Nicked Apomyoglobin: A Noncovalent Complex of Two Polypeptide Fragments Comprising the Entire Protein Chain†

Valeria Musi, Barbara Spolaore, Paola Picotti, Marcello Zambonin, Vincenzo De Filippis, and Angelo Fontana*

CRIBI Biotechnology Centre, University of Padua, Viale G. Colombo 3, I-35121 Padua, Italy

Received January 23, 2004; Revised Manuscript Received March 10, 2004

ABSTRACT: Limited proteolysis of the 153-residue chain of horse apomyoglobin (apoMb) by thermolysin results in the selective cleavage of the peptide bond Pro88−Leu89. The N-terminal (residues 1−88) and C-terminal (residues 89−153) fragments of apoMb were isolated to homogeneity and their conformational and association properties investigated in detail. Far-UV circular dichroism (CD) measurements revealed that both fragments in isolation acquire a high content of helical secondary structure, while near-UV CD indicated the absence of tertiary structure. A 1:1 mixture of the fragments leads to a tight noncovalent protein complex (1−88/89−153, nicked apoMb), characterized by secondary and tertiary structures similar to those of intact apoMb. The apoMb complex binds heme in a nativelike manner, as given by CD measurements in the Soret region. Second-derivative absorption spectra in the 250−300 nm region provided evidence that the degree of exposure of Tyr residues in the nicked species is similar to that of the intact protein at neutral pH. Also, the microenvironment of Trp residues, located in positions 7 and 14 of the 153-residue chain of the protein, is similar in both protein species, as given by fluorescence emission data. Moreover, in analogy to intact apoMb, the nicked protein binds the hydrophobic dye 1-anilinonaphthalene-8-sulfonate (ANS). Taken together, our results indicate that the two proteolytic fragments 1−88 and 89−153 of apoMb adopt partly folded states characterized by sufficiently nativelike conformational features that promote their specific association and mutual stabilization into a nicked protein species much resembling in its structural features intact apoMb. It is suggested that the formation of a noncovalent complex upon fragment complementation can mimic the protein folding process of the entire protein chain, with the difference that the folding of the complementary fragments is an intermolecular process. In particular, this study emphasizes the importance of interactions between marginally stable elements of secondary structure in promoting the tertiary contacts of a native protein. Considering that apoMb has been extensively used as a paradigm in protein folding studies for the past few decades, the novel fragment complementing system of apoMb here described appears to be very useful for investigating the initial as well as late events in protein folding.

Protein folding is a very important and yet unsolved problem of modern biology and biophysics. It has been found that, with relatively large proteins, the polypeptide chain reaches the final, biologically active structure of the native protein by a hierarchical process that involves intermediates (*1*). Partly folded or molten globule (MG)[1] states of proteins have been proposed to act as folding intermediates (*2*−*4*). MG states, defined as possessing a nativelike content of secondary structure but lacking the specific tertiary interactions of the native protein (*2*, *3*), are those adopted at equilibrium by proteins under mildly denaturing conditions, such as low pH. In a number of cases, MGs have been shown to share the molecular features of the transient intermediates observed in kinetic experiments of protein folding (*5*). However, the view that protein folding occurs through a set of well-defined intermediates has been challenged by the "folding funnel" view. Indeed, it has been proposed that there are multiple pathways of folding from an unfolded polypeptide chain to the final structure (*6*−*10*).

A quite common and useful approach to study the mechanism of protein folding involves analysis of the conformational preferences of peptide fragments of a protein (*11*, *12*). NMR studies provided evidence that relatively short, linear peptides can acquire some secondary structure, even if flickering and in rapid equilibrium with unfolded states. These observations lent support to the idea that the protein folding process is initiated by the transient formation of localized regions of secondary structure and local hydrophobic clusters (*12*, *13*). On the other hand, relatively long protein fragments can acquire various levels of both secondary and tertiary structures. In particular, autonomous folding

† This work was supported by the Italian Ministry of University and Scientific Research (FIRB Project on Protein Folding and PRIN-2003).

* To whom correspondence should be addressed: telephone +39-049-8276156; fax +39-049-8276159; e-mail angelo.fontana@unipd.it.

[1] Abbreviations: apoMb, apomyoglobin; N-fragment, apomyoglobin fragment 1−88; C-fragment, apomyoglobin fragment 89−153; nicked apoMb, noncovalent complex of apomyoglobin fragments 1−88 and 89−153; CD, circular dichroism; 3D, three-dimensional; RP-HPLC, reverse-phase high-performance liquid chromatography; MG, molten globule; NMR, nuclear magnetic resonance; Tris, tris(hydroxymethyl)-aminomethane; TFA, trifluoroacetic acid; E/S, enzyme-to-substrate ratio; [θ], mean residue ellipticity; PAGE, polyacrylamide gel electrophoresis; SDS, sodium dodecyl sulfate; UV, ultraviolet; ANS, 1-anilinonaphthalene-8-sulfonate; ESI, electrospray ionization; MS, mass spectrometry.

10.1021/bi0400103 CCC: $27.50
Published on Web 04/30/2004

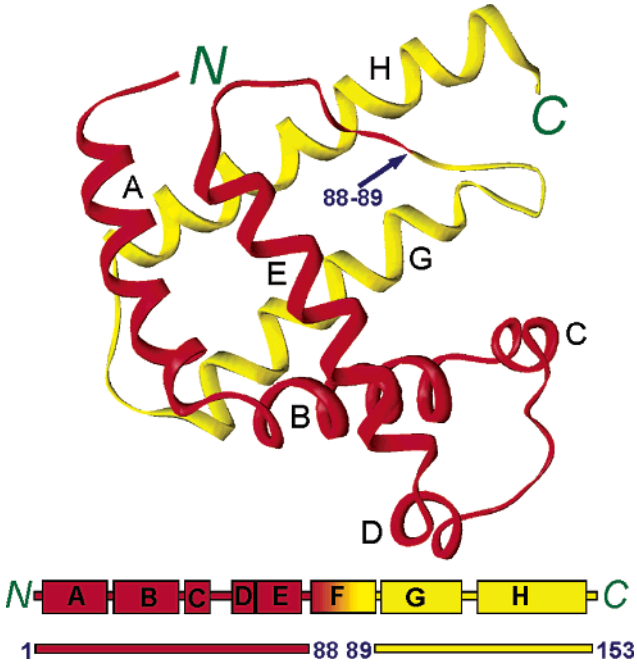

FIGURE 1: Schematic three-dimensional structure (top) of horse heart apomyoglobin (apoMb). The eight helical segments from A to H of the 153-residue chain of holomyoglobin are also shown (bottom). The arrow indicates the location of the peptide bond that is selectively cleaved by thermolysin. Chain segments 1−88 and 89−153 are colored in red and yellow, respectively. The chain segment encompassing helix F is shown to be disordered in apoMb (see text). The model was constructed from the X-ray structure (file 1YMB taken from the Brookhaven Protein Data Bank) by use of the program WebLab (Molecular Simulations Inc., San Diego, CA).

into a nativelike structure was found to occur with fragments comprising the chain region of a protein domain in the native protein (*14*−*19*). However, the most useful protein fragments for analyzing protein structure−stability−function relationships, as well as folding properties, are those that are able to associate into a stable, nativelike protein complex (*20*). Following the pioneering studies on fragments of ribonuclease A (*21*) and later staphylococcal nuclease (*11*, *22*), several nicked proteins, given usually by the association of two fragments, have been described and the results of these studies provided important information about principles underlying protein structure, dynamics and folding (*23*−*29*; see ref *30* for a review).

Apomyoglobin (apoMb), obtained by removal of heme from the holoprotein, has been extensively studied as a paradigm for protein folding in several laboratories. The best characterized myoglobin is that isolated from sperm whale, but since the structure of the homologous horse protein is essentially identical (*31*), results of experimental studies and their interpretations likely can be used interchangeably for both proteins. At neutral pH, apoMb has a well-defined globular structure, but the F helix is disordered (*32*−*37*) (Figure 1). Under mildly acidic conditions (pH 4.2), apoMb forms intermediate(s) containing mostly the A, G, and H helical regions, while the rest of the protein remains largely unfolded. The molecular and conformational features of the partly folded or MG states adopted by apoMb at low pH have been analyzed by a plethora of biochemical and biophysical techniques, making apoMb the most well characterized protein in terms of MG states (*38*−*43*) and folding mechanism. Measurements by multidimensional NMR and quench-flow hydrogen exchange provided detailed insights into the structure and dynamics of the various states that are considered to occur during the folding of apoMb (*33*, *44*).

A number of conformational studies were conducted on peptide fragments of horse or sperm whale apoMb. Initial studies utilized fragments produced by cleaving the protein by the cyanogen bromide reaction (*45*−*48*). Subsequently, a series of peptides spanning the entire 153-residue chain of the sperm whale protein were prepared by solid-phase chemical synthesis and their conformational properties were investigated by far-UV circular dichroism (CD) and NMR measurements (*49*−*52*). In a recent study, the conformational features of the sperm whale apoMb fragments 1−36, 1−77, and 1−119, prepared by recombinant methods, were analyzed by far-UV CD (*53*). Finally, a rather long fragment of human apoMb (residues 31−104) was studied for its conformational and heme-binding properties (*54*), and a functional horse "minimyoglobin" (residues 32−139) has been described (*55*).

The results of these previous studies prompted us to further apply and extend the protein fragmentation approach for investigating folding properties of apoMb. The results of our previous limited proteolysis experiments conducted on horse apoMb at neutral pH provided the starting point for the molecular dissection of the 153-residue chain of the protein (*36*, *37*). We have shown that limited proteolysis of horse apoMb by four proteases (thermolysin, subtilisin, chymotrypsin, and trypsin) preferentially occurs at the level of helix F (segment 82−97 in horse apoMb). On the other hand, the holoprotein was fully resistant to proteolysis when reacted under identical experimental conditions. These results allowed us to conclude that helix F is highly flexible or disrupted in apoMb, in agreement with the notion that limited proteolysis of globular proteins occurs at flexible sites and that a mechanism of local unfolding of a protein is required for an efficient binding and adaptation of the polypeptide substrate at the specific stereochemistry of the protease's active site (*36*, *37*, *56*, *57*). Of interest, that helix F is disrupted is confirmed by NMR studies of the structural features of apoMb (*32*−*35*), as well as by molecular dynamics simulations (*58*−*61*). While in our previous study we were interested only in locating the flexible site(s) along the apoMb chain by using proteolytic probes, here we have produced and isolated the two large fragments 1−88 and 89−153 that result from the selective proteolysis of the Pro88−Leu89 peptide bond of the protein with thermolysin. These fragments have been studied for their propensities to fold independently. Moreover, it is shown that they associate into a tight 1:1 complex in aqueous solution at neutral pH and that the resulting complex 1−88/89−153 (nicked apoMb) possesses structural features much resembling those of the intact protein. The results of this study complement and extend those previously obtained by others by use of synthetic peptides or protein fragments of apoMb (see above). It is proposed that this novel system of complementing fragments of apoMb can be used to mimic the protein folding process of the entire protein, with the difference that protein

folding via fragment association is an intermolecular process.[2]

MATERIALS AND METHODS

Materials. Horse heart myoglobin and thermolysin from *Bacillus thermoproteolyticus* were obtained from Sigma. Cyanogen bromide (CNBr), sodium dithionite, hemin, acetonitrile, trifluoroacetic acid (TFA), and 1-anilinonaphthalene-8-sulfonic acid (ANS) were purchased from Fluka. The materials used for sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis (PAGE) were from Bio-Rad. All other chemicals were analytical-grade and were obtained from Sigma or Fluka.

Apomyoglobin (apoMb) was obtained from holomyoglobin by removal of the heme by the 2-butanone extraction procedure (*62*, *63*). The possible contamination of apoMb by myoglobin was assessed spectrophotometrycally and no significant absorption was observed in the Soret region.

Preparation of Apomyoglobin Fragments. The N-terminal fragment 1−88 and C-terminal fragment 89−153 were prepared by limited proteolysis of apoMb with thermolysin (*37*). Proteolysis was conducted at 25 °C by incubating apoMb (0.5 mg/mL) in 10 mM Tris-HCl buffer, pH 7.0, containing 1 mM $CaCl_2$, with thermolysin at an E/S ratio of 1:50 (by weight). The reaction was stopped after 1 min with aqueous TFA (final concentration of TFA 0.1%). The proteolysis mixture was then separated by gel filtration on a Sephadex G-50 superfine column (1.6 × 95 cm) eluted with 5% (v/v) aqueous acetic acid at flow rate of 15 mL/h. The effluent from the column was monitored at 280 nm, and fractions of 1.5 mL were collected. The fractions containing the protein fragments were pooled and then concentrated by use of the Speed-Vac system (Savant). To avoid extensive fragment aggregation, care was taken not to dry up the apoMb fragments during the Speed-Vac treatment. Moreover, to eliminate possible aggregates, the fragment solutions utilized for spectroscopic measurements were those obtained after gel filtration on a Superdex-75 column equilibrated and eluted with Tris buffer, pH 7.0. Homogeneity of fragments 1−88 and 89−153 was determined by RP-HPLC, N-terminal sequencing, and mass spectrometry.

Electrophoresis and Chromatography. SDS−PAGE was carried out in a vertical slab gel apparatus (Mini-Protean II, Bio-Rad) with the Tricine buffer system (*64*). The gels were stained with Coomassie brilliant blue R-250. A sample of a partial CNBr cleavage of horse heart apoMb at the level of the two methionine residues in positions 55 and 131 of the 153-residue chain of the protein was used as molecular mass standard. The resulting CNBr fragments are 1−131, 56−153, 56−131, 1−55, and 132−153 (in order of decreasing molecular mass).

Reverse-phase HPLC was performed on a Vydac C_{18} column (4.6 × 250 mm) purchased from The Separations Group. Elution was carried out at a flow rate of 0.8 mL/min with a linear gradient of acetonitrile in water containing 0.05% (v/v) TFA from 40 to 60% in 25 min. The effluent was monitored by recording the absorbance at 226 nm.

Size-exclusion chromatography of the complex 1−88/89−153 (nicked apoMb) was carried out by loading a 1:1 mixture of purified apoMb fragments 1−88 and 89−153 (50 μM each) in 20 mM Tris-HCl/0.2 M NaCl buffer, pH 7.0, onto a Superdex-75 column (type HR 10/30, 1.0 × 30 cm; Pharmacia) equilibrated with the same buffer. Elution was conducted at a flow rate of 0.4 mL/min and the effluent was monitored by recording the absorbance at 226 nm. The Superdex column was calibrated with a protein mixture kit of low molecular mass (Pharmacia).

Spectroscopic Measurements. Circular dichroism (CD) measurements were made at 25 °C on a Jasco J-710 spectropolarimeter (Tokyo, Japan) equipped with a thermostatically controlled cell holder. The instrument was calibrated with *d*-(+)-10-camphorsulfonic acid. The spectra were recorded in 20 mM Tris-HCl/0.2 M NaCl, pH 7.0, at a protein or protein fragment concentration of 0.1 mg/mL in the far-UV and 0.7−1.0 mg/mL in the near-UV region, in 1-mm and 5-mm path length quartz cells, respectively. The results were expressed as mean residue ellipticity $[\theta]$ (deg cm^2 $dmol^{-1}$) calculated from the formula $[\theta] = (\theta_{obs}/10)(MRW/lc)$, where θ_{obs} is the observed ellipticity at a given wavelength, MRW is the mean residue molecular weight (molecular mass of the protein divided by the number of amino acids), l is the optical path length in centimeters, and c the protein concentration in grams per milliliter. Far-UV CD spectra were analyzed in order to estimate the percentage of protein secondary structure from the equation $\% \alpha_{helix} = 100/\{1 + [(\theta_{222nm,obs} - \theta_{222nm,H})/(\theta_{222nm,C} - \theta_{222nm,obs})]\}$, where $\theta_{222\ nm,obs}$ is the observed mean residue ellipticity at 222 nm and $\theta_{222nm,H}$ and $\theta_{222nm,C}$ are the mean residue ellipticity values for the 100% helical and coil conformation, respectively. Figures for $\theta_{222nm,H}$ and $\theta_{222nm,C}$ at 25 °C were taken as −33 500 and −485 deg·cm^2·$dmol^{-1}$, respectively, according to Scholtz et al. (*65*).

Protein and fragment concentrations were determined spectrophotometrically on a Perkin-Elmer Lambda-20 spectrophotometer. The concentrations of stock solutions of apoMb and fragments 1−88 and 89−153 were evaluated from their absorbance at 280 nm. Extinction coefficients at 280 nm for apoMb and fragments 1−88 and 89−153 were calculated according to Gill and von Hippel (*66*) and taken as 0.84, 1.16, and 0.356 mg^{-1} cm^2, respectively.

Fluorescence emission spectra were recorded at 25 °C on a Perkin-Elmer model LS-50B spectrofluorometer, equipped with a thermostatically controlled cell holder. Spectra were taken in Tris/NaCl buffer, pH 7.0, in a cuvette of 1-cm path length.

Second-derivative UV spectra of intact and nicked apoMb were recorded at 25 °C after 30 min of incubation in 20 mM Tris-HCl/0.2 M NaCl, pH 7.0, in the presence or absence of 6 M Gdn·HCl. The protein concentration was 0.12−0.17 mg/mL. The average exposure of tyrosine residues to solvent (α) was calculated according to Ragone et al. (*67*). A model-compound solution of *N*-acetyl-Trp-NH_2 and *N*-acetyl-Tyr-NH_2 at the same Tyr/Trp molar ratio (1:1) as in apoMb was employed.

Heme Binding. The binding of heme to intact and nicked apoMb was investigated by absorption and CD spectroscopy. A stock hemin solution was prepared by dissolving 0.9 mg of hemin in 1.8 mL of 0.1 N NaOH, and the exact concentration of hemin was determined by spectrophotometry

[2] This work was presented at the Fourth European Symposium of the Protein Society, 18−22 April, 2001, Paris, France [(2001) *Protein Sci. 10* (Suppl. 1), 247].

with an extinction coefficient of 5.84×10^4 $M^{-1} \cdot cm^{-1}$ at 385 nm (*68*). Heme stock solutions (usually ~0.4 mM) were diluted with 20 mM Tris-HCl/0.2 M NaCl, pH 7.0, just before use. Reduction of the protein-bound heme was obtained by adding to the heme−protein solution a 1000-fold molar excess of sodium dithionite, taken from a freshly prepared stock solution of the reagent.

ANS Binding. ANS binding to intact or nicked apoMb was monitored by fluorescence spectroscopy of a protein/ANS solution prepared by adding ANS (from a 50 μM stock solution) to intact or nicked apoMb dissolved (2 μM) in 20 mM Tris-HCl/0.2 M NaCl buffer, pH 7.0. To avoid free ANS in the mixture, an ANS/protein molar ratio of 0.36 was employed. The concentration of ANS was determined spectrophotometrically with an extinction coefficient of 5000 M^{-1} cm^{-1} at 350 nm.

Fluorescence emission spectra of ANS bound to intact or nicked apoMb (complex 1−88/89−153) were obtained upon excitation at 390 nm and the emission fluorescence was recorded from 400 to 700 nm. Fluorescence emission spectra of intact or nicked apoMb in the presence of ANS were recorded upon excitation at 280 nm and the emission fluorescence was recorded from 300 to 600 nm. The binding of ANS to intact and nicked apoMb was also evaluated by measuring CD spectra of the protein−ANS complex in the ANS absorption region (340−430 nm). The CD spectra were recorded at room temperature in 20 mM Tris-HCl/0.2 M NaCl buffer, pH 7.0, at a protein concentration of 30 μM in a 1-cm path length quartz cell.

RESULTS

Limited Proteolysis of ApoMb and Fragment Association. Thermolysin was used to fragment apoMb, since this protease cleaves apoMb with high selectivity at the level of the Pro88−Leu89 peptide bond (*37*). Figure 2 shows the SDS−PAGE and RP-HPLC analysis of the proteolysis mixture of apoMb reacted with thermolysin at an E/S ratio of 1:50 (by weight) for 1 min at 25 °C. Clearly, limited proteolysis of apoMb is very efficient and clean in producing the N-terminal fragment 1−88 and the C-terminal fragment 89−153, as shown by RP-HPLC (Figure 2). The relative height of the chromatographic peak of fragment 1−88 is much higher than that of fragment 89−153. This is expected, since the N-terminal fragment contains the two strongly absorbing chromophores of Trp7 and Trp14, while fragment 89−153 contains only Tyr103 and Tyr146 residues. For preparative purposes, it was found convenient to use open-column gel-filtration chromatography in order to isolate the apoMb fragments on a relatively large scale. To this aim, a sample of the proteolysis mixture was applied to a Sephadex G-50 column equilibrated and eluted with 5% aqueous acetic acid. The undigested apoMb is eluted first from the column, followed by fragments 1−88 and 89−153, in agreement with their molecular masses (not shown).

Figure 3 shows the chromatogram of a sample of an equimolar mixture of fragments 1−88 and 89−153 prepared in aqueous buffer at neutral pH and loaded onto a Superdex-75 column eluted with Tris/NaCl buffer, pH 7.0. It is seen that the protein material is eluted from the gel filtration column as a single chromatographic peak, despite the different molecular masses of the two fragments. Moreover,

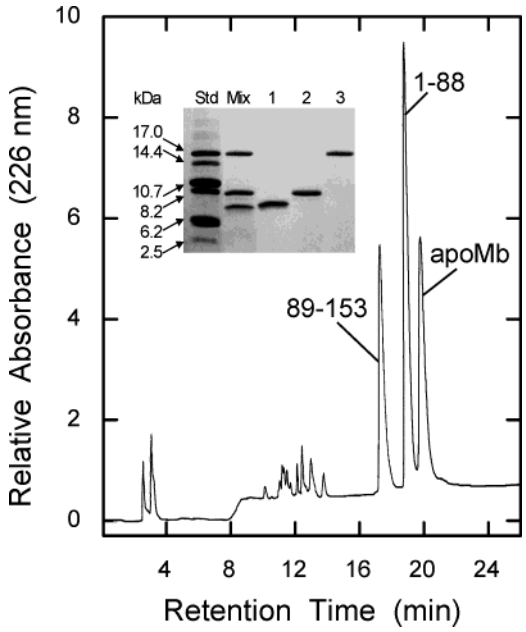

FIGURE 2: RP-HPLC analysis of the proteolysis mixture of apoMb with thermolysin. A Vydac C_{18} column was employed and elution was carried out at a flow rate of 0.8 mL/min with a linear gradient of acetonitrile in water containing 0.05% (v/v) TFA from 5 to 40% in 5 min and from 40 to 60% in 25 min. The effluent was monitored by recording the absorbance at 226 nm. (Inset) SDS−PAGE of the proteolysis mixture of apoMb with thermolysin (Mix) and of the protein material corresponding to the peaks of the RP-HPLC chromatogram. Samples 1−3 contain fragment 89−153, fragment 1−88, and intact apoMb, respectively. Std is a partial CNBr digest of apoMb (see Materials and Methods). Numbers near the electrophoretic bands of the Std sample correspond to the molecular masses in kilodaltons of the CNBr fragments of apoMb.

the elution time of the fragment mixture is almost identical or only slightly lower than that of a sample of intact apoMb analyzed by the same Superdex-75 column. Therefore, fragments 1−88 and 89−153 are eluted from the column as a single protein species at neutral pH, implying a stable complex that is not dissociated during chromatography. On the other hand, the two fragments dissociate and elute as two chromatographic peaks when a sample of the complex 1−88/89−153 (nicked apoMb), isolated after gel filtration, is analyzed by RP-HPLC in acid solution (0.05% aqueous TFA) (see Figure 3, inset). The relative heights of the two chromatographic peaks (absorption at 226 nm) in the RP-HPLC chromatogram derives from the fact that fragment 1−88 contains two Trp residues (see above).

Circular Dichroism Spectroscopy. The conformational features of nicked apoMb and its constituting fragments have been evaluated by CD spectroscopy, a technique extensively used to probe conformational features of proteins (*69*−*72*). In particular, CD measurements in the far-UV region are used to analyze the content of secondary structure of a protein (*69*, *73*), whereas in the near-UV they provide information regarding the specific environment of the aromatic chromophores in the protein fold (*74*, *75*). Holomyoglobin is an all-α globular protein comprising eight α-helices (see Figure 1), and upon removal of the heme, its CD spectrum in the 200−250 nm region retains the overall features of α-helical polypeptides with minima near 208 and 222 nm. As shown in Figure 4A, the far-UV CD spectra of intact and nicked apoMb are quite similar in terms of overall shape, but the intensity of the CD signal of the nicked species is ~10%

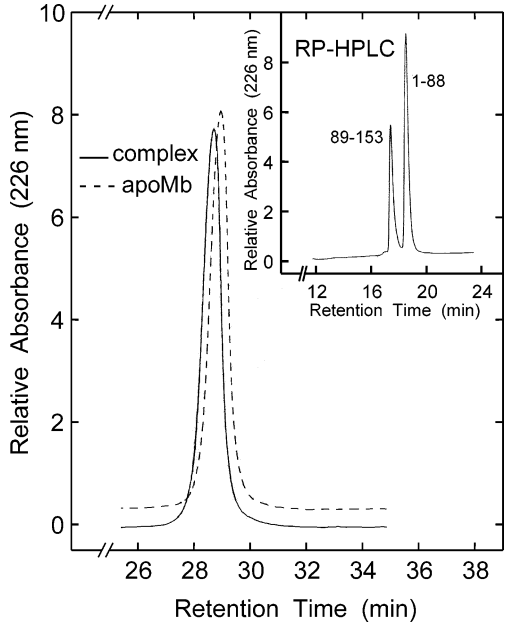

FIGURE 3: Gel filtration of the noncovalent complex 1−88/89−153 (nicked apoMb). A 1:1 mixture of purified apoMb fragments 1−88 and 89−153 in 20 mM Tris-HCl/0.2 M NaCl buffer, pH 7.0, was applied to a Superdex-75 column. Elution was conducted at a flow rate of 0.4 mL/min and the effluent was monitored at 226 nm. A sample of intact apoMb (- - -) was also analyzed by gel filtration. (Inset) RP-HPLC analysis of the noncovalent complex 1−88/89−153 eluted from the gel-filtration column. A Vydac C_{18} column was employed and the chromatographic conditions were those indicated in the caption to Figure 2.

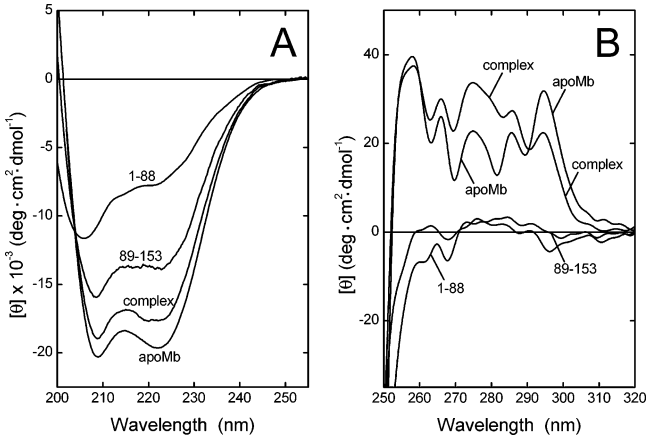

FIGURE 4: Far-UV (A) and near-UV (B) CD spectra of apoMb, nicked apoMb (1−88/89−153), and its constituting fragments 1−88 and 89−153. Spectra were recorded at room temperature in 20 mM Tris·HCl/0.2 M NaCl buffer, pH 7.0, at protein or fragment concentration of 0.1 and 0.7−1.0 mg/mL in the far- and near-UV regions, respectively.

reduced. Therefore, these CD data can be interpreted as indicating that some differences in secondary structure do exist between the nicked and intact proteins. Alternatively, considering that molecular dynamics simulations have shown that motions in protein structures do influence and reduce the CD signal of α-helices, the hypothesis can be advanced that nicked apoMb is less compact and rigid than the intact protein and experiences enhanced structural fluctuations (*58*, *76*).

Both apoMb fragments 1−88 and 89−153 display far-UV CD spectra with the features of α-helical polypeptides, since minima of negative ellipticity occur near 208 and 222 nm (Figure 4A). Analysis of the spectra of the fragments in terms of percent α-helical content allowed us to estimate that the isolated fragments 1−88 and 89−153 adopt in aqueous solvent at neutral pH approximately 25% and 46% helicity, respectively. Therefore, the helical content of the fragments is substantial but less than that expected for a nativelike conformation or independent folding (see Discussion below).

The CD spectra in the 250−320 nm region of nicked and intact apoMb, as well as of fragments 1−88 and 89−153, are shown in Figure 4B. The fragments do not show significant optical activity in the near-UV region; therefore, they do not seem to possess specific and rigid tertiary structure (*74*). On the other hand, both intact and nicked apoMb show quite intense CD signals in the 250−320 nm region arising from the contributions of several aromatic chromophores scattered along the 153-residue chain of the protein. The CD band at 295 nm is due to the Trp7 and Trp14 residues, implying that these residues are in a specific and rigid environment in both protein species (*70*, *71*, *74*, *75*). The near-UV CD spectrum of nicked apoMb is much similar to that of the intact protein, even if not identical (Figure 4B).

Fluorescence Emission Spectroscopy. The intrinsic fluorescence emission properties of apoMb provide a means of monitoring protein conformational features and transitions (*77*−*79*). The protein contains two Trp residues (Trp7 and Trp14) located at the N-terminal helix A. The λ_{max} of their fluorescence emission can probe protein structure, because the maximum wavelength of emission of the indole rings depends on their microenvironment within the protein fold (*78*−*80*). The λ_{max} of the Trp fluorescence emission of fragment 1−88 is centered at 352 nm, indicating a full exposure to solvent of the indole rings of Trp7 and Trp14 (not shown). Fragment 89−153 shows only Tyr emission fluorescence (λ_{max} 303 nm), since this fragment does not contain Trp but does have two Tyr residues at positions 103 and 146 of the 153-residue chain of apoMb. Conversely, both intact and nicked apoMb show a λ_{max} of emission at 334 nm in Tris buffer, pH 7.0 (not shown). Therefore, fluorescence emission measurements indicate that the Trp residues are in a seemingly identical environment in both protein species and rather buried in the protein interior (*78*).

Second-Derivative Spectroscopy. The use of second-derivative spectroscopy for unravelling features of the 3D structure of proteins is well-established, since this technique can provide a fingerprint of the aromatic side-chain topology for a given protein structure (*67*, *81*). In particular, the second-derivative spectrum of a protein containing both Tyr and Trp residues is influenced by the polarity of the microenvironment in which the aromatic chromophores are located, and it can be used to calculate the average exposure (α) of Tyr residues to solvent (*67*). The analysis of the second-derivative spectra of intact and nicked apoMb, recorded under native and denaturing conditions (see Figure 5), gave the same α value of 0.26, providing evidence that the Tyr residues in positions 103 and 146 are rather buried in both protein species.

Heme Binding. ApoMb is able to reversibly bind ferric heme under nondenaturing conditions (*63*). In this study, the binding of heme to the intact and nicked apoMb was investigated by absorption and CD spectroscopy. Absorption spectra of the complex 1−88/89−153 were recorded in the

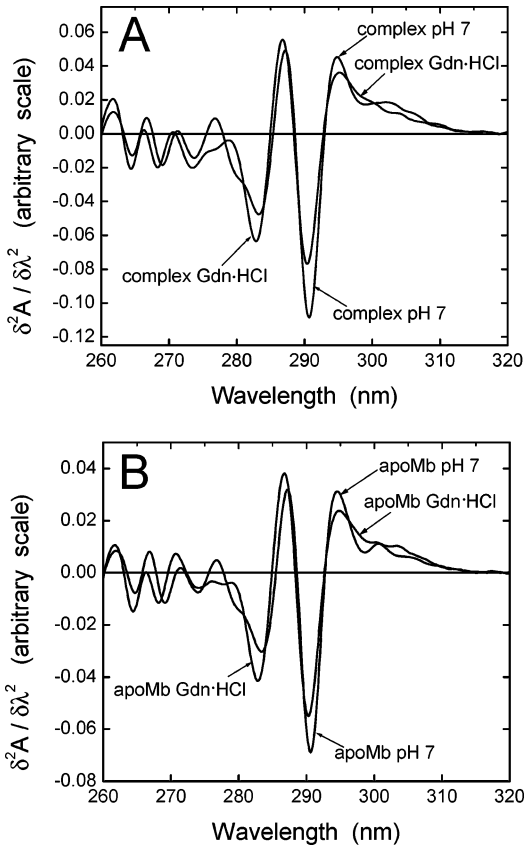

FIGURE 5: Second-derivative spectra of nicked (top) and intact (bottom) apoMb. Spectra were recorded in 20 mM Tris·HCl/0.2 M NaCl buffer, pH 7.0, at room temperature, at a protein concentration of 0.12−0.17 mg/mL. Spectra were recorded also in the presence of 6 M Gdn·HCl at pH 7.0.

presence of heme with the metal ion in the oxidized (ferric) and reduced (ferrous) form and compared with those of intact apoMb. The appearance of a sharp Soret band centered at 409 nm for both intact and nicked apoMb (Figure S1, Supporting Information) strongly indicates that the ferric-heme moiety binds to the nicked species in a nativelike manner. Of note, the absorption of ferric heme in the free form is remarkably different, with a broad band centered at 385 nm (not shown). A further proof of similarity of heme binding to nicked and intact apoMb has been obtained by reducing with a large excess of dithionite the ferric ion of heme bound to the complex 1−88/89−153. It has been found that only the ferric heme that is properly bound to apoMb in the hydrophobic heme-binding pocket can be reduced, whereas free (or incorrectly bound) heme remains in the oxidized ferric form (*63*). As a result of reduction, the Soret band at 409 nm is shifted to 434 nm. As shown in Figure S1 (Supporting Information), nicked apoMb in its oxidized or reduced state displays absorption spectra very similar to those displayed by the intact protein.

The binding of heme was also monitored by recording the variation in the CD signal of heme induced by its binding to intact or nicked apoMb. In fact, free and surface-bound heme does not display a CD signal (*63*, *82*). However, the stereospecific and rigid binding of the heme to the protein promotes a chirality and this results in a strong CD signal in the Soret region. As shown in Figure 6, in the presence of intact or nicked apoMb the CD signal of heme appears as a prominent band centered at about 410 nm, but the

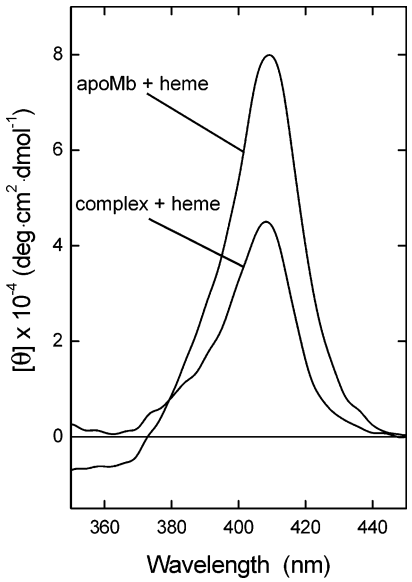

FIGURE 6: CD spectra in the Soret region of heme bound to intact or nicked apoMb (complex 1−88/89−153). To avoid the presence of free heme in the mixture, the molar ratio of protein to heme was 1.0/0.75 with both protein samples. Spectra were recorded in 20 mM Tris-HCl/0.2 M NaCl buffer, pH 7.0, at 25 °C.

intensity of the signal is about 40% reduced with nicked apoMb.

ANS Binding. ANS binds stoichiometrically to a specific site on apoMb (*83*, *84*). Since the holoprotein does not bind ANS and hemin efficiently displaces ANS from the ANS−apoMb complex, it was concluded that ANS binds to apoMb at the same site of the heme moiety (*83*). Figure 7A shows the fluorescence emission spectra in the 400−700 nm region of ANS in the presence of intact or nicked apoMb under seemingly identical experimental conditions. The same maximum of emission at 455 nm is observed with both protein species, but the intensity of emission of ANS bound to nicked apoMb is about half that observed with the intact protein.

The energy transfer from the aromatic chromophores (mostly Trp7 and Trp14) of apoMb to protein-bound ANS was next evaluated. The energy-transfer derives from the fact that in the ANS−protein complex there is a fluorescence energy transfer donor/acceptor combination due to the overlap of protein fluorescence emission, mostly from Trp7 and Trp14 (*77*, *78*, *80*). Figure 7B shows the emission spectra of the ANS−protein complex displayed by both intact and nicked apoMb, upon excitation at 280 nm under identical experimental conditions. It is seen that both complexes display ANS fluorescence in the 400−600 nm region, implying that energy transfer does occur. However, the intensity of ANS emission with nicked apoMb is about half that observed with intact apoMb, implying a less efficient energy transfer.

The ANS−apoMb complex can display dichroic activity in the ANS absorption region if the dye is bound in a stereospecific manner to the protein. In the absence of protein, the CD spectrum of ANS in the 340−430 nm region was practically coincident with the solvent baseline, whereas in the presence of either intact or nicked apoMb a negative CD band centered at about 385 nm appears (see Figure S2, Supporting Information). In the case of ANS bound to nicked

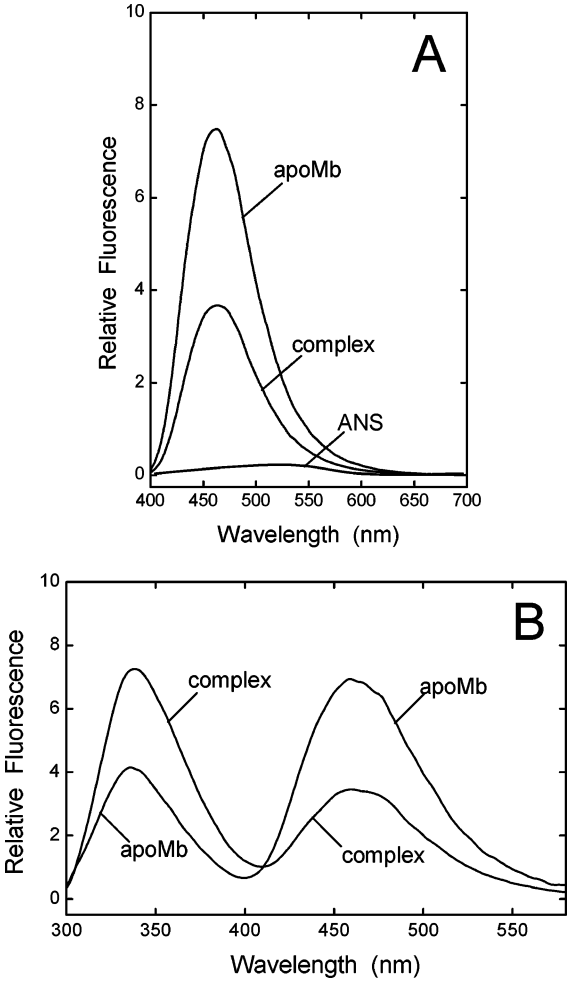

FIGURE 7: ANS binding to intact or nicked apoMb monitored by fluorescence. (A) Emission spectra of ANS mixed with intact or nicked apoMb. The excitation wavelength was 390 nm. (B) Emission spectra of intact or nicked apoMb mixed with ANS. The excitation wavelength was 280 nm. All spectra were recorded at 25 °C in 20 mM Tris-HCl/0.2 M NaCl buffer, pH 7.0. The protein concentration was 2 μM and the ANS/protein molar ratio was 0.36, to avoid the presence of free ANS in solution.

apoMb, the intensity of the CD signal at 385 nm is about 40% reduced.

We felt it would be interesting to monitor also the possible ANS binding to the individual fragments 1−88 and 89−153, considering that these fragments adopt significant secondary structure (see above). ANS has been widely used to probe partly folded or MG states of proteins (*85*). The binding of ANS to protein MGs is accompanied by a significant increase in quantum yield and a blue shift of the maximum of fluorescence emission from about 515 nm in water to 455−495 nm when bound to proteins, depending upon the hydrophobicity of the ANS binding site. Both fragments 1−88 and 89−153 are able to bind ANS, as given by the enhancement of emission intensity and maximum wavelength of emission at about 480 nm (see Figure S3, Supporting Information). These fluorescence properties of ANS are those usually considered diagnostic of protein MGs (*85*).

DISCUSSION

Complementation of protein fragments leading to nicked proteins is an excellent tool for probing mechanism(s) of protein structure, stability, and function, as well as macromolecular biorecognition phenomena in general (*20*, *30*). Fragment complementing systems have been used to scrutinize the molecular features required for fragments to associate and fold into a stable nativelike protein complex (*16*, *17*, *26*, *27*, *29*, *86*−*88*), to produce mutants of proteins by semisynthetic procedures involving the chemical or enzymatic religation of the fragments (*89*, *90*), to study autonomous folding units (*19*), and more recently, as a general method to screen for protein interactions in vivo (*91*). The novel system of apoMb-complementing fragments herein described appears to be an excellent model for further studies, in particular considering that apoMb has been used for the last three decades in numerous laboratories as a model to unravel the mechanism of protein folding (*33*, *43*). The apoMb fragments will be of value for extending these earlier studies and for dissecting the difficult problem of protein folding into the steps or stages involved in folding pathways, likely with the difference that the intramolecular process of folding of the entire polypeptide chain is intermolecular with the associating fragments (*20*, *92*). The scheme shown in Figure 8 provides an outline of the use of the protein dissection strategy in protein folding studies.

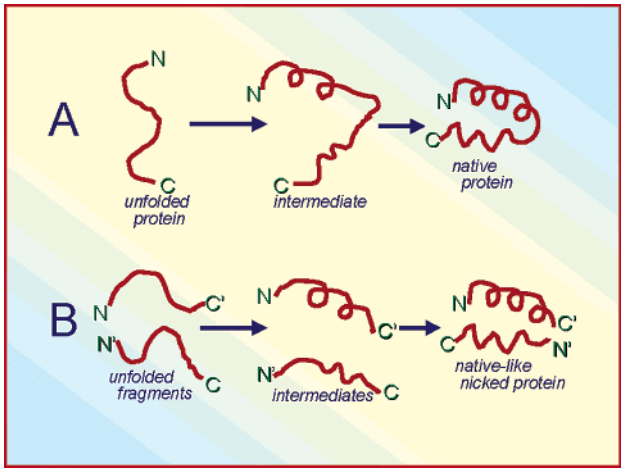

FIGURE 8: Scheme indicating the use of the protein dissection strategy in protein folding studies. (Top) The protein is shown to fold from the (fully) unfolded state (left) to the native form (right) via a folding intermediate (middle). (Bottom) Protein fragments can be prepared by enzymatic or chemical dissection of the protein, as well as by recombinant DNA methods or chemical synthesis. The fragments can attain in isolation a partly folded state and their association leads to the nicked protein. The conformation of the nicked protein species is assumed to be nativelike and quite similar to the structure of the intact native protein. N and C indicate the N- and C-terminal residues, respectively.

Conformation of Fragments 1−88 and 89−153. Spectroscopic studies by far-UV CD indicate that both fragments 1−88 and 89−153 in isolation adopt folded states in aqueous solution at neutral pH. Both fragments display minima of ellipticity near 208 and 222 nm, typical of α-helical polypeptides (*69*). On the basis and limitations of the methods used to evaluate the percent content of secondary structure of polypeptides from far-UV CD spectra (*69*, *71*, *73*, *75*), 25% and 46% helix content for fragments 1−88 and 89−153, respectively, can be estimated. The corresponding figures for an independent folding or nativelike conformation of N- and C-terminal fragment are 55% and 80%, respectively. These figures are those that can be calculated

on the basis of the crystallographic structure of the holoprotein (*31*). The CD signals of the fragments are essentially at the baseline in the near-UV (Figure 4B), and therefore, they do not appear to possess any tertiary structure. Moreover, the λ_{max} of the fluorescence emission of Trp7 and Trp14 of fragment 1−88 is centered at 352 nm (not shown), which is typical for Trp residues fully exposed to solvent, as occurs with unfolded proteins (*78*−*80*). In conclusion, the fragments appear to fold only partially. Considering that MG of proteins are defined as containing significant secondary structure but lacking tertiary structure (*2*−*4*), we may propose that fragments 1−88 and 89−153 adopt a MG state in aqueous solution at neutral pH. As a matter of fact, both fragments are able to bind ANS, a typical property of a protein MG (*85*).

It is of interest to relate the conformational features of the horse apoMb fragments 1−88 and 89−153 to those of the previously investigated synthetic peptides of the homologous sperm whale myoglobin (*52*). The synthetic peptides corresponded individually to each of the eight α-helices of the all-α protein and, with the only exception of the peptide encompassing helix H (residues 124−151), do not adopt significant secondary structure in aqueous solution. The much higher helical content of apoMb fragments 1−88 and 89−153 here observed derives from the fact that they do not correspond only to single helices but to domains of the protein given by the assembly of secondary structure elements (helices).

For example, fragment 89−153 encompasses in full the G−H helical hairpin region of myoglobin (see Figure 1). This hairpin was earlier proposed to be part of an initial folding intermediate of apoMb. However, early attempts to identify significant secondary structure in the isolated G−H helical hairpin by use of a synthetic 51-residue peptide fragment of sperm whale myoglobin failed (*49*−*51*). The 51-residue model peptide, corresponding to chain segment 100−150 of apoMb, contained four amino acid replacements with respect to the natural sequence of the protein (*51*) and it was shorter than the 65-residue fragment 89−153 herewith investigated. It could be well that in the 51-residue model some critical structure stabilizing interactions are missing with respect to those present in the natural and longer apoMb fragment 89−153, which instead displays a high content of helicity (see Figure 4A). Finally, it seems to us difficult to compare the conformational features of the N-terminal fragment 1−88 herewith investigated (see Figure 4) with those reported for the apoMb fragments 1−36, 1−77, and 1−119 (*53*). All these fragments, prepared by recombinant methods, were found to be heavily aggregated, and in particular, some of them displayed far-UV CD spectra characteristic of β-sheet secondary structure and/or aggregated polypeptides. However, these measurements were made on peptide samples recovered after lyophilization and it could well be that this treatment caused peptide aggregation phenomena. In this study, care was taken not to dry up the fragments by lyophilization, and in general, spectroscopic measurements were made on sample solutions obtained after gel filtration.

It is worth noting that surface area scans performed onto the 3D structure of native myoglobin allowed Wodak and Janin (*93*) to dissect myoglobin into two structural domains and the boundary between them was at residue 85 of the 153-residue chain of the protein. Therefore, it is gratifying to observe that limited proteolysis of apoMb occurs at the boundary between the two domains and that fragments 1−88 and 89−153, therefore, correspond almost exactly to the structural domains identified by computational dissection of the protein. It has been proposed earlier that fragments corresponding to protein domains could fold independently from the rest of the polypeptide chain, and experimental evidence for this proposal was obtained from conformational studies conducted on several protein fragments (*14*, *19*). Here, with the relatively large apoMb fragments 1−88 (88 residues) and 89−153 (65 residues), we provide evidence that they attain in isolation significant secondary structure but that they lack the specific interactions of the tertiary structure of native apoMb. Therefore, the "independent" folding of the apoMb fragments 1−88 and 89−153 reaches only that of a partially folded or MG state.

Conformational Features of Nicked Apomyoglobin. The complex of apoMb fragments 1−88 and 89−153 (nicked apoMb) has structural features quite similar to those of the intact protein, despite the Pro88−Leu89 peptide bond fission. The similarity of structures between nicked and intact apoMb has been deduced from a variety of spectroscopic measurements, including far- and near-UV CD, fluorescence emission, and second-derivative spectroscopy. The CD spectrum of the complex in the near-UV region displays positive CD signals in the 250−320 nm region, in analogy to intact apoMb. The features of the near-UV dichroic signals are associated with the microenvironment of the aromatic chromophores of a protein (*72*, *74*, *75*). It is known that proteins can display a strong variation in their near-UV CD spectra, due to the fact that the intensities and maxima of CD bonds of the aromatic chromophores are strongly influenced by the protein tertiary structure and derive from their positive and negative contributions, sometimes even canceling each other (*74*). Therefore, considering that the near-UV CD spectrum of a protein is a fingerprint of its tertiary structure, we may conclude that important and specific structural details of the 3D structure are shared between nicked and intact apoMb. However, some differences between the two protein species have been evaluated from heme and ANS binding data. We should not be surprised by the fact that the heme and ANS, while binding to both protein species, display some differences in the dichroic signal of heme in the Soret region (Figure 6) and in the intensity of ANS emission (Figure 7A), as well as energy transfer properties (Figure 7B). In fact, this is quite reasonable if one considers that intact apoMb binds one ANS molecule at the same hydrophobic site which binds the heme moiety in the holoprotein and that this site includes the chain segment encompassing helix F (residues 82−97). Cutting the polypeptide chain at the Pro88−Leu89 peptide bond, exactly at the middle of this helical segment, perturbs but does not abolish the heme and ANS binding site of the protein.

The similarity of the structure of nicked apoMb with that of the intact protein implies that the introduction of a chain fission at the level of helix F (residues 82−97) and the concomitant exchange of a neutral peptide bond for a negative carboxylate and positive ammonium group does not influence the overall structure of the protein. Taniuchi et al. (*20*) developed the concept of permissible sites of cleavage

of a protein for producing systems of complementing protein fragments and concluded that cleavages can occur at exposed, flexible loops and not at the level of chain segments in a regular secondary structure such as α-helices. In the present case of apoMb, we may observe that the cleaved peptide bond Pro88−Leu89 is located at the level of helix F. However, we assume here that helix F is unfolded (see Figure 1), as deduced from limited proteolysis experiments (*36*, *37*), NMR measurements (*32*−*35*), and molecular dynamics simulations (*58*−*61*, *76*) of apoMb. Therefore, the 88−89 peptide bond occurs at a disordered region and, according to Taniuchi et al. (*20*), is a permissible site of cleavage for developing a fragment-complementing system of apoMb.

Concluding Remarks. Here we have shown by far- and near-UV CD measurements that the apoMb fragments 1−88 and 89−153 adopt at neutral pH significant helical secondary structure but that they essentially lack tertiary structure. These conformational features are those usually ascribed to the partly folded or MG states that globular proteins adopt under mildly denaturing conditions, such as low pH (*2*−*4*). Moreover, the apoMb fragments bind the hydrophobic fluorescent dye ANS, which is usually taken as a diagnostic feature of the MG state (*85*). According to the scheme shown in Figure 8, the protein dissection strategy here applied to apoMb allowed us to separate into two parts the MG of apoMb of the intact protein. In previous studies, it has been proposed that the MG of apoMb comprises helices A, G, and H (*41*, *94*) or A, B, G, and H (*33*). Therefore, if we accept that this MG is a folding intermediate for intact apoMb (see also *44*), we may propose that the isolated fragments 1−88 and 89−153 adopt, under physiological conditions at neutral pH, a MG state and that only upon their complementation is their overall structure fine-tuned and rigidified as it occurs in intact apoMb. In the future, it will be of interest to use the apoMb fragments and their complex for conducting detailed studies by multidimensional NMR and hydrogen-exchange measurements of the type already conducted on the full-length protein (*33*). These studies will complement and extend the results herewith reported and shed light onto the initial, intermediate, and final stages of the overall folding pathway of the apoMb molecule. A specific advantage of the apoMb fragment-complementing system is that measurements can be made under physiological conditions, without the need to expose the protein to denaturing conditions in order to populate partly folded intermediates.

ACKNOWLEDGMENT

We thank Dr. Patrizia Polverino de Laureto for insightful comments on this paper and Dr. Daniele Dalzoppo for assistance with mass spectrometry measurements. The expert typing of the manuscript by Barbara Sicoli is also gratefully acknowledged.

SUPPORTING INFORMATION AVAILABLE

Three figures showing absorption spectra of the oxidized and reduced heme iron in the apoMb complex 1−88/89−153 as compared to that of intact apoM, CD spectra of ANS bound to nicked and intact apoMb, and ANS binding to apoMb fragments 1−88 and 89−153 monitored by fluorescence emission measurements. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Baldwin, R. L. (1975) Intermediates in protein folding reactions and the mechanism of protein folding, *Annu. Rev. Biochem. 44*, 453−475.
2. Ptitsyn, O. B. (1987) Protein folding: Hypothesis and experiments, *J. Protein Chem. 6*, 273−293.
3. Ptitsyn, O. B. (1995) Molten globule and protein folding, *Adv. Protein Chem. 47*, 83−229.
4. Arai, M., and Kuwajima, K. (2000) Role of the molten globule state in protein folding, *Adv. Protein Chem. 53*, 209−282.
5. Chamberlain, A. K., and Marqusee, S. (2000) Comparison of equilibrium and kinetic approaches for determining protein folding mechanisms, *Adv. Protein Chem. 53*, 283−328.
6. Onuchic, J. N., Wolynes, P. G., Luthey-Schulten, Z., and Socci, N. D. (1995) Toward an outline of the topography of a realistic protein-folding funnel, *Proc. Natl. Acad. Sci. U.S.A. 92*, 3626−3630.
7. Onuchic, J. N., Nymeyer, H., Garcia, A., Chacine, J., and Socci, N. D. (2000) The energy landscape theory of protein folding: Insights into folding mechanisms and scenarios, *Adv. Protein Chem. 53*, 87−152.
8. Bryngelson, J. D., Onuchic, J. N., Socci, N. D., and Wolynes, P. G. (1995) Funnels, pathways and the energy landscape of protein folding: A synthesis, *Proteins: Struct., Funct., Genet. 21*, 167−195.
9. Dill, K. A., and Chan, H. S. (1997) From Levinthal to pathways to funnels, *Nat. Struct. Biol. 4*, 10−19.
10. Karplus, M. (1997) The Levinthal paradox: Yesterday and today, *Folding Des. 2*, S69−S75.
11. Anfinsen, C. B., and Scheraga, H. A. (1975) Experimental and theoretical aspects of protein folding, *Adv. Protein Chem. 29*, 205−299.
12. Wetlaufer, D. B. (1981) Folding of protein fragments, *Adv. Protein Chem. 34*, 61−92.
13. Dalzoppo, D., Vita, C., and Fontana, A. (1985) Folding of thermolysin fragments: Identification of the minimum size of a carboxyl-terminal fragment that can fold into a stable nativelike structure, *J. Mol. Biol. 182*, 331−340.
14. Wright, P. E., Dyson, H. J., and Lerner, R. A. (1988) Conformation of peptide fragments of proteins in aqueous solution: Implications for initiation of protein folding, *Biochemistry 27*, 7167−7175.
15. Dyson, H. J., Merutka, G., Waltho, J. P., Lerner, R. A., and Wright, P. E. (1992) Folding of peptide fragments comprising the complete sequence of proteins: Models for initiation of protein folding. I. Myohemerythrin, *J. Mol. Biol. 226*, 795−817.
16. Gegg, C. V., Bowers, K. E., and Matthews, C. R. (1997) Probing minimal independent folding units in dihydrofolate reductase by molecular dissection, *Protein Sci. 6*, 1885−1892.
17. Llinás, M., and Marqusee, S. (1998) Subdomain interactions as a determinant in the folding and stability of T4 lysozyme, *Protein Sci. 7*, 96−104.
18. Wu, L. C., Grandori, R., and Carey, J. (1994) Autonomous subdomains in protein folding, *Protein Sci. 3*, 369−371.
19. Peng, Z. Y., and Wu, L. C. (2000) Autonomous protein folding units, *Adv. Protein Chem. 53*, 1−47.
20. Taniuchi, H., Parr, G. R., and Juillerat, M. A. (1986) Complementation in folding and fragment exchange, *Methods Enzymol. 131*, 185−217.
21. Richards, F. M., and Vithayathil, P. J. (1959) The preparation of subtilisin-modified ribonuclease and the separation of the peptide and protein components, *J. Biol. Chem. 234*, 1459−1464.
22. Anfinsen, C. B. (1973) Principles that govern the folding of protein chains, *Science 181*, 223−227.
23. Taniuchi, H. (1970) Formation of randomly paired disulfide bonds in des-(121−124)-ribonuclease after reduction and reoxidation, *J. Biol. Chem. 245*, 5459−5468.
24. Fischer, A., and Taniuchi, H. (1992) A study of core domains and the core domain−domain interaction of cytochrome *c* fragment complex, *Arch. Biochem. Biophys. 96*, 1−16.
25. Parr, G., Hantgan, R. R., and Taniuchi, H. (1978) Formation of two alternative complementing structures from cytochrome *c* heme fragment (residue 1 to 38) and the apoprotein, *J. Biol. Chem. 253*, 5381−5388.
26. Ojennus, D. D., Fleissner, M. R., and Wuttke, D. S. (2001) Reconstitution of a nativelike SH2 domain from disordered peptide fragments examined by multidimensional heteronuclear NMR, *Protein Sci. 10*, 2162−2175.

27. Ojennus, D. D., Lehto, S. E., and Wuttke, D. S. (2003) Electrostatic interactions in the reconstitution of an SH2 domain from constituent peptide fragments, *Protein Sci. 12*, 44−55.
28. de Prat-Gay, G., and Fersht, A. R. (1994) Generation of a family of protein fragments for structure−folding studies. 1. Folding complementation of two fragments of chymotrypsin inhibitor-2 formed by cleavage at its unique methionine residue, *Biochemistry 33*, 7957−7963.
29. Spolaore, B., Bermejo, R., Zambonin, M., and Fontana, A. (2001) Protein interactions leading to conformational changes monitored by limited proteolysis: Apo form and fragments of horse cytochrome *c*, *Biochemistry 40*, 9460−9468.
30. de Prat-Gay, G. (1996) Association of complementary fragments and the elucidation of protein folding pathways, *Protein Eng. 9*, 843−847.
31. 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.
32. Eliezer, D., and Wright, P. E. (1996) Is apomyoglobin a molten globule? Structural characterization by NMR, *J. Mol. Biol. 263*, 531−538.
33. 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.
34. Lecomte, J. T. J., Kao, Y. H., and Cocco, M. J. (1996) The native state of apomyoglobin described by proton NMR spectroscopy: The A−B−G−H interface of wild-type sperm whale apomyoglobin, *Proteins: Struct., Funct., Genet. 25*, 267−285.
35. Lecomte, J. T. J., Sukits, S. F., Bhattacharya, S., and Falzone, C. J. (1999) Conformational properties of native sperm whale apomyoglobin in solution, *Protein Sci. 8*, 1484−1491.
36. Fontana, A., Polverino de Laureto, P., De Filippis, V., Scaramella, E., and Zambonin, M. (1997) Probing the partly folded states of proteins by limited proteolysis, *Folding Des. 2*, R17−R26.
37. 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.
38. Barrick, D., and Baldwin, R. L. (1993) The molten globule intermediate of apomyoglobin and the process of protein folding, *Protein Sci. 2*, 869−876.
39. Fink, A. L., Oberg, K. A., and Seshadry, S. (1997) Discrete intermediates versus molten globule models for protein folding: Characterization of partially folded intermediates of apomyoglobin, *Folding Des. 3*, 19−25.
40. 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.
41. Hughson, F. M., Wright, P. E., and Baldwin, R. L. (1990) Structural characterization of a partly folded apomyoglobin intermediate, *Science 249*, 1544−1548.
42. Gilmanshin, R., Gulotta, M., Dyer, R. B., and Callender, R. H. (2001) Structures of apomyoglobin's various acid-destabilized forms, *Biochemistry 40*, 5127−5136.
43. Tcherkasskaya, O., and Uversky, V. N. (2001) Denatured collapsed states in protein folding: Example of apomyoglobin, *Proteins: Struct., Funct., Genet. 44*, 244−254.
44. Nishimura, C., Dyson, H. J., and Wright, P. E. (2002) The apomyoglobin folding pathway revisited: Structural heterogeneity in the kinetic burst phase intermediate, *J. Mol. Biol. 322*, 483−489.
45. Epand, R. M., and Scheraga, H. A. (1968) The influence of long-range interactions on the structure of myoglobin, *Biochemistry 7*, 2864−2872.
46. Hermans, J., Jr., and Puett, D. (1971) Relative effects of primary and tertiary structure on helix formation in myoglobin and alpha-lactalbumin, *Biopolymers 10*, 895−914.
47. Nakano, M., Iwamaru, H., and Tobita, T. (1982) Influence of hemin on the conformation of cyanogen bromide-cleaved peptides of apomyoglobin, *Biopolymers 21*, 805−815.
48. Waltho, J. P., Feher, V. A., Lerner, R. A., and Wright, P. E. (1989) Conformation of a T cell stimulating peptide in aqueous solution, *FEBS Lett. 250*, 400−404.
49. Shin, H. C., Merutka, G., Waltho, J. P., Tennant, L. L., Dyson, H. J., and Wright, P. E. (1993) Peptide models of protein folding initiation sites. 3. The G−H helical hairpin of myoglobin, *Biochemistry 32*, 6356−6364.
50. Shin, H. C., Merutka, G., Waltho, J. P., Wright, P. E., and Dyson, H. J. (1993) Peptide models of protein folding initiation sites. 2. The G−H turn region of myoglobin acts as a helix stop signal, *Biochemistry 32*, 6348−6355.
51. Waltho, J. P., Feher, V. A., Merutka, G., Dyson, H. J., and Wright, P. E. (1993) Peptide models of protein folding initiation sites. 1. Secondary structure formation by peptides corresponding to the G- and H-helices of myoglobin, *Biochemistry 32*, 6337−6347.
52. Reymond, M. T., Merutka, G., Dyson, H. J., and Wright, P. E. (1997) Folding propensities of peptide fragments of myoglobin, *Protein Sci. 6*, 706−716.
53. Chow, C. C., Chow, C., Raghunathan, V., Huppert, T. J., Kimball, E. B., and Cavagnero, S. (2003) Chain length dependence of apomyoglobin folding: Structural evolution from misfolded sheets to native helices, *Biochemistry 42*, 7090−7099.
54. Craik, C. S., Buchman, S. R., and Beychok, S. (1980) Characterization of globin domains: Heme binding to the central exon product, *Proc. Natl. Acad. Sci. U.S.A. 77*, 1384−1388.
55. De Sanctis, G., Ascoli, F., and Brunori, M. (1994) Folding of apominimyoglobin, *Proc. Natl. Acad. Sci. U.S.A. 91*, 11507−11511.
56. Fontana, A., Fassina, G., Vita, C., Dalzoppo, D., Zamai, M., and Zambonin, M. (1986) Correlation between sites of limited proteolysis and segmental mobility in thermolysin, *Biochemistry 25*, 1847−1851.
57. Fontana, A., Polverino de Laureto, P., De Filippis, V., Scaramella, E., and Zambonin, M. (1999) Limited proteolysis in the study of protein conformation, in *Proteolytic Enzymes: Tools and Targets* (Sterchi, E. E., and Stocker, W., Eds.) pp 257−284, Springer-Verlag, Heidelberg, Germany.
58. Hirst, J. D., and Brooks, C. L. (1995) Molecular dynamics simulations of isolated helices of myoglobin, *Biochemistry 34*, 7614−7621.
59. Brooks, C. L. (1992) Characterization of "native" apomyoglobin by molecular dynamics simulation, *J. Mol. Biol. 227*, 375−380.
60. Tirado-Rives, J., and Jorgensen, W. L. (1993) Molecular dynamics simulations of the unfolding of apomyoglobin in water, *Biochemistry 32*, 4175−4184.
61. Onufriev, A., Case, D. A., and Bashford, D. (2003) Structural details, pathways and energetics of unfolding apomyoglobin, *J. Mol. Biol. 325*, 555−567.
62. Teale, F. W. J. (1959) Cleavage of the haem-protein link by acid methylethyl ketone, *Biochim. Biophys. Acta 35*, 543.
63. Rothgeb, T. M., and Gurd, F. R. (1978) Physical methods for the study of myoglobin, *Methods Enzymol. 52*, 473−486.
64. Schägger, H., and von Jagow, G. (1987) Tricine−sodium dodecyl sulfate−polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa, *Anal. Biochem. 166*, 368−379.
65. Scholtz, J. M., Quian, H., York, E. J., Stewart, J. M., and Baldwin, R. L. (1991) Parameters of helix−coil transition theory for alanine-based peptides of varying chain lengths in water, *Biopolymers 31*, 1463−1470.
66. Gill, S. C., and von Hippel, P. H. (1989) Calculation of protein extinction coefficients from amino acid sequence data, *Anal. Biochem. 182*, 319−326.
67. Ragone, R., Colonna, G., Belestrieri, C., Servillo, L., and Irace, G. (1984) Determination of tyrosine exposure in proteins by second-derivative spectroscopy, *Biochemistry 23*, 1871−1875.
68. Dawson, R. M. C., Elliott, D. C., Elliott, W. H., and Jones, K. M. (1975) *Data for Biochemical Research*, pp 230−231, Oxford University Press, Oxford, U.K.
69. Chen, Y. H., Yang, J. T., and Chau, K. H. (1974) Determination of the helix and beta form of proteins in aqueous solution by circular dichroism, *Biochemistry 13*, 3350−3359.
70. Yang, J. T., Wu, C.-S. C., and Martinez, H. M. (1986) Calculation of protein conformation from circular dichroism, *Methods Enzymol. 130*, 208−269.
71. Johnson, W. C., Jr. (1990) Protein secondary structure and circular dichroism: A practical guide, *Proteins: Struct., Funct., Genet. 7*, 205−214.
72. Kelly, S. M., and Price, N. C. (2000) The application of circular dichroism to studies of protein folding and unfolding, *Biochim. Biophys. Acta 1338*, 161−185.
73. Venyaminov, S. Y., and Yang, J. T. (1996) Determination of protein secondary structure, in *Circular Dichroism and the Conformational Analysis of Biomolecules* (Fasman, G. D., Ed.), pp 69−107, Plenum Press, New York.

74. Strickland, E. H. (1974) Aromatic contributions to circular dichroism spectra of proteins, *CRC Crit. Rev. Biochem. 5*, 113−175.
75. Woody, R. W. (1995) Circular dichroism, *Methods Enzymol. 246*, 34−71.
76. Hirst, J. D., and Brooks, C. L. (1994) Helicity, circular dichroism and molecular dynamics of proteins, *J. Mol. Biol. 243*, 173−178.
77. Kirby, E. P., and Steiner, R. F. (1970) The tryptophan microenvironments in apomyoglobin, *J. Biol. Chem. 245*, 6300−6306.
78. Burstein, E. A., Vedenkina, N. S., and Ivkova, M. N. (1973) Fluorescence and the location of tryptophan residues in protein molecules, *Photochem. Photobiol. 18*, 263−279.
79. Lehrer, S. S., and Leavis, P. C. (1978) Solute quenching of protein fluorescence, *Methods Enzymol. 49*, 222−236.
80. Lakowicz, J. R. (1986) *Principles of Fluorescence Spectroscopy*, Plenum Press, New York.
81. De Filippis, V., Polverino de Laureto, P., Toniutti, N., and Fontana, A. (1996) Acid-induced molten globule state of a fully active mutant of human interleukin-6, *Biochemistry 35*, 11503−11511.
82. Ihara, M., Takahashi, S., Ishimori, K., and Morishima, I. (2000) Functions of fluctuation in the heme-binding loops of cytochrome b_5 revealed in the process of heme incorporation, *Biochemistry 39*, 5961−5970.
83. Stryer, L. (1965) The interaction of a naphthalene dye with apomyoglobin and apohemoglobin: A fluorescent probe of nonpolar binding sites, *J. Mol. Biol. 13*, 482−495.
84. Bismuto, E., Sirangelo, I., and Irace, G. (1992) Salt-induced refolding of myoglobin at acidic pH: Molecular properties of a partly folded intermediate, *Arch. Biochem. Biophys. 298*, 624−629.
85. Semisotnov, G. V., Rodionova, N. A., Razgulyaev, O. I., Uversky, V. N., Gripas, A. F., and Gilmanshin, R. I. (1991) Study of the "molten globule" intermediate state in protein folding by a hydrophobic fluorescent probe, *Biopolymers 31*, 119−128.
86. Chaffotte, A. F., Li, J. H., Georgescu, R. E., Goldberg, M. E., and Tasayco, M. L. (1997) Recognition between disordered states: Kinetics of the self-assembly of thioredoxin fragments, *Biochemistry 36*, 16040−16048.
87. Tasayco, M. L., Fuchs, J., Yang, X. M., Dyalram, D., and Georgescu, R. E. (2000) Interaction between two discontiguous chain segments from the beta-sheet of *Escherichia coli* thioredoxin suggests an initiation site for folding, *Biochemistry 39*, 10613−10618.
88. Berggard, T., Julenius, K., Ogard, A., Drakenberg, T., and Linse, S. (2001) Fragment complementation studies of protein stabilization by hydrophobic core residues, *Biochemistry 40*, 1257−1264.
89. Homandberg, G. A., and Laskowski, M., Jr. (1979) Enzymatic resynthesis of the hydrolyzed peptide bond(s) in ribonuclease S, *Biochemistry 18*, 586−592.
90. Wallace, C. J. A. (1993) Understanding cytochrome *c* function: Engineering protein structure by semisynthesis, *FASEB J. 7*, 505−515.
91. Remy, I., and Michnick, S. W. (1999) Clonal selection and in vivo quantitation of protein interactions with protein-fragment complementation assays, *Proc. Natl. Acad. Sci. U.S.A. 96*, 5394−5399.
92. Tsai, C. J., Xu, D., and Nussinov, A. (1998) Protein folding via binding and *vice versa*, *Folding Des. 3*, R71−R80.
93. Wodak, S. J., and Janin, J. (1981) Location of structural domains in proteins, *Biochemistry 20*, 6544−6552.
94. 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.

BI0400103