Probing the partly folded states of proteins by limited proteolysis

Angelo Fontana, Patrizia Polverino de Laureto, Vincenzo De Filippis, Elena Scaramella and Marcello Zambonin

The folding of a polypeptide chain of a relatively large globular protein into its unique three-dimensional and functionally active structure occurs via folding intermediates. These partly folded states of proteins are difficult to characterize, because they are usually short lived or exist as a distribution of possible conformers. A variety of experimental techniques and approaches have been utilized in recent years in numerous laboratories for characterizing folding intermediates that occur at equilibrium, including spectroscopic techniques, solution X-ray scattering, calorimetry and gel filtration chromatography, as well as genetic methods and theoretical calculations. In this review, we focus on the use of proteolytic enzymes as probes of the structure and dynamics of folding intermediates and we show that this simple biochemical technique can provide useful information, complementing that obtained by other commonly used techniques and approaches. The key result of the proteolysis experiments is that partly folded states (molten globules) of proteins can be sufficiently rigid to prevent extensive proteolysis and appear to maintain significant native-like structure.

Address: CRIBI Biotechnology Centre, University of Padua, Via Trieste 75, 35121 Padua, Italy.

Correspondence: Angelo Fontana
e-mail: fontana@civ.bio.unipd.it

Electronic identifier: 1359-0278-002-R0017

Folding & Design 01 Apr 1997, 2:R17–R26

 ISSN 1359-0278

Introduction

The protein folding problem is central to modern structural biology and is attracting the interest of many investigators using a variety of techniques and approaches [1,2]. The folding of a protein appears to proceed via folding intermediates, which direct the folding through a limited number of pathways, consequently leading to a final folded state within a reasonably short time [3–5]. A detailed knowledge of the structure, stability and dynamics of folding intermediates of proteins would facilitate the formulation of molecular mechanisms of protein folding. In the past, the existence of partly folded states of proteins was considered quite exceptional [6], since these states were expected to be short lived or to exist as a distribution of conformers. Kinetic experiments on the refolding process of proteins were conducted using stopped-flow methods combined with a variety of spectroscopic techniques, such as nuclear magnetic resonance (NMR), circular dichroism (CD) and fluorescence. These strategies have provided indications of the nature of the early events during folding and showed that secondary structure is formed first, followed by the formation of the specific tertiary interactions of the native state [7–11]. However, transient folding intermediates are difficult to identify and characterize. Nowadays, however, we observe a flow of experimental studies on partly folded states that can be generated by exposing proteins to specific and mildly denaturing solvent conditions and that occur at equilibrium, thus allowing the analysis of their structural features [12,13]. The characteristics of the equilibrium folding intermediates are related to those of the transient intermediates detected by kinetic measurements [8].

The most popular equilibrium intermediate is the molten globule state [14–16], which has been described for numerous proteins and can be obtained by dissolving proteins in the presence of moderate concentrations of denaturant, in acid and in the presence of salts or by removing protein-bound metal ions or cofactors [4,5]. According to the original definition, the key characteristics of a molten globule include a well-defined and native-like secondary structure, lack of specific tertiary interactions and a more expanded and flexible structure with respect to that of the native protein [17,18]. The possible biological significance of the molten globule state of proteins has been proposed recently [15]. Unfortunately, the term ‘molten globule’ has been used to describe a wide range of partly folded or nonnative states of proteins not necessarily consistent with the original definition, thus giving rise to conflicting interpretations of experimental results [19] and to a lively debate on the significance of the term [5,20,21].

There is growing evidence that protein folding intermediates display significant structural variability, ranging from highly unfolded to near native-like structures [5]. Recently, structural studies have been conducted on partly folded states of proteins dissolved in mixed organic/aqueous solvents, such as aqueous trifluoroethanol (TFE) [22–25]. In aqueous TFE, proteins acquire a state characterized by a high content of secondary structure (helix) in the absence of persistent tertiary structure. In a number of studies, it has been shown that even in high concentrations of denaturing agents, such as urea or guanidine hydrochloride, proteins can maintain some residual structure and clear-cut evidence has been provided that at

least some local clusters of hydrophobic residues persist even under these extreme conditions [26–31].

The structural analysis of partly folded states of proteins (or molten globules) is expected to be difficult, since these states are flexible and are an ensemble of conformations not amenable to structural elucidation by either X-ray crystallography or NMR. Nevertheless, recent developments in hydrogen/deuterium exchange, combined with two-dimensional NMR spectroscopy, have provided very useful experimental tools for analyzing intermediates even at the level of atomic resolution [7,32–35]. Advances in NMR measurements include the use of multi-dimensional hetero-nuclear NMR techniques utilizing ^{13}C- or ^{15}N-labeled proteins [36,37]. Also, the solution X-ray scattering technique has been used to obtain accurate data on the size, shape and, in some cases, even tertiary fold of compact nonnative states of proteins [38–40].

In this review, we will focus on the use of proteolytic enzymes as probes of the structure and dynamics of protein intermediates [41–43]. This approach relies on the fact that proteolysis of a protein substrate can occur only if the polypeptide chain can bind and adapt to the specific stereochemistry of the protease active site [44–48]. This is difficult to achieve with native rigid protein structures, whereas unfolded proteins are degraded much faster. Nevertheless, even native folded proteins can be attacked by a protease. In this case, usually limited proteolysis occurs at flexible sites of the protein chain, leading often to a nicked protein species constituting two (or very few) fragments associated in a stable complex [49–51] or, with multi-domain proteins, to the dissection of the protein into the individual domains [42,52,53]. Here, we show that proteolytic probes can provide useful and novel information on the structural characteristics of partly folded states of proteins, nicely complementing and extending information reached by the use of other techniques.

The limited proteolysis approach

The limited proteolysis approach for probing protein conformation implies that the proteolytic event should be dictated by the stereochemistry and flexibility of the protein substrate and not by the specificity of the attacking protease. For this aim, the most suitable proteases are those displaying broad substrate specificity, such as subtilisin, thermolysin, proteinase K and pepsin. These endopeptidases display a moderate preference for hydrolysis at hydrophobic or neutral amino acid residues, but cleavages often occur at other residues as well [54]. The recommended approach is to perform trial experiments of proteolysis on the protein of interest in order to find out the most useful protease, the optimal substrate : protease ratio and the effect of temperature and time of incubation. Possible ways to control proteolysis are the use of a low concentration of protease, short reaction times and low temperature. It is not easy to predict in advance the most useful experimental conditions for conducting a limited proteolysis experiment, since these depend upon the structure, dynamics and stability/rigidity properties of the protein substrate and also upon the actual aim of the experiment, i.e. isolation of the rigid core of the protein or location of the sites of protein flexibility [41,43,55].

What are the structural and dynamic properties of the (specific) peptide bond(s) cleaved in a globular protein by limited proteolysis? It is plausible that limited proteolysis derives from the fact that a specific chain segment of the compact folded protein substrate is sufficiently exposed to bind at the active site of the protease. However, the notion of exposure/protrusion/accessibility [56,57] is a required property, but clearly not sufficient to explain the limited proteolysis phenomenon, since it is evident that even in a small globular protein there are many exposed sites that could be targets of proteolysis. Enhanced chain flexibility (segmental mobility) appears to be the key feature of the site of limited proteolysis, as shown by the results of systematic proteolysis and autolysis experiments conducted on the thermophilic protease thermolysin [49]. In this study, it has been demonstrated that there is a clear-cut correlation between sites of limited proteolytic attack and sites of enhanced chain flexibility, this last reflected by the crystallographic temperature factors (*B*-values) [58–61]. The results obtained with thermolysin are in line with those derived from limited proteolysis experiments conducted on a variety of other proteins of known three-dimensional structure [48]. In many cases, limited proteolysis was observed to occur at sites in the polypeptide chain displaying high segmental mobility [59,60] or poorly resolved in the electron density map, implying significant static/dynamic disorder. The proposal was advanced, therefore, that limited proteolysis of a globular protein occurs at flexible loops and, in particular, that chain segments in a regular secondary structure (such as helices) are not sites of limited proteolysis [46–49].

In a more recent study, Hubbard *et al.* [62] conducted modelling studies of the conformational changes required for proteolytic cleavages and concluded that the sites of limited proteolysis require a large conformational change (local unfolding) of a stretch of up to 12 residues. A possible explanation of the fact that helices and, in general, elements of regular secondary structure are not easily hydrolyzed by proteolytic enzymes can also be given on the basis of energetic considerations. If proteolysis is occurring at the centre of the helical segment, the helix is probably fully destroyed by end-effects and consequently all hydrogen bonds, which cooperatively stabilize it, are broken. On the other hand, a peptide bond fission at a disordered flexible site probably does change much of the energetics of that site, since the peptide hydrolysis can easily be compensated by some hydrogen bonds with

water [63]. It can be proposed, therefore, that proteolysis of rigid elements of secondary structure is thermodynamically very disadvantageous.

The concept of flexibility/loose structure in dictating the limited proteolysis event can be used to explain the fact that often, in relatively large proteins, cleavage occurs at chain segment(s) connecting the structural domains [42,52]. These 'hinge' regions between domains are usually more exposed and flexible than the rest of the polypeptide chain forming the globular units of the domains. Indeed, limited proteolysis is the classic procedure for producing, from large proteins, individual autonomously folding domains for further structural and functional characterization [52]. Limited proteolysis has also been used to define the boundaries of a domain by removing its flexible and unstructured parts [64,65]. Finally, the chain flexibility notion explains the fact that, in certain proteins, the rather loose parts can be removed by limited proteolysis, allowing the isolation of the 'core' of the protein as the most proteolysis-resistant moiety and thus as a stable protein entity [66].

α-Lactalbumin

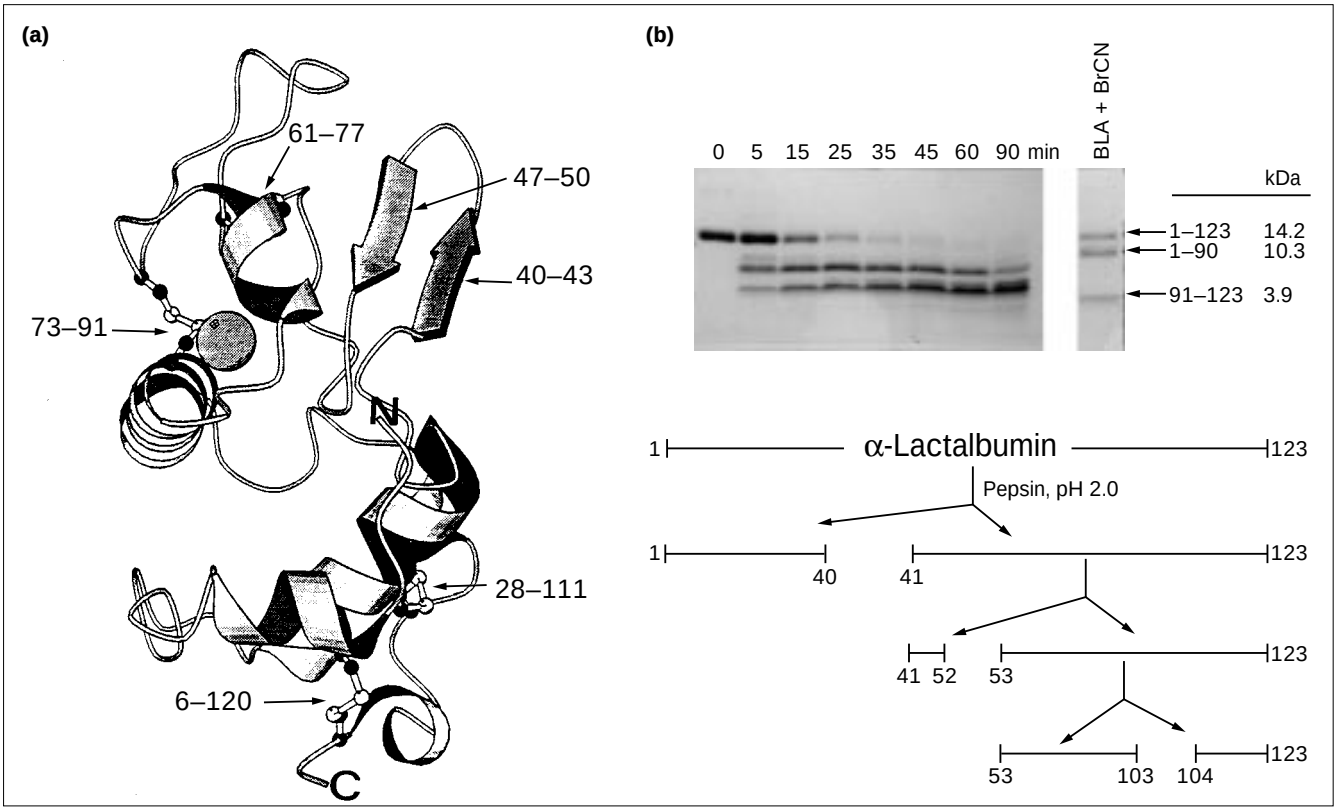
The most extensively characterized molten globule state is the one obtained by acid-induced unfolding of α-lactalbumin at pH 2.0 (A-state). These studies took advantage of the detailed knowledge of the three-dimensional structure of the baboon [67] and human [68] protein species at 1.7 Å resolution (Fig. 1). A great number of different techniques and approaches have been used over the years to characterize the A-state, which can be considered a prototype molten globule ([18] and references therein). NMR spectroscopy was used to characterize the A-state at the atomic level, providing evidence that this state is a dynamic conformational state possessing stable regions of localized and native-like helical secondary structure, but largely disordered tertiary structure [9,69–71]. In particular, evidence has been provided that the chain region encompassing the β-sheet is missing the structural constraints of the native protein and displays enhanced chain flexibility (disorder), as given by hydrogen/deuterium exchange data [9].

The results of limited proteolysis experiments conducted on bovine α-lactalbumin dissolved in acid solution are summarized in Figure 1 [72]. For this study, pepsin was used as the proteolytic probe, since this protease is optimally active in acid and shows broad substrate specificity [54]. The timecourse analysis by SDS-PAGE [73] of the peptic digestion of α-lactalbumin reveals that the protein can be digested at very few peptide bonds, because quite large fragments are formed during proteolysis (see Fig. 1). The protein fragments produced were isolated by chromatography and their identity determined by a variety of analytical methods [72]. The results of these analyses indicated that the peptic digestion of α-lactalbumin in acid is a sequential process. The initial fast nicking occurs at peptide bond Ala40–Ile41, followed by cleavage at Leu52–Phe53, leading to the excision from the protein of a 12-residue chain segment encompassing the β-sheet region of the native protein and formation of a nicked protein species constituted by fragments 1–40 and 53–123 crosslinked by the four disulfide bonds of the protein. A subsequent slower cleavage occurs at peptide bond Tyr103–Trp104.

The initial proteolytic events are the most useful and critical for unravelling structural and dynamic features of a protein species, since the protein fragments or the nicked protein (if the fragments remain associated) are different conformational and dynamic entities from the intact protein and are more easily degraded by proteolysis as a result of their enhanced flexibility [48]. The peptide bonds Ala40–Ile41 and Leu52–Phe53 suffering fast initial proteolysis in α-lactalbumin in acid are located at chain segments flanking the single antiparallel β-sheet of the protein (residues 40–43 and 47–50; see Fig. 1), thus implying that this region in the A-state is flexible by the criteria of the limited proteolysis. The results of proteolysis experiments are therefore in agreement with the NMR data of the A-state indicating that the region encompassing the β-sheet region is disordered [9,69–71].

The molten globule state of α-lactalbumin can also be generated by removing the single protein-bound calcium ion by the use of chelating agents [18,74]. Recently, we have used proteolytic probes for studying the structural features of the apo form of α-lactalbumin, with a view to comparing the intermediate state of the calcium-free protein at neutral pH with that observed in acid (A-state). For this study, proteinase K and chymotrypsin were employed, displaying a very broad and a somewhat preferential substrate specificity [54], respectively, and an optimum of activity at neutral pH. Whereas the holo form of the protein was resistant to proteolysis, the apo form was efficiently digested by the two proteases. From a timecourse analysis of the proteolysis reactions, it was concluded that both proteinase K and chymotrypsin initially cleave the Gln39–Ala40 and Phe53–Gln54 peptide bonds, followed by slower cleavages at other bonds (A Fontana *et al.*, unpublished data). Proteolytic probes therefore indicate that the partly folded states of α-lactalbumin in acid or upon removal of protein-bound calcium at neutral pH are similar, since in both cases the chain region suffering limited proteolysis events is the same (see above for the results with pepsin). In this context, it is interesting to observe that Matsumura *et al.* [75] studied the susceptibility of bovine α-lactalbumin to the incorporation of primary amines catalyzed by transglutaminase. It was found that Gln54 can be selectively modified by this enzymatic reaction only if the protein is in its calcium-free form. It seems, therefore, that both protease

Figure 1



(a) Schematic representation of the crystal structure of α-lactalbumin. The 123-residue chain of α-lactalbumin is crosslinked by four disulfide bridges, which are labeled by the cysteine residues that they connect (6–120, 28–111, 61–77 and 73–91). The anti-parallel β-sheet region (arrows) encompasses residues 40–43 and 47–50. Calcium is represented by a solid sphere. The model was produced using the program MOLSCRIPT [104] utilizing the crystallographic coordinates of baboon α-lactalbumin [67]. Adapted from [105]. (b) Timecourse analysis by SDS-PAGE [73] of the peptic digestion of α-lactalbumin in acid (A-state). Proteolysis was conducted at room temperature with pepsin (E : S ratio 1:750, by weight) at pH 2.0 in the presence of 0.1 M NaCl. Samples were taken from the reaction mixture at intervals and dissolved in the sample buffer of the electrophoresis systems [73]. Under these solvent conditions, proteolysis is stopped by denaturation of pepsin. A partial BrCN-digest of α-lactalbumin at the level of the Met90 residue in the chain and leading to fragments 1–90 and 91–123 served as a marker of molecular masses. The protein fragments formed at the initial stages of proteolysis were isolated to homogeneity by chromatography and their identity established by quantitative amino acid analysis after acid hydrolysis, N-terminal sequencing and mass analysis by mass spectrometry. The results of these analyses, when compared with the known sequence of α-lactalbumin, allowed in the definition of the sites of initial peptide bond fissions along the protein chain. Analysis of the timecourse of the proteolytic reaction by pepsin indicates that the digestion follows a sequential process, the first peptide bond cleaved being Ala40–Ile41, followed by Leu52–Phe53 and by a much slower cleavage at Tyr103–Trp104 [72]. Of interest, the single chain proteolytic fragment 53–103 containing two disulfide bonds (61–77 and 73–91) and the binding sites for the calcium ion and the two fragments 1–40 and 104–123 covalently linked by two disulfide bonds (6–120 and 28–111) are the species most resistant to further proteolysis by pepsin and accumulating in the proteolytic mixture. Fragments 53–103 and (1–40)(104–123) were isolated to homogeneity by chromatography and their conformational properties investigated.

and transglutaminase probe the same flexible region of the apo form of α-lactalbumin, as reasonably expected from the fact that in both cases the process involves a protein–protein recognition phenomenon.

The proteolysis of α-lactalbumin in acid by pepsin leads, upon prolonged reaction, to fragment 53–103 and to the two chain species constituted by fragments 1–40 and 104–123 covalently linked by disulfide bonds (see Fig. 1). These two fragment species are relatively resistant to further proteolysis and accumulate in the reaction mixture, implying that they possess sufficient stability and rigidity to prevent fast proteolysis. Observations of this type are usually taken as a clear-cut indication of the existence of folded protein domains [52]. On this basis, it can be proposed that the experiments of peptic digestion identify, in the A-state of α-lactalbumin in acid solution, the existence of two relatively stable and somewhat rigid domains. The interesting observation is that the two proteolytic domains 53–103 and (1–40)(104–123) are exactly those one would predict simply from visual inspection of the three-dimensional structure of native α-lactalbumin in

its crystal state (see Fig. 1). The two proteolytic domains encompass the complete polypeptide chain of α-lactalbumin, with the exception of the short segment 41–52 arranged, in the native protein, in a β-sheet secondary structure, which appears to be quite exposed and isolated from the rest of the protein.

The experimental dissection of α-lactalbumin into domains by proteolysis contrasts with the usual description of α- and β-domains of α-lactalbumin utilized by many investigators in discussing structural properties of this protein in its native or partly folded state, since they attribute the α-helical segment 86–99 (helix C) to the α-domain. In particular, a recombinant model of the α-domain constituted by fragments 1–39 and 81–123 connected by a linker of three glycines was produced and characterized [76]. Proteolysis data, as well as inspection of the protein model shown in Figure 1, suggest that perhaps a more suitable domain model would have been a species constituted by fragments 1–40 and 104–123. In this context, it is interesting to observe that an improved algorithm for the automatic identification and location of protein domains from X-ray coordinates of a protein structure has been developed recently [77]. When applied to baboon α-lactalbumin, the algorithm predicts that this protein is constituted by two domains, one given by the continuous chain segment 38–104 and the other by the two segments 1–37 and 105–122. If we consider that the β-sheet portion is likely to be a subdomain within the predicted domain 38–104 (see Fig. 1), the correspondence between domains experimentally isolated by proteolysis and domains predicted by the computer algorithm is indeed striking. The two proteolytic domains produced by peptic digestion of bovine α-lactalbumin have been isolated and their conformational properties investigated. Of interest, domain 53–103 contains all the binding sites for the calcium ion and in the presence of calcium acquires significant structure, as revealed by CD measurements (A Fontana *et al.*, unpublished data).

Summing up, the results of proteolysis experiments indicate that, although α-lactalbumin becomes a much more dynamic entity in acid or by removal of Ca^{2+}, it maintains its overall domain topology and suffers initial nicking at sites flanking the small β-sheet subdomain, which is excised from the protein by proteolysis. Proteolytic probes therefore appear to detect in the molten globule state of α-lactalbumin significant native-like structure. In this context, it should be mentioned that a similar conclusion was reached recently by Barron and co-workers [78,79] on the basis of vibrational Raman optical activity measurements conducted on α-lactalbumin at pH 2.0 over the temperature range 2 to 45°C. These authors concluded that the protein is almost completely native-like at 2°C, still highly native-like at 25°C and with the helical content decreased and the tertiary structure lost at 45°C.

Apomyoglobin

Apomyoglobin (apoMb; myoglobin without the heme) is a small monomeric protein of 153 amino acid residues that has long been used by numerous investigators as a model for studies of protein folding and stability. The structure of apoMb has not yet been determined by X-ray methods and for simplicity, in the past, the structure of apoMb has often been assumed to be similar to that of holoMb. Indeed, the results of a variety of spectroscopic studies have indicated that, in solution and at neutral pH, apoMb retains much of the secondary structure observed in the crystal structure of the heme-containing myoglobin (holoMb). Nonetheless, CD measurements provided evidence that the helical content of apoMb is reduced with respect to that of holoMb, which in the crystal state is highly helical and is constituted by eight helices (A–H; see Fig. 2; [37,80] and references therein). Recently, Lin *et al.* [81] have shown that site-specific mutations of apoMb lead to dramatic effects on the CD and fluorescence emission spectra of the protein and concluded that apoMb at neutral pH is not a well-structured protein and possesses characteristics of a molten globule.

With a view to probing the structure and dynamics of apoMb utilizing the limited proteolysis approach, apoMb was subjected to a series of proteolysis experiments with subtilisin, thermolysin, chymotrypsin and trypsin [80]. At neutral pH and 25°C, these proteases initially cleave apoMb at a very restricted region of the 153-residue polypeptide chain. When reacted under identical experimental conditions, holoMb was instead fully resistant to proteolysis, in line with the proposal that native holoMb is expected to be cleaved by proteases 10^{19} times more slowly than the fully unfolded protein [82]. From each proteolytic digest, two major fragments accounting for the full length of the 153-residue chain of apoMb were isolated. As shown in Figure 2, proteolytic cleavages occur at chain segment 89–96, which in the crystal state of native holoMb encompasses helix F (residues 82–97) [83]. From the proteolytic digests of apoMb with trypsin and thermolysin, small amounts of fragments 1–31 and 1–32, respectively, were also isolated. The results of limited proteolysis experiments therefore indicate that, with apoMb in solution and at neutral pH, the chain segment corresponding to helix F in holoMb is at least highly flexible or disrupted.

The conclusions reached from proteolysis experiments on apoMb are nicely and precisely supported by a number of additional experiments and theoretical studies on the structure and dynamics of apoMb. First of all, hydrogen exchange experiments by NMR [84] fail to detect protection in the chain segment encompassing helix F, whereas the rest of the hydrophobic helical core of apoMb shows structural and dynamic properties similar to those of native holoMb. Moreover, the results of molecular dynamics

simulations conducted on apoMb provided evidence that helix F is substantially more mobile than the core of the protein [85–87]. The additional and minor proteolytic cleavages at chain segment 30–33 (see above) can be taken as an indication that the heme removal from holoMb causes some chain motility also at this site of the chain, even if minor in respect to that of region 89–96. Of interest, NMR data indicate that helix B in sperm whale apoMb is folded up to residue 30 and that "past this residue the constraints are few and fraying is seen" [84]. Molecular dynamics simulations of the helices of apoMb indicate that "helix B shows greater unwinding of the C-terminus" [86]. It is quite reassuring that the results of proteolysis experiments are in full agreement with those derived from other physicochemical measurements, emphasizing, therefore, the reliability of proteolytic probes of protein structure and dynamics.

Limited proteolysis data indicate that, with the exception of the region suffering fast proteolysis, the rest of the polypeptide chain of apoMb appears to maintain a native-like structure, characterized by sufficient rigidity to prevent extensive proteolysis. The results of limited proteolysis experiments conducted on apoMb are not different from those obtained with many other proteins subjected to limited proteolysis, in the sense that a variety of nicked proteins have been described in the literature, as a result of a proteolytic nicking at flexible or disordered chain regions [41,43,48,55]. There is no reason, therefore, to consider the state of apoMb at neutral pH a molten globule [81], since this concept clearly does not apply to the many other proteins that, in their native state, suffer specific peptide bond fission(s) by proteolysis. In fact, in a recent study, Eliezer and Wright [37] utilized multi-dimensional hetero-nuclear NMR spectroscopy to analyze the structure of isotopically labelled recombinant apomyoglobin and concluded that the state of the protein at neutral pH possesses the typical characteristics expected for a native globular protein, with the exception of a fluctuating or disordered chain segment encompassing helix F. We conclude, therefore, that apoMb at neutral pH is not a molten globule and we concur with the opinion expressed also by others [21,37] that there is no reason to introduce a new term of 'highly ordered molten globule' [88,89] to describe the state of apoMb, since it appears that extending the term molten globule to include all partly folded sates of proteins would devalue its significance and utility.

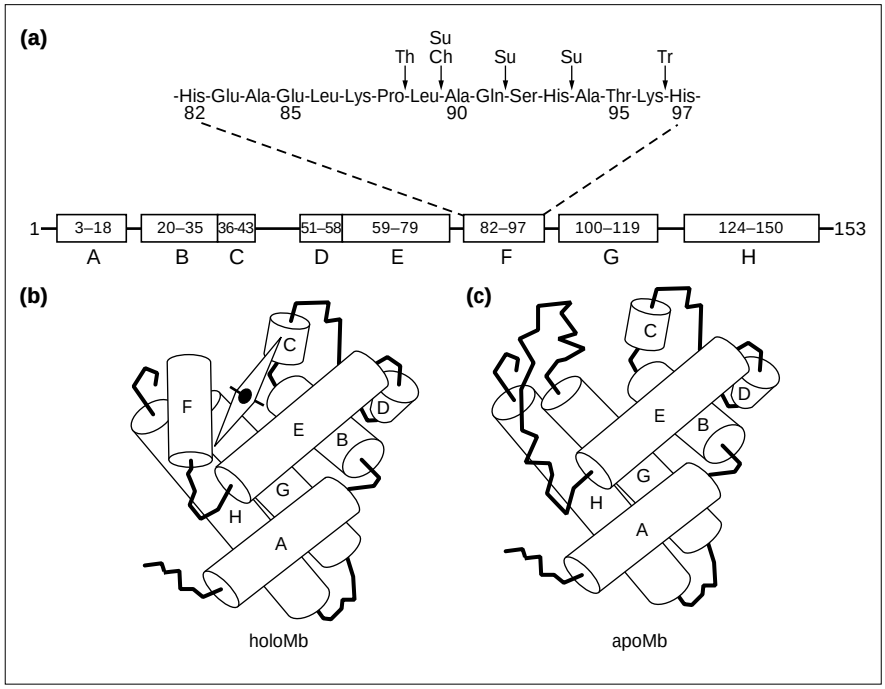
Proteins in aqueous trifluoroethanol

Weakly polar alcohols including TFE have been shown to disrupt the native conformation of globular proteins, but the resulting denatured state was much different from the random-coil state observed in the presence of chemical denaturants such as urea or guanidine hydrochloride. Recent studies, based on CD and NMR measurements, have shown that the TFE state of a protein, e.g. as obtained by dissolving the protein in 50% (v/v) aqueous TFE, is a stable partially folded state with a high content of α-helical conformation, but lacking the specific tertiary interactions of the native protein [22–25]. This TFE state appears to be a noncompact expanded conformational state characterized by an ensemble of fluctuating helices. The TFE state of proteins has been analyzed recently by a limited proteolysis approach utilizing the thermophilic protease thermolysin. This protease appeared to be the most suitable proteolytic probe because of its noteworthy stability under relatively harsh solvent conditions, including organic solvents [90], and broad substrate specificity [91,92].

A variety of model proteins have been reacted with thermolysin in 50% (v/v) TFE at neutral pH and 20–52°C for several hours and then the protein fragments isolated and analyzed for their identity. The results obtained with five model proteins, bovine α-lactalbumin [72], pancreatic ribonuclease A (RNase) [93], horse cytochrome *c* [94], hen egg-white lysozyme [95] and horse apomyoglobin (A Fontana *et al.*, unpublished data), were quite striking, since they indicated that the procedure of cleaving proteins in their TFE state by thermolysin is highly selective. Proteolysis was found to occur at the amino-side of leucine, isoleucine and phenylalanine only, even if thermolysin in aqueous solution cleaves at a variety of other residues as well [91,92]. Nevertheless, even if the model proteins investigated contain many leucine, isoleucine and phenylalanine residues, thermolysin cleaves specifically at a few sites only. In each case, it was possible to identify one peptide bond being cleaved first by thermolysin, followed by much slower cleavages at a few additional peptide bonds. In the case of lysozyme [95] and RNase [93], nicked proteins were isolated to homogeneity and their functional, conformational and stability properties investigated. Accepting our view that helical chain segments are not cleaved by the proteolytic probe (see above), the selective proteolytic digestion of the model proteins in their TFE state by thermolysin is consistent with the fact that this conformational state is highly helical. The slow but selective proteolysis is also due to the fact that thermolysin is much less active in the presence of TFE and since proteases, in the presence of organic solvents, catalyze the reverse reaction, i.e. the synthesis instead of the hydrolysis of peptide bonds [96], thus rendering proteolysis much slower. Nevertheless, it was possible to interpret the proteolysis data of proteins in their TFE state on the basis of features of the structure and dynamics of the native proteins in aqueous solution, implying that the helical TFE state is related to the native state. As an example, we discuss below the case of RNase.

RNase is cleaved by thermolysin in 30–50% (v/v) TFE at the level of the peptide bond Asn34–Leu35, followed by a

Figure 2



Limited proteolysis of horse apomyoglobin (apoMb) by subtilisin (Su), thermolysin (Th), chymotrypsin (Ch) and trypsin (Tr) [80]. Proteolysis of apoMb by the four proteases (E : S ratio 1:100, by weight) was conducted at 25°C for up to 15 min at neutral pH. At intervals, aliquots were removed from the reaction mixture and, after stopping proteolysis by acidification to pH ~2, the samples were analyzed by SDS-PAGE [73] and reverse phase HPLC. Protein fragments were isolated by micropreparative HPLC and analyzed by N-terminal sequencing and mass spectrometry in order to determine their identity and thus the sites of initial proteolytic cleavages along the 153-residue chain of apoMb. **(a)** The A–H helices of native horse holoMb [83] are indicated by boxes. The amino acid sequence of helix F (residues 82–97) is shown and the sites of initial proteolytic cleavage are indicated by arrows. Below are shown the schematic three-dimensional structure of **(b)** holoMb and **(c)** apoMb. The location of the heme is shown in holoMb and the eight helices are indicated by letters A–H. Note that helix F in the model of apoMb is assumed to be unfolded. The detailed structure of the corresponding chain segment is not known, but the results of limited proteolysis experiments provide evidence that this helix is largely disrupted. Reproduced from [80].

slower cleavage at Thr45–Phe46 [93]. On the other hand, thermolysin does not cleave RNase when reacted at 20–22°C in aqueous buffer only. Two nicked RNase species, resulting from cleavages at one or two peptide bonds and thus constituting two (1–34 and 35–124) and three (1–34, 35–45 and 46–124) fragments linked covalently by the four disulfide bonds of the protein, were isolated and characterized. Nicked RNase with one peptide bond cleaved was found to be folded in a native-like conformation and retained 20% of the catalytic activity of the native enzyme, while the protein species with two peptide bonds hydrolyzed was unfolded and inactive.

The highly preferential cleavage of peptide bond Asn34–Leu35 in RNase in its TFE state by thermolysin implies that this protein state is folded and rigid, preventing proteolysis but still displaying a flexible region amenable to proteolysis. Of interest, the site of initial proteolysis is located, in native RNase, at a flexible loop displaying the largest dispersion values of the backbone angles of the protein, as given by NMR measurements [97]. Even the refined X-ray structure of native RNase revealed that "Asn34 is located in a surface region with uncertain electron density and its conformation is still unknown," thus indicating a static/dynamic disorder of this site in the intact native protein [98]. Moreover, hydrogen exchange measurements by NMR of RNase revealed that chain segment 30–36 exchanged protons at a much faster rate than other sites of the protein and that the rate was too fast to be measured [99]. All these data therefore indicate that the chain region encompassing peptide bond Asn34–Leu35 displays higher flexibility than the rest of the protein chain both in the native state and TFE state of RNase, implying some similarity of structure and dynamics between the two states.

Conclusions

This review summarizes the results that have been obtained so far utilizing proteolytic probes of the structure and dynamics of partly folded states or molten globules of proteins. The technique is simple to use, modest in demands for protein sample requirements, instrumentation and experimental effort and, moreover, it provides data on the solution structure of a protein, even if the data do not reach the high-resolution level of other physicochemical techniques. The approach has been utilized to probe nonnative states of proteins exposed to different solvent conditions and has provided evidence that these protein intermediates possess significant native-like properties, in agreement with the results obtained recently utilizing vibrational Raman optical activity [78,79] or multi-dimensional hetero-nuclear NMR measurements [37]. Even the conformational features of proteins dissolved in solutions containing a high concentration of TFE can be probed by proteolysis, in this case utilizing the thermophilic protease thermolysin [58]. Proteolysis of proteins in their TFE state can be very specific, implying that proteins maintain in aqueous TFE a somewhat rigid

structure, preventing extensive proteolysis but still possessing site(s) of sufficient flexibility to permit the attack and hydrolysis by the proteolytic probe.

The limited proteolysis approach was used to analyze the folding intermediates exhibited by two well-characterized proteins (α-lactalbumin and apomyoglobin) utilized over the years by many investigators as prototype model proteins for addressing the issues of existence and properties of molten globules of proteins with the aid of a variety of other experimental techniques and approaches [5,16–18,100]. The most important result emerging from proteolysis studies is that both α-lactalbumin and apomyoglobin suffer limited proteolysis at very few peptide bonds along their polypeptide chain, whereas native α-lactalbumin and holomyoglobin are resistant to proteolysis. That molten globules of proteins are relatively rigid protein entities and thus show resistance to extensive proteolysis was not easily anticipated, since these protein intermediates have been shown before by hydrogen/deuterium exchange to be highly dynamic protein entities, if compared with the corresponding native ones [9,70,71,101]. The protection factors of amides in various partly folded states of proteins, including the TFE state [22], have been found to be up to 500, whereas the same amides in the native protein are 10^5–10^7 [22,27,101]. Nevertheless, proteolysis is selective and initially occurs at sites of the chain previously shown to be flexible or disordered by NMR and hydrogen exchange measurements, as expected from the fact that proteolytic probes pinpoint in globular proteins sites devoid of regular secondary structure and displaying high segmental mobility [49,102,103]. It seems, therefore, that proteolytic probes can detect residual, sufficiently persistent structure in proteins, even if showing marginal protection from solvent exchange.

The route by which an unfolded polypeptide chain acquires its native folded state is still poorly understood. We suggest that the fundamental and difficult problem of protein folding should be investigated by taking advantage of a variety of experimental techniques, including the simple biochemical approach relying on the use of proteolytic probes of protein structure and dynamics.

Acknowledgements

This work was supported in part by the Italian National Council of Research (CNR). The authors thank Anna Mocavero for the careful typing of the manuscript.

References

1. Creighton, T.E. (ed.) (1992). *Protein Folding.* Freeman & Co., New York.
2. Fersht, A.R. & Dill, K.A. (ed.) (1994). Folding and binding. *Curr. Opin. Struct. Biol.* 4, 67–129.
3. Baldwin, R.L. (1994). Finding intermediates in protein folding. *Bioessays* 16, 207–210.
4. Creighton, T.E., Darby, N. & Kemmink, J. (1996). The roles of partly folded intermediates in protein folding. *FASEB J.* 10, 110–118.
5. Privalov, P.L. (1996). Intermediate states in protein folding. *J. Mol. Biol.* 258, 707–725.
6. Privalov, P.L. (1979). Stability of proteins: small globular proteins. *Adv. Protein Chem.* 33, 167–241.
7. Englander, S.W. & Maine, L. (1992). Protein folding studied using hydrogen exchange labeling and two-dimensional NMR. *Annu. Rev. Biophys. Biomol. Struct*. 21, 243–265.
8. Jennings, P.A. & Wright, P.E. (1993). Formation of a molten globule intermediate early in the kinetic folding pathway of apomyoglobin. *Science* 262, 892–896.
9. Balbach, J., Forge, V., Van Nuland, N.A.J., Winder, S.L., Hore, P.J. & Dobson, C.M. (1995). Following protein folding in real time using NMR spectroscopy. *Nat. Struct. Biol*. 2, 865–870.
10. Matthews, C.R. (1993). Pathway of protein folding. *Annu. Rev. Biochem*. 62, 653–683.
11. Roder, H. & Colòn, W. (1997). Kinetic role of early intermediates in protein folding. *Curr. Opin. Struct. Biol*. 7, 15–28.
12. Fink, A.L. (1995). Compact intermediate states in protein folding. *Annu. Rev. Biophys. Biomol. Struct.* 24, 495–522.
13. Freire, E. (1995). Thermodynamics of partly folded intermediates in proteins. *Annu. Rev. Biophys. Biomol. Struct*. 24, 141–165.
14. Ptitsyn, O.B. (1992). The molten globule state. In *Protein Folding.* (Creighton, T.E., ed.), pp. 243–300, Freeman & Co., New York.
15. Ptitsyn, O.B. (1995). Molten globule and protein folding. *Adv. Protein Chem*. 47, 83–229.
16. Kuwajima, K. (1989). The molten globule state as a clue for understanding the folding and cooperativity of globular-protein structure. *Proteins* 6, 87–103.
17. Ptitsyn, O.B. (1995). Structures of folding intermediates. *Curr. Opin. Struct. Biol*. 5, 74–78.
18. Kuwajima, K. (1996). The molten globule state of α-lactalbumin. *FASEB J.* 10, 102–109.
19. Creighton, T.E. (1997). How important is the molten globule for correct protein folding? *Trends Biol. Sci.* 22, 6–10.
20. Ewbank, J.J., Creighton, T.E., Hayer-Hartl, M.K. & Hartl, F.U. (1995). What is the molten globule? *Nat. Struct. Biol*. 2, 10.
21. Okazaki, A., Ikura, T. & Kuwajima, K. (1995). What is the molten globule? Reply. *Nat. Struct. Biol*. 2, 10–11.
22. Buck, M., Radford, S.E. & Dobson, C.M. (1993). A partially folded state of hen egg-white lysozyme in trifluoroethanol: structural characterization and implications for protein folding. *Biochemistry* 32, 669–678.
23. Alexandrescu, A.T., Ng, Y.-L. & Dobson, C.M. (1994). Characterization of a TFE-induced partially folded state of α-lactalbumin. *J. Mol. Biol*. 235, 587–599.
24. Buck, M., Schwalbe, H. & Dobson, C.M. (1995). Characterization of conformational preferences in a partly folded protein by heteronuclear NMR spectroscopy: assignment and secondary structure analysis of hen egg-white lysozyme in trifluoroethanol. *Biochemistry* 3, 13219–13232.
25. Shiraki, K., Nishikawa, K. & Goto, Y. (1994). Trifluoroethanol-induced stabilization of the α-helical structure of β-lactoglobulin: implications for non-hierarchical protein folding. *J. Mol. Biol.* 245, 180–194.
26. Neri, D., Billeter, M., Wider, G. & Wüthrich (1992). NMR determination of residual structure in a urea-denatured protein, the 434-repressor. *Science* 257, 1559–1563.
27. Radford, S.E., Buck, M., Topping, K.D., Dobson, C.M. & Evans, P.A. (1992). Hydrogen exchange in native and denatured states of hen egg-white lysozyme. *Proteins* 14, 237–248.
28. Wüthrich, K. (1994). NMR assignments as a basis for structural characterization of denatured states of globular proteins. *Curr. Opin. Struct. Biol*. 4, 93–99.
29. Logan, T.M., Thériault, Y. & Fesik, S.W. (1994). Structural characterization of the FK506-binding protein unfolded in urea and guanidine hydrochloride. *J. Mol. Biol.* 236, 637–648.
30. Smith, L.J., Fiebig, K.M., Schwalbe, H. & Dobson, C.M. (1996). The concept of random coil: residual structure in peptides and denatured proteins. *Fold. Des.* 1, R95–R106; 1359-0278-001-R0095.
31. Shortle, D. (1996). The denatured state (the other half of the folding equation) and its role in protein stability. *FASEB J.* 10, 27–34.
32. Baldwin, R.L. (1993). Pulsed H/D-exchange studies of folding intermediates. *Curr. Opin. Struct. Biol.* 4, 87–92.
33. Dobson, C.M. (1994). Solid evidence for molten globules. *Curr. Biol.* 4, 636–640.
34. Gladwin, S.T. & Evans, P.A. (1996). Structure of very early protein folding intermediates: new insights through a variant of hydrogen exchange labelling. *Fold. Des*. 1, 407–417; 1359-0278-001-00407.
35. Improta, S., Molinari, H., Pastore, A., Consonni, R. & Zetta, L. (1995). Probing protein structure by solvent perturbation of NMR spectra:

photochemically induced dynamic nuclear polarization and paramagnetic perturbation techniques applied to the study of the molten globule state of α-lactalbumin. *Eur. J. Biochem*. **227**, 87–96.
36. Thériault, Y., Pochapsky, T.C., Dalvit, C., Chiu, M.L., Sligar, S.G. & Wright, P.E. (1994). ^{1}H and ^{15}N resonance assignments and secondary structure of the carbon monoxide complex of sperm whale myoglobin. *J. Biomol. NMR* **4**, 491–504.
37. Eliezer, D. & Wright, P.E. (1996). Is apomyoglobin a molten globule? Structural characterization by NMR. *J. Mol. Biol*. **263**, 531–538.
38. Kataoka, M., Nishii, I., Fujisawa, T., Ueki, T., Tokumaga, F. & Goto, Y. (1995). Structural characterization of molten globule and native states of apomyoglobin by solution X-ray scattering. *J. Mol. Biol*. **249**, 215–228.
39. Kataoka, M., Kuwajima, K., Tokunaga, F. & Goto, Y. (1997). Structural characterization of the molten globule of α-lactalbumin by solution X-ray scattering. *Protein Sci*. **6**, in press.
40. Kataoka, M. & Goto, Y. (1996). X-ray solution scattering studies of protein folding. *Fold. Des.* **1**, R107–R114; 1359-0278-001-R0107.
41. Mihalyi, E. (1978). *Application of Proteolytic Enzymes to Protein Structure Studies*. CRC Press, Boca Raton, FL.
42. Neurath, H. (1980). Limited proteolysis, protein folding and physiological regulation. In *Protein Folding.* (Jaenicke, R., ed.), pp. 501–504, Elsevier/North Holland Biomedical Press, Amsterdam/New York.
43. Price, N.C. & Johnson, C.M. (1990). Proteinases as probes of conformation of soluble proteins. In *Proteolytic Enzymes. A Practical Approach.* (Beynon, R.J. & Bond, J.S., eds.), pp. 163–180, IRL Press, Oxford.
44. Schechter, I. & Berger, A. (1967). On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Res. Commun*. **27**, 157–162.
45. Herschlag, D. (1988). The role of induced fit and conformational changes of enzymes in specificity and catalysis. *Bioorganic Chem.* **16**, 62–96.
46. Fontana, A. (1989). Limited proteolysis of globular proteins occur at exposed and flexible loops. In *Highlights of Modern Biochemistry*. (Kotyk, A., Skoda, J., Pacek, V. & Kostka, W., eds.), pp. 1711–1726, VSP International Publications, Amsterdam.
47. Fontana, A., Vita, C., Dalzoppo, D. & Zambonin, M. (1989). Limited proteolysis as a tool to detect structure and dynamic features of globular proteins: studies on thermolysin. In *Methods in Protein Sequence Analysis.* (Wittman-Liebold, B., ed.), pp. 315–324, Springer-Verlag, Berlin.
48. Fontana, A., Polverino de Laureto, P. & De Filippis, V. (1993). Molecular aspects of proteolysis of globular proteins. In *Protein Stability and Stabilization*. (Van den Tweel, W., Harder, A. & Buitelaar, M., eds.), pp. 101–110, Elsevier Science Publications, Amsterdam.
49. Fontana, A., Fassina, G., Vita, C., Dalzoppo, D., Zamai, M. & Zambonin, M, (1986). Correlation between sites of limited proteolysis and segmental mobility in thermolysin. *Biochemistry* **25**, 1847–1851.
50. Fassina, G., Vita, C., Dalzoppo, D., Zamai, M., Zambonin, M. & Fontana, A. (1986). Autolysis of thermolysin: isolation and characterization of a folded three-fragment complex. *Eur. J. Biochem*. **156**, 221–228.
51. Vita, C., Dalzoppo, D. & Fontana, A. (1985). Limited proteolysis of thermolysin by subtilisin: isolation and characterization of a partially active enzyme derivative. *Biochemistry* **24**, 1798–1806.
52. Wetlaufer, D.B. (1981). Folding of protein fragments. *Adv. Protein Chem.* **34**, 61–92.
53. Cohen, S.L. (1996). Domain elucidation by mass spectrometry. *Curr. Biol*. **4**, 1013–1016.
54. Bond, J.S. (1990). Commercially available proteases. In *Proteolytic Enzymes: A Practical Approach*. (Beynon, R.J. & Bond, J.S. eds.), pp. 232–240, IRL Press, Oxford.
55. Wilson, J.E. (1991). The use of monoclonal antibodies and limited proteolysis in elucidation of structure-function relationships in proteins. *Methods Biochem. Anal*. **351**, 207–250.
56. Novotny, J. & Bruccoleri, R.E. (1987). Correlation among sites of limited proteolysis, enzyme accessibility and segmental mobility. *FEBS Lett.* **211**, 185–189.
57. Hubbard, S.J., Campbell, S.F. & Thornton, J.M. (1991). Molecular recognition: conformational analysis of limited proteolytic sites and serine proteinase protein inhibitors. *J. Mol. Biol.* **220**, 507–530.
58. Holmes, M.A. & Matthews, B.W. (1982). Structure of thermolysin refined at 1.6 Å resolution. *J. Mol. Biol.* **160**, 623–639.
59. Frauenfelder, H., Petsko, G.A. & Tsernoglou, D. (1979). Temperature-dependent X-ray diffraction as a probe of protein structural dynamics. *Nature* **280**, 558–563.
60. Sternberg, M.J.E., Grace, D.E.P. & Phillips, D.C. (1979). Dynamic information from protein crystallography: an analysis of temperature factors from refinement of the hen egg-white lysozyme. *J. Mol. Biol*. **130**, 231–253.
61. Ringe, D. & Petsko, G.A. (1985). Mapping protein dynamics by X-ray diffraction. *Prog. Biophys. Mol. Biol*. **45**, 197–235.
62. Hubbard, S.J., Eisenmenger, F. & Thornton, J.M. (1994). Modeling studies of the change in conformation required for cleavage of limited proteolytic sites. *Protein Sci*. **3**, 757–768.
63. Krokozsynska, I. & Otlewski, J. (1996). Thermodynamic stability effects of single peptide bond hydrolysis in protein inhibitors of serine proteases. *J. Mol. Biol.* **256**, 793–802.
64. Dalzoppo, D., Vita, C. & Fontana, A. (1985). Folding of thermolysin fragments: identification of the minimum size of a carboxyl-terminal fragment that can fold into a stable native like structure. *J. Mol. Biol*. **182**, 331–340.
65. Darby, N.J., Kemmink, J. & Creighton, T.E. (1996). Identifying and characterizing a structural domain of protein disulfide isomerase. *Biochemistry* **35**, 10517–10528.
66. Vindigni, A., De Filippis, V., Zanotti, G., Visco, C., Orsini, G. & Fontana, A. (1994). Probing the structure of hirudin from *Hirudinaria manillensis* by limited proteolysis: isolation, characterization and thrombin-inhibitory properties of N-terminal fragments. *Eur. J. Biochem.* **226**, 323–333.
67. Acharia, K.R., Stuart, D.I., Walker, N.P.C., Lewis, M. & Phillips, D.C. (1989). Refined structure of baboon α-lactalbumin at 1.7 Å resolution. *J. Mol. Biol*. **208**, 99–127.
68. Acharia, K.R., Ren, J., Stuart, D.I. & Phillips, D.C. (1991). Crystal structure of human α-lactalbumin at 1.7 Å resolution. *J. Mol. Biol*. **221**, 571–581.
69. Baum, J., Dobson, C.M., Evans, P.A. & Hanley, C. (1989). Characterization of a partly folded protein by NMR methods: studies on the molten globule state of guinea pig α-lactalbumin. *Biochemistry* **28**, 7–13.
70. Alexandrescu, A.T., Evans, P.A., Pitkeathly, M., Baum, J. & Dobson, C.M. (1993). Structure and dynamics of the acid-denatured molten globule state of α-lactalbumin: a two-dimensional NMR study. *Biochemistry* **32**, 1707–1718.
71. Chyan, C.L., Wormald, C., Dobson, C.M., Evans, P.A. & Baum, J. (1993). Structure and stability of the molten globule state of guinea-pig α-lactalbumin: a hydrogen exchange study. *Biochemistry* **32**, 5681–5691.
72. Polverino de Laureto, P., De Filippis, V., Di Bello, M., Zambonin, M. & Fontana, A. (1995). Probing the molten globule state of α-lactalbumin by limited proteolysis. *Biochemistry* **34**, 12596–12604.
73. Schägger, H. & Von Jagow, G. (1987). Tricine sodium dodecyl sulfate polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 kDa to 100 kDa. *Analyt. Biochem*. **166**, 368–379.
74. Kronman, M.J. (1989). Metal-ion binding and the molecular conformational properties of α-lactalbumin. *Crit. Rev. Biochem. Mol. Biol*. **24**, 565–667.
75. Matsumura, J., Chanyongvorakul, Y., Kumazawa, Y., Ohtsuka, T. & Mori, T. (1996). Enhanced susceptibility to transglutaminase reaction of α-lactalbumin in the molten globule state. *Biochim. Biophys. Acta* **1292**, 69–76.
76. Peng, Z. & Kim, P.S. (1994). A protein dissection study of a molten globule. *Biochemistry* **33**, 136–141.
77. Siddiqui, A.S. & Barton, G.J. (1995). Continuous and discontinuous domains: an algorithm for the automatic generation of reliable protein domain definitions. *Protein Sci.* **4**, 872–884.
78. Wilson, G., Ford, S.J., Cooper, A., Hecht, L., Wen, Z.Q. & Barron, L.D. (1995). Vibrational Raman optical activity of α-lactalbumin: comparison with lysozyme and evidence for native tertiary folds in molten globule states. *J. Mol. Biol*. **254**, 747–760.
79. Wilson, G., Hecht, L. & Barron, L.D. (1996). The native tertiary fold in molten globule α-lactalbumin appears to be controlled by a continuous phase transition. *J. Mol. Biol*. **261**, 341–347.
80. Fontana, A., Zambonin, M., Polverino de Laureto, P., De Filippis, V., Clementi, A. & Scaramella, E. (1997). Probing the conformational state of apomyoglobin by limited proteolysis. *J. Mol. Biol*. **266**, 223–230.
81. Lin, L., Pinker, R.J., Forde, K., Rose, G.D. & Kallenbach, N.R. (1994). Molten globular characteristics of the native state of apomyoglobin. *Nat. Struct. Biol.* **1**, 447–451.
82. McLendon, C. & Radany, E. (1978). Is protein turnover thermodynamically controlled? *J. Biol. Chem.* **253**, 6335–6337.
83. Evans, S.V. & Brayer, G.D. (1990). High-resolution study of the three-dimensional structure of horse heart metmyoglobin. *J. Mol. Biol*. **213**, 885–897.

84. Lecomte, J.T.J., Kao, Y.-H. & 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* **25**, 267–285.
85. Brooks, C.L. (1992). Characterization of native apomyoglobin by molecular dynamics simulation. *J. Mol. Biol*. **227**, 375–380.
86. Hirst, J.D. & Brooks, C.L. (1995). Molecular dynamics simulations of isolated helices of myoglobin. *Biochemistry* **34**, 7614–7621.
87. Tirado-Rives, J. & Jorgensen, W.L. (1993). Molecular dynamics simulations of the unfolding of apomyoglobin in water. *Biochemistry* **32**, 4175–4184.
88. Redfield, C., Smith, R.A.G. & Dobson, C.M. (1994). Structural characterization of a highly-ordered molten globule at low pH. *Nat. Struct. Biol*. **1**, 23–29.
89. Laidig, K.E. & Daggett, V. (1996). Molecular dynamics simulations of apocytochrome b_{562} – the highly ordered limit of molten globules. *Fold. Des*. **1**, 335–346; 1359-0278-001-00335.
90. Welinder, K.G. (1988). Generation of peptides suitable for sequence analysis by proteolytic cleavage in reversed-phase high-performance liquid chromatography solvents. *Analyt. Biochem*. **174**, 54–64.
91. Heinrikson, R.L. (1977). Applications of thermolysin in protein structural analysis. *Methods Enzymol*. **47**, 175–189.
92. Keil, B. (1982). Enzymic cleavage of proteins. In *Methods in Protein Sequence Analysis*. (Elzinga, M., ed.), vol IV, pp. 291–304, Humana Press, Clifton, NJ.
93. Polverino de Laureto, P., Scaramella, E., De Filippis, V., Bruix, M., Rico, M. & Fontana, A. (1997). Limited proteolysis of ribonuclease A with thermolysin in trifluoroethanol. *Protein Sci*. **6**, in press.
94. Fontana, A., Zambonin, M., De Filippis, V., Bosco, M. & Polverino de Laureto, P. (1995). Limited proteolysis of cytochrome c in trifluoroethanol. *FEBS Lett.* **362**, 266–270.
95. Polverino de Laureto, P., De Filippis, V., Scaramella, E., Zambonin, M. & Fontana, A. (1995). Limited proteolysis of lysozyme in trifluoroethanol: isolation and characterization of a partially active enzyme derivative. *Eur. J. Biochem.* **230**, 779–787.
96. Wayne, S.I. & Fruton, J.S. (1983). Thermolysin-catalyzed peptide bond synthesis. *Proc. Natl. Acad. Sci. USA* **80**, 3241–3244.
97. Santoro, J., *et al*., & Rico, M. (1993). High-resolution three-dimensional structure of ribonuclease A in solution by nuclear magnetic resonance spectroscopy. *J. Mol. Biol.* **229**, 722–734.
98. Wlodawer, A., Bott, R. & Sjölin, L. (1982). The refined crystal structure of ribonuclease A at 2.0 Å resolution. *J. Biol. Chem*. **257**, 1325–1332.
99. Kiefhaber, T. & Baldwin, R.L. (1995). Kinetics of hydrogen bond breakage in the process of unfolding of ribonuclease A measured by pulsed hydrogen exchange. *Proc. Natl. Acad. Sci. USA* **92**, 2657–2661.
100. Evans, P.A. & Radford, S.E. (1994). Probing the structure of folding intermediates. *Curr. Opin. Struct. Biol.* **4**, 100–106.
101. Schulman, B.A., Redfield, C., Peng, Z.Y., Dobson, C.M. & Kim, P.S. (1995). Different subdomains are most protected from hydrogen exchange in molten globule and native states of human α-lactalbumin. *J. Mol. Biol.* **253**, 651–657.
102. Signor, G., *et al*., & Grandi, G. (1990). Structural features of neutral protease from *Bacillus subtilis* deduced from model-building and limited proteolysis experiments. *Eur. J. Biochem.* **189**, 221–227.
103. Polverino de Laureto, P., Toma, S., Tonon, G. & Fontana, A. (1995). Probing the structure of human growth hormone by limited proteolysis. *Intl. J. Pept. Protein Res*. **45**, 200–208.
104. Kraulis, J.P. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr*. **24**, 946–950.
105. Hendrix, T.M., Griko, Y. & Privalov, P. (1996). Energetics of structural domains in α-lactalbumin. *Protein Sci*. **5**, 923–931.