doi:10.1016/j.jmb.2003.10.082 J. Mol. Biol. (2004) 336, 253–262

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

The Regions of the Sequence Most Exposed to the Solvent Within the Amyloidogenic State of a Protein Initiate the Aggregation Process

Maria Monti[1], Barbara Lelj Garolla di Bard[1], Giulia Calloni[2] Fabrizio Chiti[2], Angela Amoresano[1], Gianpietro Ramponi[2] and Piero Pucci[1]*

[1]*Dipartimento di Chimica Organica e Biochimica Università di Napoli Federico II via Cinthia 6, 80126 Naples Italy*

[2]*Dipartimento di Scienze Biochimiche, Università degli Studi di Firenze Viale Morgagni 50 50134 Firenze, Italy*

**Corresponding author*

Formation of misfolded aggregates is an essential part of what proteins can do. The process of protein aggregation is central to many human diseases and any aggregating event needs to be prevented within a cell and in protein design. In order to aggregate, a protein needs to unfold its native state, at least partially. The conformational state that is prone to aggregate is difficult to study, due to its aggregating potential and heterogeneous nature. Here, we use a systematic approach of limited proteolysis, in combination with electrospray ionisation mass spectrometry, to investigate the regions that are most flexible and solvent-exposed within the native, ligand-bound and amyloidogenic states of muscle acylphosphatase (AcP), a protein previously shown to form amyloid fibrils in the presence of trifluoroethanol. Seven proteases with different degrees of specificity have been used for this purpose. Following exposure to the aggregating conditions, a number of sites along the sequence of AcP become susceptible to proteolytic digestion. The pattern of proteolytic cleavages obtained under these conditions is considerably different from that of the native and ligand-bound conformations and includes a portion within the N-terminal tail of the protein (residues 6-7), the region of the sequence 18–23 and the position 94 near the C terminus. There is a significant overlap between the regions of the sequence found to be solvent-exposed from the present study and those previously identified to be critical in the rate-determining steps of aggregation from protein engineering approaches. This indicates that a considerable degree of solvent exposure is a feature of the portions of a protein that initiate the process of aggregation.

© 2003 Elsevier Ltd. All rights reserved.

Keywords: mass spectrometry; limited proteolysis; acylphosphatase; aggregation; amyloid

Introduction

Protein aggregation is a key aspect of protein chemistry. Prior to attaining the functional native state, a polypeptide chain exists as an ensemble of unfolded conformations. Partially folded states may also be populated during the folding process or as a consequence of partial denaturation of the native state resulting from misprocessing of the polypeptide chain, mutation or perturbation of the microenvironment conditions. In such ensembles of unfolded or partially folded states many of the hydrophobic and amide groups are partially exposed to the solvent and susceptible to form the intermolecular interactions that pave the way to the formation of large protein aggregates.[1]

Protein aggregation is an undesired phenomenon that is responsible for many human diseases ranging from various forms of systemic amyloidoses to neurodegenerative disorders such as Alzheimer's and Parkinson's diseases. These aggregation events lead ultimately to the formation of stable protein aggregates, called amyloid fibrils,

Abbreviations used: AcP, acylphosphatase; TFE, trifluoroethanol.

E-mail address of the corresponding author: pucci@unina.it

0022-2836/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.

that are a unifying pathological feature of the diverse forms of amyloidosis.

Until very recently it was thought that amyloid formation was a characteristic limited to the few polypeptide chains associated with such disorders. Nevertheless, a number of independent observations have indicated that this behaviour is indeed shared by a number of polypeptide chains not involved in disease under partially denaturing conditions.[2–9] This has led to the hypothesis that amyloid aggregation is a generic property of proteins.[1] The investigation of the fundamental aspects of protein aggregation using a large number of model systems is relevant not just for the elucidation of a process that is central to human illness, but also because it creates new perspectives for understanding how natural proteins have achieved the goal of escaping effectively from aggregation under conditions close to physiological. Indeed, the observation that protein aggregates forming early on the route to amyloid formation can be inherently highly cytotoxic[10] suggests that Nature has actively combated against uncontrolled aggregation by designing mechanisms of quality control, but also by selecting amino acid sequences with a sufficiently low propensity to aggregate. The investigation of the aggregation process by disease-unrelated proteins and of the strategies devised by natural amino acid sequences to circumvent this process under conditions close to physiological has implications for understanding how protein sequences have evolved and for establishing new rules for successful protein design. The intrinsic potential of protein sequences to aggregate into ordered amyloid fibrils and the direct observation that folding and aggregation may occur concomitantly under physiological conditions[11–14] indicate that the ability of a protein sequence to fold properly depends on a delicate balance of thermodynamic and kinetic factors governing the various steps of these two processes.

In spite of its relevance, the investigation of protein aggregation at a residue level is technically challenging. The initial monomeric state from which the aggregates form is also difficult to characterise because of its heterogeneous nature (it is often an ensemble of conformations) and the difficulty to preserve the monomeric condition at the high protein concentrations required for detailed structural investigation by NMR. Radford and co-workers have recently utilised urea titration in combination with NMR to elegantly determine the level of residual structure present in the amyloidogenic state.[15] Systematic protein engineering approaches, in which single residues are substituted or deleted one by one, provide detailed information on the regions of the sequence that are involved in the rate-determining steps of aggregation.[16–19] Nevertheless, they are not able to clarify the precise mechanism of aggregation, nor the structure of the amyloidogenic monomeric state or the resulting aggregates.

Here, we use a systematic approach of limited proteolysis in combination with mass spectrometry to investigate the regions that are solvent-exposed, or flexible enough to be accessible to protein–protein interactions, in the native and amyloidogenic states of muscle acylphosphatase (AcP), a protein previously shown to form amyloid fibrils in the presence of moderate concentrations of trifluoroethanol (TFE).[4] Although this approach provides low-resolution data, it is amenable to the analysis of conformational changes in protein structure under different experimental conditions[20–22] and to investigate transient species[23] or partly folded intermediates.[24] The identification of the sites along a polypeptide chain that are most flexible and exposed to the solvent help identify the dynamic properties of the amyloidogenic state of AcP and gain insight into the key events of the early steps of aggregation.

AcP is known to aggregate and form amyloid fibrils in the presence of moderate concentrations of trifluoroetahnol (TFE).[4] This results from the ability of this co-solvent to denature the native fold of AcP without precluding the possibility of forming non-covalent interactions such as backbone hydrogen bonds.[4] Following the addition of TFE, the native protein unfolds very rapidly to form an ensemble of denatured conformations rich in α-helical structure. These convert slowly into protein aggregates, which have the appearance of ~4 nm wide granules and short protofibrils. These aggregates are able to bind to amyloid-specific dyes, such as Congo red and thioflavine T. Moreover, they are stabilised by extensive β-sheet structure as revealed by far-UV circular dichroism (CD) and Fourier transform infra-red (FT-IR) spectroscopies. Within several days the protofibrils elongate to become rigid protofilaments that ultimately assemble in a parallel manner to form rope-like fibrils[4] The object of the present investigation is the structure and dynamics of the "amyloidogenic" state of AcP, i.e. the denatured ensemble of monomeric conformations that have the ability to initiate the process of aggregation. We will show that the regions of the sequence that are most exposed to the solvent and susceptible to proteolytic attack coincide to a good approximation with those that influence the rate of aggregation.

Results

AcP was purified from an *Escherichia coli*-based expression systems as described.[25] The protein used here has the cysteine residue at position 21 replaced by a serine residue in order to avoid complexities arising from a free cysteine residue.[26] An additional Gly residue is present at the N terminus as a result of the fact that AcP is expressed as a protein fused to glutathione *S*-transferase (GST), the latter being removed subsequently with specific proteolytic cleavage.[25] Such N-terminal

Gly residue is assigned position 0, while the others are correctly assigned positions 1–98. The recombinant protein was analysed by HPLC and revealed the occurrence of one major peak and two minor ones. ESI MS analyses carried out on the major peak accounting for a fraction higher than 90% of total protein showed the presence of a single component displaying a molecular mass of 11,049.7(±0.2) Da, in excellent agreement with the theoretical mass value expected on the basis of the AcP amino acid sequence (11,049.4 Da).

The other two peaks, accounting together for a fraction less than 10% of total protein, represent variants of AcP that are either N-terminally truncated by one to three residues or oxidised at the level of a Met residue. The large prevalence of the intact protein renders the protein sample suitable for the study presented here.

Limited proteolysis on native AcP

Limited proteolysis experiments on AcP were first carried out in the absence of TFE or any other denaturant, at pH 7.5 and 25 °C. Under these conditions AcP is known to be stable and adopt its native conformation.[27] The far-UV circular dichroism spectrum acquired for AcP under these conditions (Figure 1) was very similar to that previously obtained for native AcP and typical of an α/β protein.[28] This confirms the native topology of the AcP sample used here.

Limited proteolysis experiments were performed using seven different proteases with the aim to create conditions in which the selectivity of the cleavage was not related to, or limited by, the specificity of the enzyme. For each protease, the appropriate enzyme:AcP ratio was accurately determined to generate a limited number of proteolytic events and to address protease action towards the most flexible and solvent-exposed sites. The peptide fragments released from native AcP were separated by reverse-phase HPLC and identified by ESI MS. This allowed the positions of the cleavage sites to be assigned.

As an example, Figure 2 shows the HPLC chromatograms of the aliquots withdrawn following 15 minutes and 30 minutes of trypsin digestion. AcP was immediately cleaved at Arg74, releasing the complementary fragments 0-74 (8140.4(±0.8) Da) and 75-98 (2925.8(±0.5) Da) (peaks 4 and 1 in Figure 2, respectively). Moreover, peak 2 was identified as the peptide 8-98 with a molecular mass of 10,277.8(±0.9) Da, indicating the occurrence of a second cleavage site at Lys7.

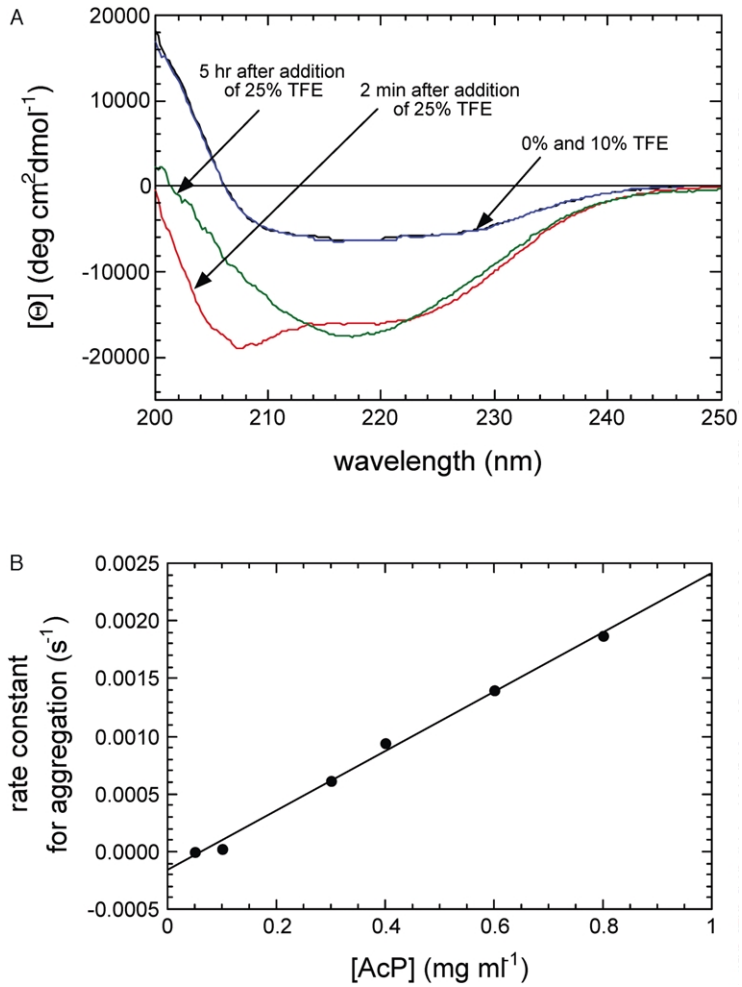

Figure 1. A, CD spectra of AcP in the presence of various concentrations of TFE. Dark, 0% TFE; blue, 10% TFE after five hours incubation; red, 25% TFE after two minutes incubation; green, 25% TFE after five hours incubation. The spectra recorded in the absence and presence of 10% TFE overlie each other and are typical of the native state of AcP. The spectrum in the presence of 25% TFE changes from that of an ensemble of denatured conformations rich with α-helical structure to that of an aggregated state containing β-sheet structure. B, Rate constant for the conversion of an ensemble of denatured and monomeric conformations into β-sheet-containing aggregates as a function of protein concentration.

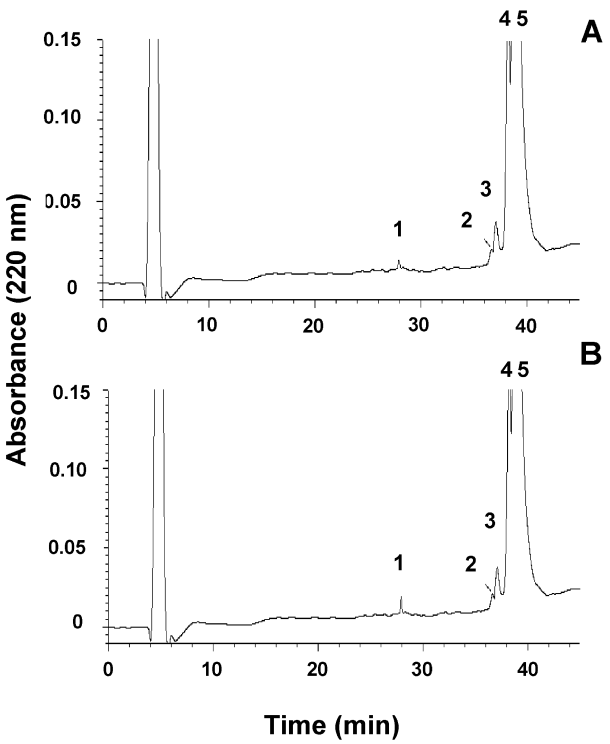

Figure 2. Limited proteolysis of native AcP with trypsin. HPLC analysis of the aliquots withdrawn following 15 minutes (A) and 30 minutes (B) of trypsin incubation of native AcP. The indicated fractions were manually collected and the eluted peptides were identified by ESI MS.

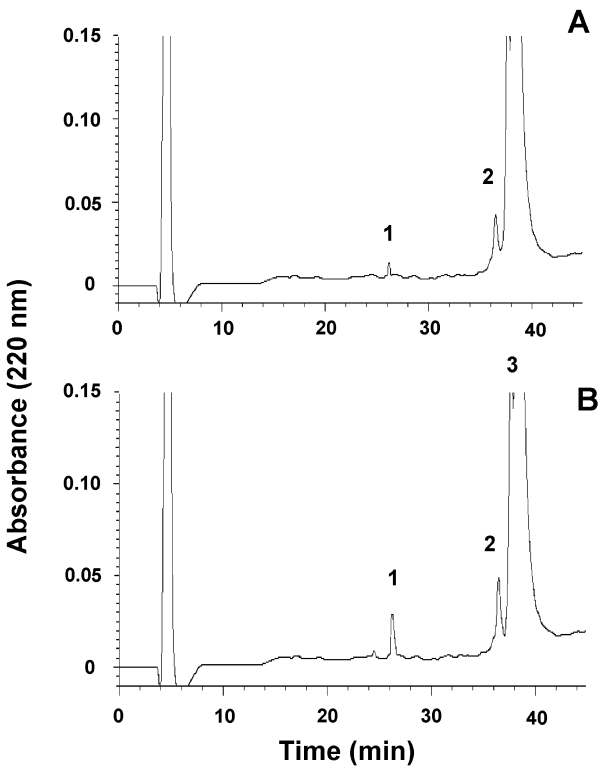

Figure 3. Limited proteolysis of native AcP with chymotrypsin. HPLC analysis of the aliquots withdrawn following 15 minutes (A) and 30 minutes (B) of chymotryptic digest of native AcP. The indicated fractions were manually collected and the eluted peptides were identified by ESI MS.

Similar results were obtained when the proteolytic experiments were carried out with the endoproteases AspN and V8. Preferential proteolytic sites were identified near the N and C termini of the protein, at Glu12, Asp76, Glu83 and Glu90, with only AspN showing a cleavage site in the central region of AcP at Asp56.

Further limited proteolysis experiments were carried out by incubating AcP with broader specificity proteases such as chymotrypsin, thermolysin, elastase and subtilisin. When chymotrypsin was used as a proteolytic agent (Figure 3), the HPLC time-course analysis showed the occurrence of the complementary peptide pair 0-80 and 81-98 (fractions 4 and 1, mass values 8887.3(±1.1) Da and 2180.1(±0.4) Da, respectively), indicating the occurrence of a single proteolytic event at Phe80.

The overall data from the limited proteolysis experiments are summarised in Figure 5A and Table 1. Preferential proteolytic sites clustered into two separate regions of the protein, the N-terminal segment encompassing residues 4–12 and another region near the C terminus, comprising β-strand 4 and part of the subsequent loop (residues 74–90). Among the proteases utilised here, only Asp56 showed accessibility in the central region of the protein. These results indicated that in native AcP these segments are particularly flexible and exposed, with the central region and a few residues immediately before the C terminus constituting a compact protein core.

Limited proteolysis on the amyloidogenic state of AcP

The same approach was employed to investigate the conformations of AcP in the presence of 10% and 25% (v/v) TFE under the same conditions of pH, temperature and AcP concentration. The enzymatic activity of the various proteases was evaluated at different TFE concentrations prior to limited proteolysis analyses by using synthetic substrates (data not shown). Preliminary proteolytic experiments showed that thermolysin, V8 protease and endoprotease AspN displayed only a very low proteolytic activity in the presence of the co-solvent and hence were not used. The proteases employed for the experiments in the presence of TFE were then trypsin, chymotrypsin, elastase and subtilisin. In order to account for the decrease of proteolytic activity observed at higher percentages of TFE the enzyme:AcP ratios were increased as compared to those used in the corresponding experiments on native AcP (see Materials and Methods for further details).

The CD spectrum of AcP acquired after five hours incubation in the presence of 10% TFE was very similar to that recorded in the absence of

co-solvent, suggesting that the protein retains a native-like topology under these conditions (Figure 1A). The results of the limited proteolysis experiments were in agreement with the CD data, as the pattern of proteolytic cleavages was identical with that obtained for the native protein in the absence of TFE, as summarised in Figure 5B and Table 1.

The native fold of AcP rapidly denatures in the presence of 25% TFE to generate, within five seconds, an ensemble of denatured, monomeric conformations presenting a relatively high content of α-helical structure.[4] At this stage no aggregates were detected using electron microscopy, including spherical, 4 nm wide, protofibrils that were observed with electron microscopy during aggregation of other systems.[29] Amyloid protofibrils and fibrils form relatively slowly from such an ensemble after some hours.[4] Accordingly, the CD spectrum recorded two minutes after addition of TFE, generates a typically α-helical spectrum (Figure 1A). However, after five hours incubation in 25% TFE, the CD spectrum of AcP is largely β in character, indicating that aggregates are present at this time (Figure 1A). The transition from α to β-structure follows apparent single-exponential kinetics.[19] The observed rate constant was determined here at different protein concentrations (Figure 1B). In the presence of 25% TFE aggregation is relatively slow ($k \leq 0.0001\ s^{-1}$) at the AcP concentration of 0.1 mg ml^{-1} used for the proteolysis experiments described below, with no substantial aggregates accumulating within the time required for enzymatic hydrolysis (15 minutes). The α-helical spectrum of the initial state rules out the presence of β-sheet containing oligomer precursors to amyloid formation. Hence, the conformational state of AcP that was studied here in the presence of 25% TFE by the limited proteolysis approach is either an ensemble of denatured monomeric conformations or an ensemble of soluble oligomers in which the various molecules are only weakly bound to each other and need major conformational rearrangements to form the β-sheet structure typical of the protofibrils. This conformational state will be referred to as the amyloidogenic state.

Figure 4 shows the HPLC chromatograms of the aliquots withdrawn after 15 minutes of trypsin (A) and chymotrypsin (B) digestions of AcP in the presence of 25% TFE. The ESI MS analysis of individual tryptic peptides (Figure 4A) identified fractions 1 and 2 as the complementary peptide pairs 0-23 (2617.2(±0.6) Da) and 24-98 (8448.5(±0.6) Da). This clearly indicates that a proteolytic event occurred at the level of Arg23. The molecular mass of fraction 3 was measured as 10,277.2(±0.9) Da, corresponding to the peptide 8-98. This confirmed Lys7 as a further preferential tryptic cleavage site also in the presence of 25% TFE.

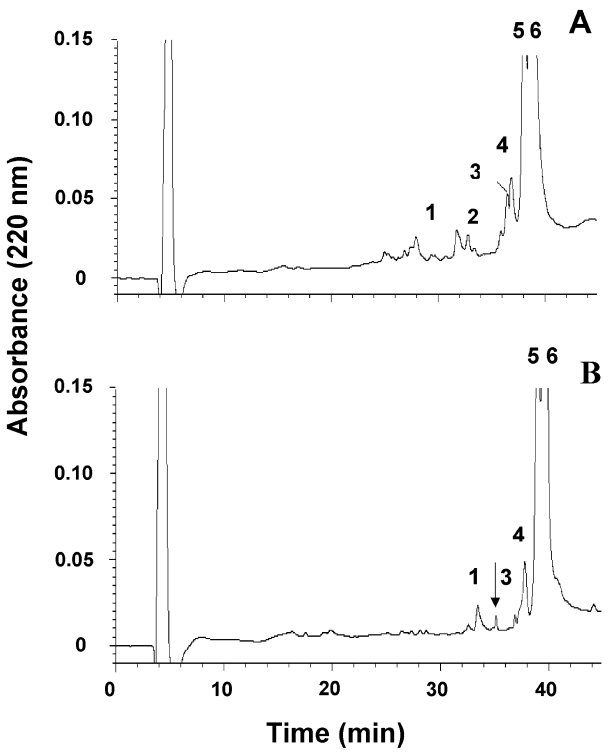

Figure 4. Limited proteolysis experiments of AcP in the amyloidogenic state. HPLC chromatograms of the aliquots withdrawn following 15 minutes of trypsin (A) and chymotrypsin (B) digestions of AcP in the presence of 25% TFE. The indicated fractions were manually collected and the eluted peptides were identified by ESI MS.

The mass spectral analysis of chymotrypsin fractions 1 and 2 (Figure 4B) revealed the occurrence of the peptide pair 23-98 (8604.7(±0.9) Da) and 0-22 (2461.1(±0.2) Da) generated by a proteolytic cleavage at Phe22. Peak 3 displayed a molecular mass of 10,405.5(±0.9) Da corresponding to the fragment 7-98 and fraction 6 showed a mass value of 10,529.9(±0.9) Da corresponding to the peptide 0-94. The presence of these fragments indicates

Table 1. Preferential hydrolysis sites obtained for the native and amyloidogenic states of AcP

TFE (%)	Trypsin	Chymotrypsin	Endoprotease V8	Endoprotease AspN	Elastase	Subtilisin	Thermolysin
0	Lys7, Arg74	Phe80	Glu12, Glu83, Glu90	Asp56, Asp76	Thr78	Leu6, N-terminal	Ser5, Asn79
10	Lys7, Arg74	Phe80	n.d.	n.d.	Thr78	Leu6, N-terminal	n.d.
25	Arg23	Phe22, Phe94, Leu6	n.d.	n.d.	Leu6, Lys7, Val20, Arg23	Leu6, Gln18, Phe94	n.d.

n.d., not determined.

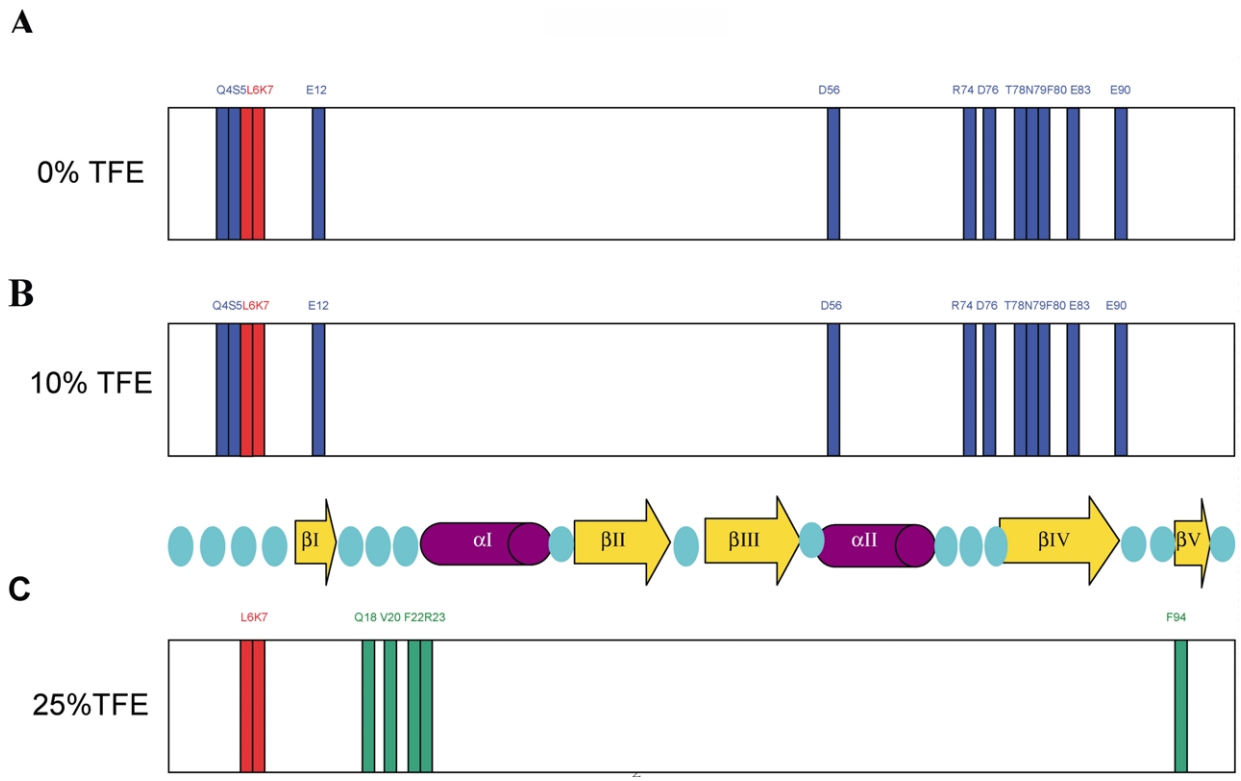

Figure 5. Location of proteolytic sites in native and amyloidogensic AcP. Schematic representation of the results obtained from limited proteolysis experiments on AcP in the absence (A) and in the presence of 10% (B) and 25% (C) TFE. Preferential proteolytic sites occurring only in the amyloidogenic state, i.e. at 25% TFE, are highlighted in green; those recognised by proteases at both 0% and 10% TFE are indicated in blue and those identified in all conditions are in red.

that the protein was preferentially cleaved at Leu6 and Phe94. Similar results were obtained with the non-specific proteases subtilisin and elastase with preferential proteolytic sites occurring at Leu6, Lys7, Gln18, Val20, Arg23 and Phe94.

The overall results of the limited proteolysis experiments carried out on AcP in 25% TFE are summarised in Figure 5C and Table 1. When these data are compared to those obtained on native AcP in the absence of TFE or in the presence of 10% co-solvent, a completely different pattern of proteolytic cleavages emerged, indicating that increasing TFE concentrations induced substantial conformational changes in the structure of AcP. In the presence of 25% co-solvent, the N-terminal region of AcP seems less flexible than the corresponding region of the native molecule and the segment encompassing residues 74–90, which was easily cleaved in the native protein, resulted in completely resistant to proteases. On the contrary, the region comprising residues 18–23 became accessible to proteolytic enzymes in the presence of TFE. Finally, a new cleavage site appeared in the C-terminal region at the level of Phe94.

Limited proteolysis in the presence of a stabilising ligand

The sensitivity of AcP to proteolytic digestion in 25% TFE was also assessed in the presence of 50 mM phosphate, a ligand that specifically binds to the active site of native AcP.[30] A ligand binding specifically to the native state of a protein acts as an inhibitor of aggregation and amyloid formation as a consequence of the stabilisation of the native state resulting from complex formation.[31,32] A phosphate concentration of 50 mM is able to stabilise the native state of AcP by ca 15 kJ mol^{-1} and to slow aggregation dramatically under these conditions of TFE concentration.[32]

The limited proteolysis experiments were performed using trypsin and chymotrypsin as proteolytic probes under the same conditions described above, i.e. 0.1 mg ml^{-1} AcP, pH 7.5, 25 °C. Native AcP was incubated in 50 mM phosphate buffer for 15 minutes prior to protease addition and the extent of proteolysis was monitored on a time-course basis. No protein fragments were released from AcP using the usual concentrations of either protease, demonstrating that the ligand exerted a very strong protection to the conformational changes that lead to partial unfolding and then to aggregation. When the protease:AcP ratio was increased by ten times, a distribution of preferential proteolytic sites similar to that found on native AcP in the absence of TFE was obtained (data not shown).

Discussion

AcP is a protein able to aggregate into amyloid fibrils morphologically related to those associated with a number of human diseases.[4] The possibility of inducing the process of fibrillogenesis *in vitro* under controlled conditions, i.e. in the presence of moderate concentrations of TFE, renders AcP a suitable model system for investigating the fundamental aspects of protein aggregation. The object of the present study is the "amyloidogenic" state, i.e. the ensemble of denatured and monomeric conformations that is produced immediately after unfolding of native AcP in TFE prior to the formation of any stable protein aggregates. Different conformational states of AcP, such as the native, ligand-bound and TFE-denatured states, were probed by limited proteolysis combined with mass spectrometric methodologies. The overall strategy is based on the evidence that amino acid residues located within exposed and flexible regions of the protein can be recognised by proteases, leading to a reasonably good imprinting of the AcP conformation in solution.[33] Since the surface topology of the protein is affected by conformational changes, when comparative experiments were carried out on AcP either at different TFE concentrations or in the presence of phosphate, differential peptide maps were obtained from which protein regions with preferential flexibility or exposure to the solvent could be inferred.

In the absence of denaturant or in the presence of 10% TFE, AcP adopts a native conformation. The same distribution of proteolytic sites was consequently obtained under these conditions. It is worth noticing, however, that in the presence of 10% TFE, AcP appeared more susceptible to proteolytic digestion, even if the cleavage site pattern was identical. This indicates that a greater flexibility is gained upon the addition of small amounts of TFE. In both cases, the limited proteolysis experiments revealed that the N-terminal region (residues 1–12) and a region that encompasses β-strand IV and part of the subsequent loop (residues 74–90) are accessible to the proteases as indicated in Figure 5. These segments correspond to regions that are rather mobile in the native conformation[27] and are totally unstructured in the transition state for folding and unfolding.[34] This suggests that the unfolding process of AcP is initiated by the local unfolding of these portions of the protein. These events of local unfolding cause the remainder of the protein to be destabilised and are therefore likely to trigger the cooperative unfolding of the entire protein.

The results of limited proteolysis performed in the presence of 25% TFE provide new insight into the structure and dynamics of the amyloidogenic state and the mechanism of aggregation of AcP. Denaturation of AcP in the presence of 25% TFE uncovers a number of proteolytic sites along the polypeptide chain. These are located within the N-terminal residues 6-7 and the segment 18-23. A proteolytic site at positions 94, within the β-strand V at the C terminus, was also probed by two distinct proteases. All other regions do not appear to be susceptible to proteolytic attack. Preferential proteolysis at these sites does not result from the specificity of the proteases employed here. Trypsin, for example, is known to hydrolyse peptide bonds preceding residues of lysine or arginine. AcP possesses six arginine residues and nine lysine residues. These do not cluster within one or a few regions of the sequence, but are distributed rather uniformly along the sequence. Preferential cleavage of this enzyme at Arg23 is a consequence of the exposure of this residue as compared to others. Similarly, chymotrypsin preferentially cleaves peptide bonds preceding hydrophobic and aromatic residues that are highly represented along the sequence of AcP. Elastase and subtilisin have even broader specificity.

A previous study had indicated the two regions of the sequence 16–31 and 87–98 as the key sites for promoting the aggregation process of AcP from the TFE-denatured state.[19] Indeed, the aggregation rate of AcP is highly sensitive to single-point mutations within these regions. In addition, synthetic peptides corresponding to these two stretches of the AcP sequence are highly prone to aggregate.[19] Interestingly, proteolytic sites are present within these regions in the amyloidogenic state (Figure 5). This suggests that the importance of these two regions in the aggregation process arises from their solvent-exposure and flexibility, which render them accessible for intermolecular interactions. The N-terminal region is also solvent-exposed and vulnerable to proteolytic attack (Figure 5). Nevertheless, the small values of hydrophobicity and β-sheet propensity render this region insignificant in the aggregation process as revealed by the full solubility of the peptide corresponding to this region and the unchanged aggregation rate of AcP when this region is mutated.[19]

It is particularly interesting to note that no proteolytic sites have been identified in the amyloidogenic state of AcP within the region corresponding to β-hairpin 34-53. This region was shown by the mutational study to be very important in the folding process of AcP but not in that of aggregation.[19,34] The lack of importance of this sequence region in the aggregation process, which seemingly contrasted with its high hydrophobicity and β-sheet propensity, finds an explanation in the finding reported here that this portion of the AcP sequence is relatively shielded from the solvent and therefore inaccessible for establishing the first intermolecular contacts needed for aggregation.

Limited proteolysis experiments performed in the presence of 50 mM phosphate, a ligand binding specifically to the active site of native AcP,[30] showed an almost absolute resistance of AcP to the proteolytic attack even when the experiments were performed in 25% TFE. When the protease: AcP ratio was increased by ten times, a distribution

of preferential proteolytic sites similar to that found in the absence or in the presence of 10% TFE was obtained. Since the phosphate ion binds specifically to the active site of native AcP, the native state of AcP is stabilised in its presence to an extent sufficient to be largely populated even in 25% TFE.[36] The absence of proteolytic events within the region 18-23 and at position 94 indicates that the phosphate-induced folding of AcP favours the partial shield of regions critical to aggregation from the solvent. This result experimentally supports previous suggestions that addition of ligands binding specifically to the native state of a protein is a strategy of general value to prevent amyloid formation.[31,32]

Many of the proteolytic sites detected in the amyloidogenic state of AcP are located in a region of the sequence encompassing residues 18–23 (Gln18-Gly19-Val20-Cys21-Phe22-Arg23). Proteolysis at these sites is by itself an indication that these residues of AcP have a relatively high propen-sity to interact with other protein molecules. In addition to its relatively high solvent exposure and/or flexibility, this continuous stretch of amino acid residues is highly hydrophobic and possesses a relatively high propensity to form β-sheet structure.[19] Molecular chaperones of the Hsp70 family, ubiquitous proteins that bind to polypeptide chains during biosynthesis and cellular stress to prevent their aggregation, have been shown to bind to unfolded stretches of the target proteins, through hydrophobic side-chain interactions and hydrogen bonds with the peptide backbone.[35] Hsp70s recognise exposed hydrophobic side-chains, in conjunction with an accessible polypeptide backbone.[35] These features are all present in the 18-23 region of AcP in the amyloidogenic state. Moreover, residues 19–23 consist of the typical consensus motif of Hsp70s, that is a hydrophobic core of four to five residues flanked by basic residues.[37] This suggests that chaperones normally present in the cell can effectively inhibit aggregation of AcP by binding to the regions of the sequence that play a key role in the aggregation of this protein.

In conclusion, regions of the sequence previously recognised from a protein engineering approach to promote the process of aggregation of AcP appear to be solvent-exposed and/or flexibile enough to be accessible to intermolecular interactions. The observation that such regions may be preferential sites for the interaction with molecular chaperones of the Hsp70 family allows the mechanistic description of the aggregation of AcP to be related to the well-known binding properties of chaperones.

Materials and Methods

Materials

Trypsin, chymotrypsin, thermolysin, elastase and subtilisin were purchased from Sigma. V8 endoprotease and AspN endoprotease are from Roche. TFE and Tris–HCl are from Sigma. All other chemicals and HPLC solvents are from Baker.

Limited proteolysis experiments

Limited proteolysis experiments were carried out by incubating AcP with trypsin, chymotrypsin, V8 endoprotease, AspN endoprotease, thermolysin, elastase and subtilisin. Enzymatic digestions were all performed in 100 mM Tris–HCl (pH 7.5) at 25 °C using enzyme-to-substrate ratios ranging between 1:70 and 1:200 (w/w). Enzymatic digestions at different concentrations of TFE were carried out under the same conditions in the presence of 10% and 25% TFE using enzyme-to-substrate ratios ranging from 1:10 to 1:380 (w/w). The extent of the reaction was monitored on a time-course basis by sampling the incubation mixture at different time intervals. Proteolytic fragments were fractionated by reverse-phase HPLC on a Phenomenex Jupiter C18 column and the individual fractions were manually collected and analysed by electrospray ionisation mass spectrometry (ESI MS).

Chromatographic separation of peptides

Peptide mixtures from the different proteolysis experiments were fractionated by reverse-phase HPLC on a Phenomenex Jupiter C18 column (250 mm × 2.1 mm, 300 Å pore size) with a multistep gradient from 10% to 30% acetonitrile/isopropanol (2:1, v/v), in 0.07% trifluoroacetic acid (TFA) over 12 minutes and from 30% to 50% in 32 minutes. The flow rate was kept at 200 μl/minute. Elution was monitored at 220 nm. Fractions were collected and identified by ESI MS.

Mass spectrometry

Proteolytic fragments were analysed by ESI MS on an API-100 single quadrupole instrument (Perkin Elmer). Samples were injected directly into the ion source by a Harward pump syringe at a flow rate of 3 μl/minute. Data were acquired and processed using Biomultiviewer (Perkin Elmer) software. Horse heart myoglobin was used to calibrate the instrument (average molecular mass 16,951.5 Da); all masses are reported as average masses.

Circular dichroism

AcP was incubated at a concentration of 0.4 mg ml^{-1} in the presence of 0%, 10% and 25% (v/v) TFE, acetate buffer, pH 5.5, 25 °C. After five hours incubation under these conditions CD spectra were acquired by means of a Jasco J-720 spectropolarimeter and cuvettes of 1 mm path length. The spectrum immediately after addition of 25% TFE was also acquired.

Acknowledgements

This work was supported by MIUR (PRIN 2002, FIRB 2002) and Regione Campania (L5/2002).

References

1. Dobson, C. M. (1999). Protein misfolding, evolution and disease. *Trends Biochem. Sci.* **24**, 329–332.
2. Guijarro, J. I., Sunde, M., Jones, J. A., Campbell, I. D. & Dobson, C. M. (1998). Amyloid fibril formation by an SH3 domain. *Proc. Natl Acad. Sci. USA*, **95**, 4224–4228.
3. Litvinovich, S. V., Brew, S. A., Aota, S., Akiyama, S. K., Haudenschild, C. & Ingham, K. C. (1998). Formation of amyloid-like fibrils by self-association of a partially unfolded fibronectin type III module. *J. Mol. Biol.* **280**, 245–258.
4. Chiti, F., Webster, P., Taddei, N., Clark, A., Stefani, M., Ramponi, G. & Dobson, C. M. (1999). Designing conditions for *in vitro* formation of amyloid protofilaments and fibrils. *Proc. Natl Acad. Sci. USA*, **96**, 3590–3594.
5. Konno, T., Murata, K. & Nagayama, K. (1999). Amyloid-like aggregates of a plant protein: a case of a sweet-tasting protein, monellin. *FEBS Letters*, **454**, 122–126.
6. Kallberg, Y., Gustafsson, M., Persson, B., Thyberg, J. & Johansson, J. (2001). Prediction of amyloid fibril-forming proteins. *J. Biol. Chem.* **276**, 12945–12950.
7. Yutani, K., Takayama, G., Goda, S., Yamagata, Y., Maki, S., Namba, K. *et al.* (2000). The process of amyloid-like fibril formation by methionine aminopeptidase from a hyperthermophile, *Pyrococcus furiosus*. *Biochemistry*, **39**, 2769–2777.
8. Ramirez-Alvarado, M., Merkel, J. S. & Regan, L. (2000). A systematic exploration of the influence of the protein stability on amyloid fibril formation *in vitro*. *Proc. Natl Acad. Sci. USA*, **97**, 8979–8984.
9. Kallijarvi, J., Haltia, M. & Baumann, M. H. (2001). Amphoterin includes a sequence motif which is homologous to the Alzheimer's beta-amyloid peptide (Abeta), forms amyloid fibrils *in vitro*, and binds avidly to Abeta. *Biochemistry*, **40**, 10032–10037.
10. Bucciantini, M., Giannoni, E., Chiti, F., Baroni, F., Formigli, L., Zurdo, J. *et al.* (2002). Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature*, **416**, 507–511.
11. Eliezer, D., Chiba, K., Tsuruta, H., Doniach, S., Hodgson, K. O. & Kihara, H. (1993). Evidence of an associative intermediate on the myoglobin refolding pathway. *Biophys. J.* **65**, 912–917.
12. Silow, M. & Oliveberg, M. (1997). High-energy channeling in protein folding. *Biochemistry*, **36**, 7633–7637.
13. Silow, M., Tan, Y. J., Fersht, A. R. & Oliveberg, M. (1999). Formation of short-lived protein aggregates directly from the coil in two-state folding. *Biochemistry*, **38**, 13006–13012.
14. van den Berg, B., Ellis, R. J. & Dobson, C. M. (1999). Effects of macromolecular crowding on protein folding and aggregation. *EMBO J.* **18**, 6927–6933.
15. McParland, V. J., Kalverda, A. P., Homans, S. W. & Radford, S. E. (2002). Structural properties of an amyloid precursor of β2-microglobulin. *Nature Struct. Biol.* **9**, 326–331.
16. Wood, S. J., Wetzel, R., Martin, J. D. & Hurle, M. R. (1995). Prolines and amyloidogenicity in fragments of the Alzheimer's peptide beta/A4. *Biochemistry*, **34**, 724–730.
17. Moriarty, D. F. & Raleigh, D. P. (1999). Effects of sequential proline substitutions on amyloid formation by human amylin 20-29. *Biochemistry*, **38**, 1811–1818.
18. Tjernberg, L. O., Callaway, D. J., Tjernberg, A., Hahne, S., Lilliehook, C., Terenius, L. *et al.* (1999). A molecular model of Alzheimer amyloid beta-peptide fibril formation. *J. Biol. Chem.* **274**, 12619–12625.
19. Chiti, F., Taddei, N., Baroni, F., Capanni, C., Stefani, M., Ramponi, G. & Dobson, C. M. (2002). Kinetic partitioning of protein folding and aggregation. *Nature Struct. Biol.* **9**, 137–143.
20. Orrù, S., Dal Piaz, F., Casbarra, A., Biasiol, G., De Francesco, R., Steinkuhler, C. & Pucci, P. (1999). Conformational changes in the NS3 protease from hepatitis C virus strain BK monitored by limited proteolysis and mass spectrometry. *Protein Sci.* **8**, 1445–1454.
21. Bianchi, E., Orrù, S., Dal Piaz, F., Ingenito, R., Casbarra, A., Biasiol, G. *et al.* (1999). Conformational changes in human hepatitis virus NS3 protease upon binding of product-based inhibitors. *Biochemistry*, **38**, 13844–13852.
22. Piccoli, R., De Lorenzo, C., Dal Piaz, F., Pucci, P. & D'Alessio, G. (2000). Trypsin sheds light on the singular case of seminal rnase, a dimer with two quaternary conformations. *J. Biol. Chem.* **275**, 8000–8006.
23. Esposito, G., Michelutti, R., Verdone, G., Viglino, P., Hernandez, H., Robinson, C. V. *et al.* (2000). Removal of the N-terminal hexapeptide from human beta2-microglobulin facilitates protein aggregation and fibril formation. *Protein Sci.* **9**, 831–845.
24. Birolo, L., Dal Piaz, F., Pucci, P. & Marino, G. (2002). Structural characterization of the M* partly folded intermediate of wild-type and P138A EcAspAT. *J. Biol. Chem.* **277**, 17428–17437.
25. Modesti, A., Taddei, N., Bucciantini, M., Stefani, M., Colombini, B., Raugei, G. & Ramponi, G. (1995). Expression, purification, and characterization of acylphosphatase muscular isoenzyme as fusion protein with glutathione *S*-transferase. *Protein Expr. Purif.* **6**, 799–805.
26. van Nuland, N. A., Chiti, F., Taddei, N., Raugei, G., Ramponi, G. & Dobson, C. M. (1998). Slow folding of muscle acylphosphatase in the absence of intermediates. *J. Mol. Biol.* **283**, 883–891.
27. Pastore, A., Saudek, V., Ramponi, G. & Williams, R. J. (1992). Three-dimensional structure of acylphosphatase. Refinement and structure analysis. *J. Mol. Biol.* **224**, 427–440.
28. Chiti, F., Taddei, N., Webster, P., Hamada, D., Fiaschi, T., Ramponi, G. & Dobson, C. M. (1999). Acceleration of the folding of acylphosphatase by stabilization of local secondary structure. *Nature Struct. Biol.* **6**, 380–387.
29. Lashuel, H. A., Hartley, D., Petre, B. M., Walz, T. & Lansbury, P. T., Jr (2002). Neurodegenerative disease: amyloid pores from pathogenic mutations. *Nature*, **418**, 291.
30. Stefani, M., Taddei, N. & Ramponi, G. (1997). Insights into acylphosphatase structure and catalytic mechanism. *Cell. Mol. Life Sci.* **53**, 141–151.
31. Baures, P. W., Peterson, S. A. & Kelly, J. W. (1998). Discovering transthyretin amyloid fibril inhibitors by limited screening. *Bioorg. Med. Chem.* **6**, 1389–1401.
32. Chiti, F., Taddei, N., Stefani, M., Dobson, C. M. & Ramponi, G. (2001). Reduction of the amyloidogenicity of a protein by specific binding of ligands to the native conformation. *Protein Sci.* **10**, 879–886.
33. Zappacosta, F., Pessi, A., Bianchi, E., Venturini, S., Sollazzo, M., Tramontano, A. *et al.* (1996). Probing the tertiary structure of proteins by limited

proteolysis and mass spectrometry: the case of Minibody. *Protein Sci.* **5**, 802–813.
34. Chiti, F., Taddei, N., White, P. M., Bucciantini, M., Magherini, F., Stefani, M. & Dobson, C. M. (1999). Mutational analysis of acylphosphatase suggests the importance of topology and contact order in protein folding. *Nature Struct. Biol.* **6**, 1005–1009.
35. Hartl, F. U. & Hayer-Hartl, M. (2002). Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science*, **295**, 1852–1858.
36. Rudiger, S., Germeroth, L., Schneider-Mergener, J. & Bukau, B. (1997). Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. *EMBO J.* **16**, 1501–1507.

Edited by P. T. Lansbury Jr

(Received 24 February 2003; accepted 2 October 2003)