1−85 and 1−83, and the inflection point was pH 6.29 (data not shown). Small variability of enzymatic reactivity in each dataset was corrected by intensity normalization using corresponding *I*86−153 as an internal standard. The average value pH 6.2 (±0.1) is very close to the pK_a of the imidazole group of histidine (pK_a = 6.04), and the calculated pK_a values of the His81 and the His82 in native apoMb (6.2 and 6.5, respectively, with ionic strength 0.1 M) (*29*). Since the F-helix in apoMb is fully disordered, we assume that environmental influence altering the pK_a of His can be minimized. This result demonstrates that the protection of Glu83 in the molten globule is pH-driven, and that one of the His residues in the EF-loop and the F might be involved in this transition.

To prove the participation of the His82 to this rigidity, MD (molecular dynamics) simulation studies were performed. Figure 5 shows the average trajectory structure from the MD simulation (from Lys79 to Glu83 shown). Two hydrogen bonds formed during the dynamics simulation, one between His82 and Lys79 with a length of 1.71 Å, the other between His81 and Glu83 with 2.50 Å.

The protonation of the His residues at lower pH enhanced the affinity to bind the oppositely charged oxygens to form the hydrogen bonds. These hydrogen bonds, in turn, made a structural change in that region to make the cleavage of the residue Glu83 difficult. It has been reported that capping stabilizes α-helices in both proteins (*30*, *31*) and peptides (*32*), and these capping motifs can determine supersecondary structure (*25*). Helix capping by hydrogen bonding between the His82 and the Lys79 can strengthen not only the helicity of the E-helix but also the tertiary structure of the molten globule. Therefore, the helix capping at Glu83 shows that the lower pH built up the local rigidity, contributing to the unusual stability of the apoMb in the acidic molten globule (*24*, *27*, *28*).

Helix H. The H-helix has been known as a part of the hydrophobic A-G-H core of I_1 state, so we expected this helix would not be unfolded at pH 4.2. The H-helix contains two probe sites at Glu136 and Glu148. The first nickings of both sites were found in this study. The 18th and 19th peaks contain peptides 1−136 and 1−148, which are produced by nicking Glu136 and Glu148 in the H-helix. This result clearly shows that the Glu148 is the first nicking site. However, it

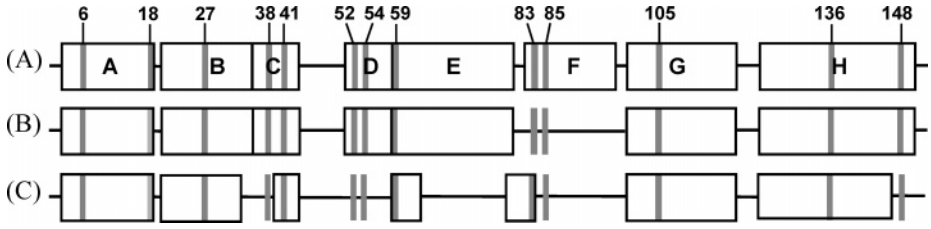

FIGURE 6: (A) Schematic block diagram of a native holoMb. (B) Schematic block diagram of a native apoMb at pH 8.0. Since the F-segment is frayed after heme extraction, both of the vulnerable Glu83 and Glu85 are to be cleaved by endoproteinase Glu-C. (C) Schematic block diagram of a molten globule apoMb at pH 4.2.

is not clear that the Glu136 is the first nicking site because 1−136 and 137−153 can also be products of additional cleavages after the first nicking at Glu148. Therefore, we suggest that at least the C-terminal end of the H-helix is fluctuating, and there is a possibility that the extended region of the H-helix is flexible.

HoloMb. As a control, native holoMb was also digested at both pH 8.0 and 4.2. No cleavage was found at either pH even after 2 days (data not shown) since no Glu is located in any of the loop regions. This result convinces that the cleavages in apoMb were driven by structural flexibility.

Summary of the Structure of ApoMb at pH 4.2. Local helicities of three different states, holoMb, native apoMb, and molten globule, are shown in Figure 6 based on our study combined with the complimentary works reported before (*2*).

The A-helix is structured; at least the center of B is structured; the first half of the C-helix is disordered, and the second half of the C is structured. The entire D is unfolded; the N-terminal end of the E-helix was structured, and the C-terminal end of the E was also structured by the aid of pH-induced helix capping; the F is unfolded, although the very first part of the F gains the rigidity attributed to the C-cap of the E-helix; the G-helix is structured; the first half of the H-helix is ambiguous, but the C-terminal end of the H is unfolded.

ACKNOWLEDGMENT

We acknowledge Dr. Henry M. Fales, Dr. D. Eric Anderson, and Dr. Sonja Hess at the National Institutes of Health (Bethesda, MD) for their helpful discussions and advice.

REFERENCES

1. Ogasahara, K., and Yutani, K. (1994) Unfolding-refolding kinetics of the tryptophan synthase alpha subunit by CD and fluorescence measurements, *J. Mol. Biol. 236*, 1227−1240.
2. Eliezer, D., Yao, J., Dyson, H. J., and Wright, P. E. (1998) Structural and dynamic characterization of partially folded states of apomyoglobin and implications for protein folding, *Nat. Struct. Biol. 5*, 148−155.
3. Wolynes, P. G., Onuchic, J. N., and Thirumalai, D. (1995) Navigating the folding routes, *Science 267*, 1619−1620.
4. Polverino de Laureto, P., De Filippis, V., Di Bello, M., Zambonin, M., and Fontana, A. (1995) Probing the molten globule state of alpha-lactalbumin by limited proteolysis, *Biochemistry 34*, 12596−12604.
5. Garcia, C., Nishimura, C., Cavagnero, S., Dyson, H. J., and Wright, P. E. (2000) Changes in the apomyoglobin folding pathway caused by mutation of the distal histidine residue, *Biochemistry 39*, 11227−11237.
6. Barrick, D., and Baldwin, R. L. (1993) Three-state analysis of sperm whale apomyoglobin folding, *Biochemistry 32*, 3790−3796.
7. Hughson, F. M., Wright, P. E., and Baldwin, R. L. (1990) Structural characterization of a partly folded apomyoglobin intermediate, *Science 249*, 1544−1548.
8. Jennings, P. A., and Wright, P. E. (1993) Formation of a molten globule intermediate early in the kinetic folding pathway of apomyoglobin, *Science 262*, 892−896.
9. Eliezer, D., and Wright, P. E. (1996) Is apomyoglobin a molten globule? Structural characterization by NMR, *J. Mol. Biol. 263*, 531−538.
10. Yang, A. S., and Honig, B. (1994) Structural origins of pH and ionic strength effects on protein stability. Acid denaturation of sperm whale apomyoglobin, *J. Mol. Biol. 237*, 602−614.
11. Fontana, A., Zambonin, M., Polverino de Laureto, P., De Filippis, V., Clementi, A., and Scaramella, E. (1997) Probing the conformational state of apomyoglobin by limited proteolysis, *J. Mol. Biol. 266*, 223−230.
12. Loh, S. N., Kay, M. S., and Baldwin, R. L. (1995) Structure and stability of a second molten globule intermediate in the apomyoglobin folding pathway, *Proc. Natl. Acad. Sci. U.S.A. 92*, 5446−5450.
13. Miranker, A., Robinson, C. V., Radford, S. E., Aplin, R. T., and Dobson, C. M. (1993) Detection of transient protein folding populations by mass spectrometry, *Science 262*, 896−900.
14. Cha, S. S., Kim, J. S., Cho, H. S., Shin, N. K., Jeong, W., Shin, H. C., Kim, Y. J., Hahn, J. H., and Oh, B. H. (1998) High resolution crystal structure of a human tumor necrosis factor-alpha mutant with low systemic toxicity, *J. Biol. Chem. 273*, 2153-2160.
15. Zappacosta, F., Pessi, A., Bianchi, E., Venturini, S., Sollazzo, M., Tramontano, A., Marino, G., and Pucci, P. (1996) Probing the tertiary structure of proteins by limited proteolysis and mass spectrometry: the case of Minibody, *Protein Sci. 5*, 802−813.
16. Novotny, J., and Bruccoleri, R. E. (1987) Correlation among sites of limited proteolysis, enzyme accessibility and segmental mobility, *FEBS Lett. 211*, 185−189.
17. Hubbard, S. J., Campbell, S. F., and Thornton, J. M. (1991) Molecular recognition. Conformational analysis of limited proteolytic sites and serine proteinase protein inhibitors, *J. Mol. Biol. 220*, 507−530.
18. Tremblay, J. M., Helmkamp, G. M., and Yarbrough, L. R. (1996) Limited proteolysis of rat phosphatidylinositol transfer protein by trypsin cleaves the C terminus, enhances binding to lipid vesicles, and reduces phospholipid transfer activity, *J. Biol. Chem. 271*, 21075−21080.
19. Drapeau, G. R., Boily, Y., and Houmard, J. (1972) Purification and properties of an extracellular protease of *Staphylococcus aureus*, *J. Biol. Chem. 247*, 6720−6726.
20. Evans, S. V., and Brayer, G. D. (1990) High-resolution study of the three-dimensional structure of horse heart metmyoglobin, *J. Mol. Biol. 213*, 885−897.
21. Rothgeb, T. M., and Gurd, F. R. (1978) Physical methods for the study of myoglobin, *Methods Enzymol. 52*, 473−486.
22. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. T., Dnola, A. J., Haak, R. (1984) Molecular dynamics with coupling to an external bath, *J. Chem. Phys. 81*, 3684−3690.
23. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. J. (1983) CHARMM: a program for macromolecular energy, minimization, and dynamics calculations, *J. Comput. Chem. 4*, 187−217.
24. Lanigan, M. D., Tudor, J. E., Pennington, M. W., and Norton, R. S. (2001) A helical capping motif in ShK toxin and its role in helix stabilization, *Biopolymers 58*, 422−436.

25. Aurora, R., and Rose, G. D. (1998) Helix capping, *Protein Sci. 7*, 21−38.
26. Presta, L. G., and Rose, G. D. (1988) Helix signals in proteins, *Science 240*, 1632−1641.
27. Tripet, B., and Hodges, R. S. (2002) Helix capping interactions stabilize the N-terminus of the kinesin neck coiled-coil, *J. Struct. Biol. 137*, 220−235.
28. Iovino, M., Falconi, M., Petruzzelli, R., and Desideri, A. (2001) Role of the helix capping in the stability of the mouse prion (180−213) segment: investigation through molecular dynamics simulations, *J. Biomol. Struct. Dyn. 19*, 237−246.
29. Yang, A.-S., and Honig, B. (1994) Structural origins of pH and ionic strength effects on protein stability. Acid denaturation of sperm whale apomyoglobin, *J. Mol. Biol. 237*, 602−614.
30. Serrano, L., Fersht, A. R. (1989) Capping and alpha-helix stability, *Nature 342*, 296−299.
31. Bell, J. A., Becktel, W. J., Sauer, U., Baase, W. A., and Matthews, B. W. (1992) Dissection of helix capping in T4 lysozyme by structural and thermodynamic analysis of six amino acid substitutions at Thr 59, *Biochemistry 31*, 3590−3596.
32. Reymond, M. T., Huo, S., Duggan, B., Wright, P. E., and Dyson, H. J. (1997) Contribution of increased length and intact capping sequences to the conformational preference for helix in a 31-residue peptide from the C terminus of myohemerythrin, *Biochemistry 36*, 5234−5244.

BI050305W