UNIVERSITA' DEGLI STUDI DI FIRENZE

DIPARTIMENTO DI SCIENZE BIOCHIMICHE

DOTTORATO DI RICERCA IN BIOCHIMICA E BIOLOGIA APPLICATA

XXI CICLO

Settore disciplinare BIO 10

Studies on the native-like aggregation of acylphosphatases

Coordinatore: Prof. Gianni Cappugi

Docente guida: Prof. Fabrizio Chiti

Candidato: Gemma Soldi

To my dearest ones

(who knows, knows).

"No man is an island entire of itself; every man is a piece of the continent, a part of the main; if a clod be washed away by the sea, Europe is the less, as well as if a promontory were, as well as any manner of thy friends or of thine own were; any man's death diminishes me, because I am involved in mankind. And therefore never send to know for whom the bell tolls; it tolls for thee."

John Donne

...Serendipità è dunque - filosoficamente - lo scoprire una cosa non cercata e imprevista mentre se ne sta cercando un'altra. Ma il termine non indica solo fortuna: per cogliere l'indizio che porterà alla scoperta occorre essere aperti alla ricerca e attenti a riconoscere il valore di esperienze che non corrispondono alle originarie aspettative....

"With dreams to be a king, first one should be a man
I call about and charge them all with a life that is a lie"
Manowar, "Warriors of the world"

List of abbreviations

AcP	Acylphosphatase
AcPDro2	Acylphosphatase 2 from *D. Melanogaster*
AFM	Atomic Force Microscopy
ANS	1-anilino-8-naphtalene-sulfonic acid
CD	Circular Dichroism
CR	Congo Red
DLS	Dynamic Light Scattering
DTT	Dithiothreitol
EDTA	Ethylen-diamino-tetraacetic acid
Far-UV CD	Far Ultraviolet Circular Dichroism
GdnHCl	Guanidine Hydrochloride
GST	Glutahione-S-Transferase
HypF-N	N-terminal domain of HypF protein from *E. Coli*
IPTG	Isopropyl-β-D-1-Thiogalactopyranoside
mtAcP	Human muscular type Acylphosphatase
Near-UV CD	Near Ultraviolet Circular Dichroism
PCR	Polymerase Chain Reaction
SOD-1	Superoxide Dismutase type 1
SS-NMR	Solid State Nuclear Magnetic Resonance Spectroscopy
Sso AcP	Acylphosphatase from *S. Solfataricus*
TEM	Transmission Electron Microscopy
TFA	Trifluoroacetic acid
TFE	Trifluoroethanol
ThT	Thioflavin T
TM-AFM	Tapping Mode Atomic Force Microscopy
TTR	Transthyretin

Table of Contents

CHAPTER I – INTRODUCTION

1.1 Protein deposition diseases: a brief overview

1.1.1 The protein misfolding/misfunction problem

One of the most challenging issues in modern molecular medicine is probably to gain a full understanding of the equilibria between fully folded and functional structures of proteins and potentially dangerous states, including partially or fully unfolded states as well as native-like states, that may lead to formation of misfolded structures ultimately causing disease. Each protein attains a unique correct and functional fold, that is the so-called "native state". The native state is usually reached shortly after translation, often with the co-operation of other proteins with chaperone functions. It is widely accepted that the bare amino acidic sequence of a protein contains the essential information to acquire the definitive three-dimensional structure. It is also well known that the native state is, in most cases, the thermodynamically most stable of all the conformations that the polypeptide chain can theoretically adopt (1,2). On the other hand, it is obvious that the number of possible conformations is so huge that a search of the correct one, simply by making attempts, would take absolutely too much time. In past years, it has been clarified that the folding process does not involve a linear series of mandatory steps between specific partly folded states, but rather a stochastic search of the many conformations

accessible to a polypeptide chain in a range of possibilities known as "energy landscape" (Fig. 1.1) (3-5). Considering that native-like interactions are usually more stable than non-native ones, the folding process is very likely to be seen as a series of progressive "guided" attempts, rather than a random sampling of possibilities.

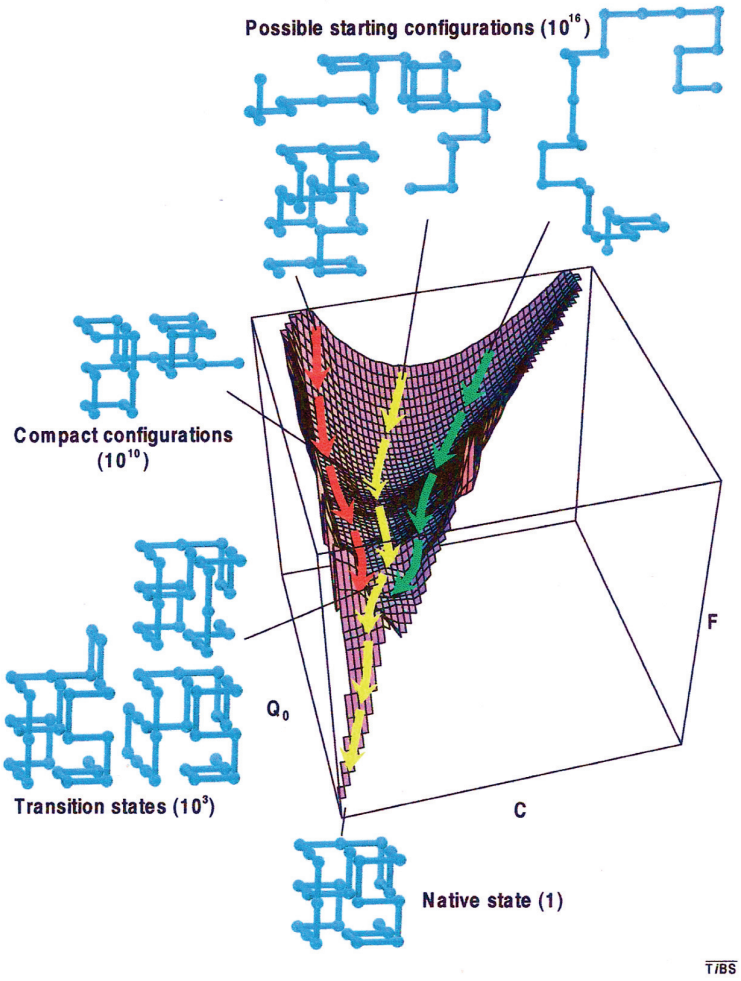


Figure 1.1 Free-energy (F) surface of a 27-mer model protein as a function of the number of native contacts (Q_0) and the total number of (native and non-native) contacts (C) obtained by sampling the accessible configuration space using Monte Carlo simulations. The yellow trajectory shows the average path traced by structures in 1000 independent trials that each began in a different random conformation. The other two trajectories (green and red) show a range of two standard deviations around the average and are thus expected to include ca. 95% of the trajectories. The structures illustrate the various stages of the reaction. From one of the 10^{16} possible random starting conformations, a folding chain collapses rapidly to a disorderd globule It then makes a slow, non-directed search among the 10^{10} semi-compact conformations for one of the approximately 10^{3} transition states that lead rapidly to the unique native state. (Figure and caption from(6))

Both prokaryotic and eukaryotic cells have developed a very efficient cellular machinery , both at translational and post-translational level, that can monitor the protein folding process from the ribosome to the fully functional native state (7-10). If the correct folding process fails, the resulting polypeptide chain can be conformationally re-modeled by specific chaperones in an attempt to give it the correct conformation, or be just ubiquitinated and degraded back to amino acids. Sometimes the complex machine that control protein folding homeostasis is poorly functional or even impaired, due to unfavourable environmental conditions like pH, temperature (heat shock), oxidative stress, aging or it can be damaged by chemical interactions or gene mutations, or even uncorrect folding/assembly of the chaperones itself. In these cases the misfolded polypeptides/proteins are not eliminated and remain in the cell or in the extracellular space, where they can pack together and accumulate giving rise to protein assemblies (11,12). If there is a massive degradation of non-functional forms of a protein, this can lead to a significant reduction of the physiological level of the native state and then to a lack of the associated metabolic function; a typical example is cystic fibrosis (13,14). In other cases, an uncorrect folding causes an improper binding and trafficking of the polypeptide, as in familial hypercholesterolemia (15) or early-onset emphysema (16). In a great number of cases the misfolded proteins can self-assemble into insoluble fibrillar structures that form deposits in the cytosol or, more commonly, in the extracellular space, giving rise to a variety of pathological conditions well known as "protein deposition diseases". These pathological conditions include some of the most invalidant syndromes of our century, such

as Parkinson's disease, type-II diabetes, spongiform encephalopaties and systemic amyloidosis.

1.1.2 Protein deposition diseases arising from amyloid aggregation

There is a large subset of conformational diseases that show, as a common feature, the presence of non-functional and highly ordered fibrillar aggregates with amyloid or amyloid-like structure. This term was given after the observation that these fibrils, at first sight, looked like starch crystals, but it is well known today that they possess a well-defined structure and morphology (see section 1.2). These structures usually originate from proteins that are normally soluble and functional, either as a fully folded state or as intrinsically disordered proteins (17). Although the term "amyloid" is generally used to refer to extracellular deposits, it is increasingly recognized that even intracellular proteinaceous inclusions share the same structural and morphological features (18,19). A basical overview of protein deposition diseases should consider three fundamental categories: neuropathic conditions, non-neuropathic localized amyloidoses, and non-neuropathic sistemic amyloidoses. In the first group of pathologies, aggregation occurs in the central nervous system; in the other types of amyloidoses, protein deposition occurs in a single type of tissue other than the brain, or in multiple tissues in the whole body, respectively (17). This wide range of pathologies includes both sporadic and hereditary forms, the latter arising from the presence of specific mutations, often geographically localized (e.g. British and Danish dementia, Finnish hereditary amyloidosis, etc…). In addition, there are also some forms of spongiform encephalopathies that have been demonstrated to be transmissible.

Other diseases arise from medical treatment, such as dialysis-related amyloidosis and injection-localized amyloidosis. Table 1.1 lists some of the most widely studied protein deposition diseases associated with amyloid or amyloid-like structure (17).

Table 1.1 Human diseases associated with formation of extracellular amyloid deposits or intracellular inclusions with amyloid-like characteristics (17).

Disease	Aggregating protein or peptide	Number of residues of protein or peptide[a]	Native Structure of protein or peptide[b]
Neurodegenerative diseases			
Alzheimer's disease[c]	Amyloid β peptide	40 or 42[f]	Natively unfolded
Spongiform encephalopathies[c,e]	Prion protein or fragments thereof	253	Natively unfolded (residues 1-120) and α-helical (121-230)
Parkinson's disease[c]	α-Synuclein	140	Natively unfolded
Dementia with Lewy bodies[c]	α-Synuclein	140	Natively unfolded
Frontotemporal dementia with Parkinsonism[c]	Tau	352-441[f]	Natively unfolded
Amyotrophic lateral sclerosis[c]	Superoxide dismutase 1	153	All-β, Ig like
Huntington's disease[d]	Huntingtin with polyQ expansion	3144[g]	Largely natively unfolded
Spinocerebellar ataxias[d]	Ataxins with polyQ expansion	816[g,h]	All-β, AXH domain (residues 562-694); the rest are unknown
Spinocerebellar ataxia 17[d]	TATA box-binding protein with polyQ expansion	339[g]	α+β, TBP like (residues 159-339); unknown (1-158)
Spinal and bulbar muscular atrophy[d]	Androgen receptor with polyQ expansion	919[g]	All-α, nuclear receptor ligand-binding domain (669-919); the rest are unknown
Hereditary dentatorubral-pallidoluysian atrophy[d]	Atrophin-1 with polyQ expansion	1185[g]	Unknown

Familial British dementia[d]	ABri	23	Natively unfolded
Familial Danish dementia[d]	ADan	23	Natively unfolded
Nonneuropathic systemic amyloidoses			
AL amyloidosis[c]	Immunoglobulin light chains or fragments	~ 90[f]	All-β, IG like
AA amyloidosis[c]	Fragments of serum amyloid A protein	76-104[f]	All-α, unknown fold
Familial Mediterranean fever[c]	Fragments of serum amyloid A protein	76-104[f]	All-α, unknown fold
Senile systemic amyloidosis[c]	Wild-type transthyretin	127	All-β, prealbumin like
Familial amyloidotic polyneuropathy[d]	Mutants of transthyretin	127	All-β, prealbumin like
Hemodialysis-related amyloidosis[c]	β2-microglobulin	99	All-β, Ig like
ApoAI amyloidosis[d]	N-terminal fragments of apolipoprotein AI	80-93[f]	Natively unfolded
ApoAII amyloidosis[d]	N-terminal fragment of apolipoprotein AII	98[i]	Unknown
ApoAIV amyloidosis[c]	N-terminal fragment of apolipoprotein AIV	~ 70	Unknown
Finnish hereditary amyloidosis[d]	Fragments of gelsolin mutants	71	Natively unfolded
Lysozyme amyloidosis[d]	Mutants of lysozyme	130	α+β, lysozyme fold
Fibrinogen amyloidosis[d]	Variants of fibrinogen α-chain	27-81[f]	Unknown
Icelandic hereditary cerebral amyloid angiopathy[d]	Mutant of cystatin C	120	α+β, cystatin like
Nonneuropathic localized diseases			
Type II diabetes[c]	Amylin, also called islet amyloid polypeptide (IAPP)	37	Natively unfolded
Medullary carcinoma of the thyroid[c]	Calcitonin	32	Natively unfolded
Atrial amyloidosis[c]	Atrial natriuretic factor	28	Natively unfolded
Hereditary cerebral haemorrhage with amyloidosis[d]	Mutants of amyloid β peptide	40 or 42[f]	Natively unfolded
Pituitary prolactinoma	Prolactin	199	All-α, 4-helical cytokines

Injection-localized amyloidosis[c]	Insulin	21 + 30[j]	All-α, insulin like
Aortic medial amyloidosis[c]	Medin	50[k]	Unknown
Hereditary lattice corneal dystrophy[d]	Mainly C-terminal fragments of kerato-epithelin	50-200[f]	Unknown
Corneal amylodosis associated with trichiasis[c]	Lactoferrin	692	α+β, periplasmic-binding protein like II
Cataract[c]	γ-Crystallins	Variable	All-β, γ-crystallin like
Calcifying epithelial odontogenic tumors[c]	Unknown	~ 46	Unknown
Pulmonary alveolar proteinosis[d]	Lung surfactant protein C	35	Unknown
Inclusion-body myositis[c]	Amyloid β peptide	40 or 42[f]	Natively unfolded
Cutaneous lichen amyloidosis[c]	Keratins	Variable	Unknown

[a]Data refer to the number of residues of the processed polypeptide chains that deposit into aggregates, not of the precursor proteins.

[b]According to Structural Classification Of Proteins (SCOP), these are the structural class and fold of the native states of the processed peptides or proteins that deposit into aggregates prior to aggregation.

[c]Predominantly sporadic, although in some cases hereditary forms associated with specific mutations are well documented.

[d]Predominantly hereditary although in some cases sporadic forms are documented.

[e] 5% of the cases are infectious (iatrogenic).

[f]Fragments of various lengths are generated and have been reported to be present in *ex vivo* fibrils.

[g]Lengths shown refer to the normal sequences with nonpathogenic traits of polyQ.

[h]Length shown is for ataxin-1.

[i]The pathogenic mutation converts the stop codon into a Gly codon, extending the 77-residue protein by 21 additional residues.

[j]Human insulin consists of two chains (A and B with 21 and 30 residues, respectively) covalently linked by disulfide bridges.

[k]Medin is the 245-294 fragment of human lactadherin.

This correlation between amyloid fibril formation and such a wide range of socially relevant and often incurable pathologies has lead to a great interest in unraveling the mechanisms of aggregation processes, both from a biomedical and a biophysical point of view. Evidence is now emerging that amyloid

aggregation is in fact one of the fundamental structural states of polypeptide chains that arises from very basic physical principles (4,20). In fact, it is now evident that amyloidogenic proteins do not share any obvious sequence or structural similarity in their native states, and they are also very different in secondary structure composition. Moreover, a large variety of proteins that are not associated with disease, even syntethic polymers, may are able to form supramolecular assemblies similar to amyloids, thus suggesting that these aggregation phenomena share common grounds (17,21).

1.2 Amyloid fibril structure

Amyloid fibrils are supramolecular structures whose main components are twisted rods called protofilaments, that are made of protein monomers stacked one on another. These protofilaments are are usually rich in β-sheet structure, and measure about 2-5 nm in diameter (22). In the fibril, the β-strands are usually perpendicular to the main axis (23). The protofilaments assemble as fibrils by generally twisting like an helix into rope-like structures, about 7-13 nm wide (22,23).

Classical and wide used techniques to image amyloid fibrils were microscopies such as Transmission Electron Microscopy (TEM) or Atomic Force Microscopy (AFM), and X-ray fiber diffraction (22-24).

Amyloid fibrils possess three main hallmarks that are present independently of the precursor protein that form them:

(i) A well-defined fibrillar morphology observable with TEM or AFM (Fig. 1.2) (23,25-27);

(ii) The presence of a cross-β structure observed by X-ray fibre diffraction (Fig. 1.3) (25)

(iii) A green birefringence observed under cross-polarized light after Congo Red (CR) staining (Fig. 1.4) (28).

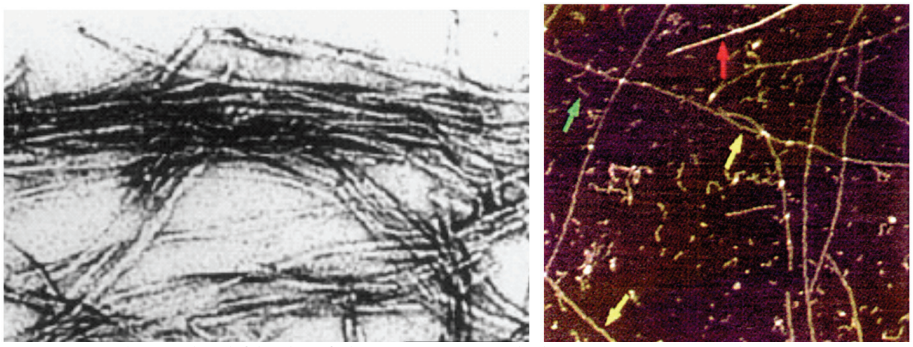


Figure 1.2 a) amyloid fibrils of the $A\beta_{1-40}$ peptide by transmission electron microscopy (26); b) amyloid fibrils of $A\beta_{1-42}$ peptide by atomic force microscopy (27).

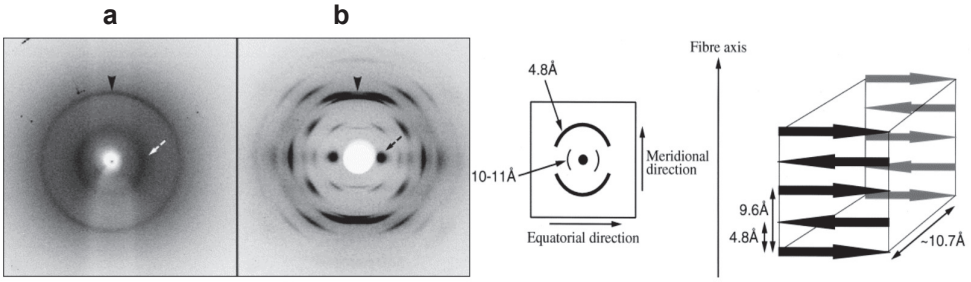


Figure 1.3 (left) X-ray fibre diffraction patterns of (a) a partially aligned sample of Aβ1-40 amyloid ¢brils and (b) a magnetically aligned sample of Aβ1-40 amyloid fibrils. (right) the characteristic cross-L spacings from X-ray ¢bre di¡raction from amyloid ¢brils. A strong 4.8 Ä reflection on the meridian corresponds to the hydrogen bonding distance between L-strands (shown right), and a more diffuse 10-11 Ä reflection on the equator shows the intersheet distance of about 10.7 Ä. A spacing of 9.6 Ä would correspond to the repeat distance for an anti-parallel arrangement of L-strands. Figure and captions from (25)

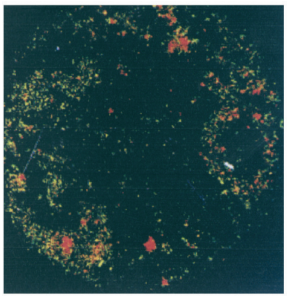


Figure 1.4 Cβ2microglobulin fibrils stained with 120 µM Congo Red solution. When viewed under cross-polarized light, the sample exhibits apple-green birefringence typical for amyloid fibrils. (28)

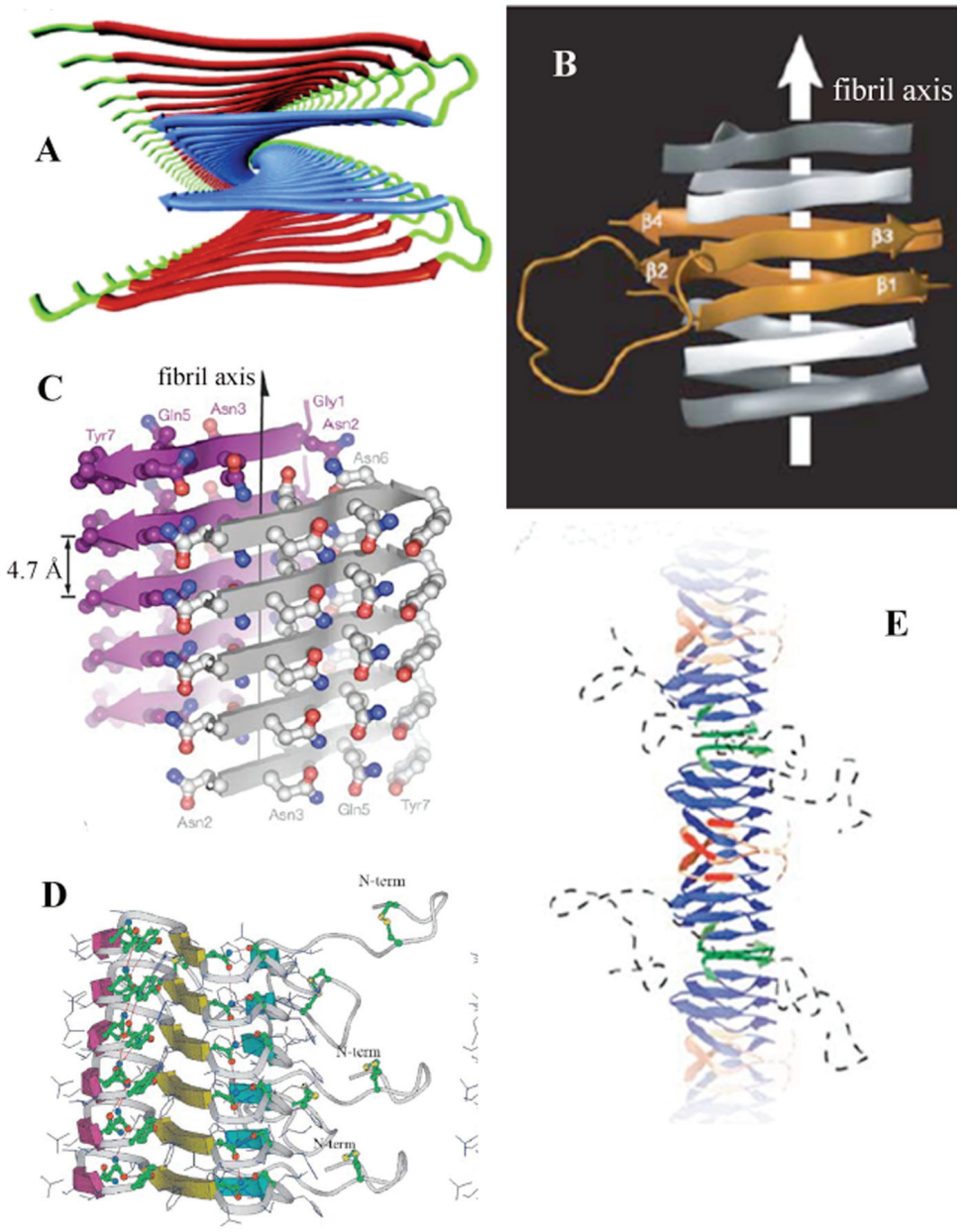


Figure 1.5 Three dimensional models of fibrillar amyloid aggregates.
a) The protofilament of the Aβ peptide as viewed along the axis of the fibril. b) The fibril from the C-terminal domain of the fungal prion protein HET-s. c) Atomic structure of the microcrystals assembled from the GNNQQNY peptide, where each peptide forms one β-strand. d)The protofilament from amylin, with an un-structured N-terminal tail bearing a disulfide bridge. e)The fibril from the NM region of Sup35p.
Figure and caption adapted from (17).

For a long time the ultrastructure of amyloid fibrils has been poorly characterized because they are too large and insoluble to be used for solution

NMR experiments and do not crystallize properly for X-ray crystallography measurements. Recently these limitations have been partially overcomed follwing the application of solid-state NMR (SSNMR) spectroscopy (29-31) and the characterization of peptide fragments that model essential structural features of amyloid fibrils and can be studied by mean of X-ray microcrystallography (32,33). Examples of well-characterized structural model obtained using different techniques are given in Figure 1.5 (17) .

1.3 The utility of using model proteins to study the fundamentals of protein aggregation

The principal aim of studying the mechanism of protein aggregation is to obtain a deeper knowledge on a great number of diseases of very high social and economical impact. Nevertheless, it is well known that a huge number of proteins, if not all, can give rise to amyloid aggregation (34). Several proteins that are not involved in disease were shown to form amyloid-like aggregates using particular organic cosolvents, a high temperature, high pressure, or pH extremes (35-38). More recently, a combination of agitation and high salt concentration has been used to form amyloid fibrils from one of such proteins (39). A few examples of such proteins are reported in Table 1.2 (21).

Table 1.2 Non-disease related amyloidogenic proteins and peptides

Non-disease-related amyloidogenic proteins and peptides

Protein (peptide), reference	Type of structure	Protein (peptide), reference	Type of structure
Betabellins 15D and 16D [294]	β-Sandwich	Prothymosin α [41]	Natively unfolded
Cytochrome c_{552} [133]	α-Helical	Myoglobin [38]	α-Helical
Methionine aminopeptidase [35]	α-Helical	Muscle acylphosphatase [32]	α/β
Phosphoglycerate kinase [34]	α/β	Hen egg white lysozyme [33]	α + β
Hen egg white lysozyme, β-domain [33]	β-Sheet	Acidic fibriblast growth factor [42]	β-Barrel
PI3-SH3 domain [27]	β-Barrel	OspA protein, BH^{9-10} peptide [36]	β-Turn
β-Lactoglobulin [44]	β-Sheet (predominantly)	De novo αtα peptide [37]	α-Helix-turn-α-helix
Monellin [29]	α/β	Lung surfactant protein C [40]	α-Helix
Immunoglobulin light chain LEN [87,88]	β-Sheet	α-Lactalbumin [45]	α + β
HypF, N-terminal domain, [295]	α/β	V_L domain of mouse antibody F.1 [296]	β-Sheet
Human complement receptor 1, 18–34 fragment [297]	Unfolded	Apolipoprotein C-II [213–215]	Natively unfolded
Human stefin B [298]	α/β	Cold shock protein A [299]	β-Barrel
GAGA factor [300]	Natively unfolded	Protein G, B1 Ig-binding domain [79,301]	Four-stranded β-sheet with a flanking α-helix
Yeast prion Ure2p [302]	α-Helical/unfolded	Cold shock protein B, 1–22 fragment [303]	Unfolded
Herpes simplex virus glycoprotein B fragment [304]	β-Structural	De novo proteins from combinatorial library [305]	β-Structural
The fiber protein of adenovirus, 355–396 peptide from shaft [306]	Fibrillar	Soluble homopolypeptides, [262]: poly-L-lysine poly-L-glutamic acid poly-L-threonine	Unordered

Studies based on small, low-molcular weight proteins or even shorter polypeptides, including ex-novo designed ones, can help investigator to focus on one aspect at a time, i.e. the presence of particular conformations on the backbone, small variations of charge, the presence of small sequence repeats, etc. (40-44). In addition, the study of large numbers of systems, as opposed to the few systems associated with disease, allows statistics of the various determinants involved in the process to be obtained (45,46). Performing experiments on relatively small proteins and peptides is also of great help in translating experimental observations into computational models and algorithms, because the possibilities of evaluate correctly the importance of each parameter involved in the process increase as the system complexity decreases (47-50).

1.4 Mechanism of amyloid fibril formation

1.4.1 General features on aggregation

The obtainement of a complete a description of the aggregation process of a protein is not certainly a simple task. There can be a quite large number of different phases and intermolecular associations in the aggregation process, different kinds of oligomeric species and even different types of amyloid fibrils co-populated at the final stage. An important issue is also the determination of the thermodynamics and kinetics of all the discrete phases involved in the process in which conformational changes occur and the structures of all the monomeric and oligomeric species that form during aggregation.

The time course of amyloid fibril formation usually consists of an initial phase, called "lag" phase, during which no detectable formation of fibrils occurs (51). This phase is due to the time required for the first aggregation nuclei to constitute. During the lag phase a variety of oligomeric aggregates form, not just the nuclei for the formation of mature fibrils, but also dimers, small oligomers, protofibrils, etc. (26,27). After this phase, which can vary greatly in time, from a few minutes or hours to days or even weeks, a rapid exponential growth of fibrils occur (52-54) by association of monomers or other oligomers to the pre-formed nuclei (fig. 1.6). Aggregation of protein into amyloid fibrils is therefore thought to be consistent with a "nucleated growth" mechanism.

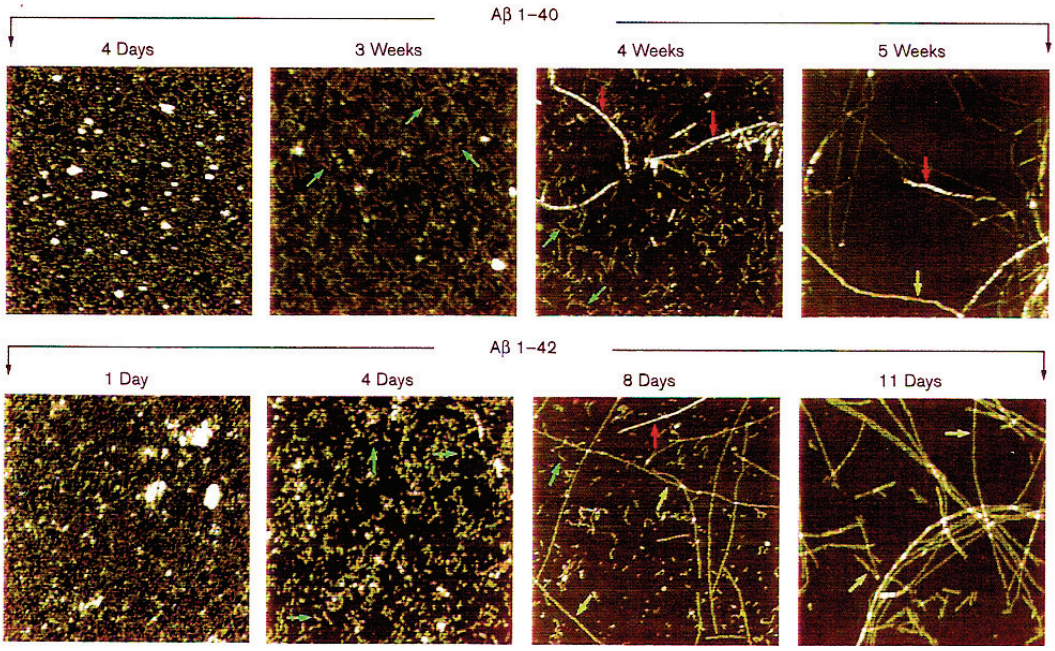


Figure 1.6 AFM images of the time course of fibril formation by Aβ 1-40 (at 45 μM, top row) and Aβ 1-42 (at 20 μM, bottom row). For both Aβ 1-40 and Aβ 1-42, the progression from small protofibrils to longer protofibrils to fibrils of two types is clear, as is the absence of protofibrils at the latest time points. The progression from protofibrils to fibrils occurs more rapidly for Aβ 1-42, despite the fact that protein concentration is lower. Figure and caption adapted from (24).

In analogy with other nucleated growth systems (for example crystal growth), it is possible to "seed" the solution with ready made fibrils to increase the rate of fibril growth: this avoids the need for nucleation and consequently the lag phase results to be shortened, or even canceled (52,53).

In same experimental conditions the lag phase can be influenced or deleted, probably creating a situation in which nucleation in no longer rate-limiting (55,56). Anyway, the absence of the lag phase does not ensure us on the absence of a nucleated growth: it can simply mean that the main fibril formation occurs on time-scale comparable to that of first nucleation process (17).

1.4.2 Oligomers and protofibrils

The oligomeric species that preceed the formation of amyloid fibrils have been

the subject of great scientific attention recently (57-61). They are considered very important for two reasons. First, they are the main intermediate species between monomers and amyloid fibrils, and their study and comprehension can shed light on aggregation pathways. Second, they are thought to be the real active forms responsible for the onset of many protein deposition diseases, the amyloid fibrils representing rather an inert dead-end product (62,63). Among the various prefibrillar species, the most commonly observed aggregates are those assemblies that are often termed protofibrils (24,26). These species show a significant structural regularity. For example, the aggregation process of the amyloid β peptide involves the early appearance of a variety of nonfibrillar species, visible using AFM or TEM (24,26). The simplest "unit" seems to be a spherical aggregate, of about 20 monomers, that assembles into chains that can be straight, curved, or even circular (24,26,64). In all cases protofibrils have a diameter of ca. 2-5 nm. Protofibrils are able to bind to ThT and CR and possess an extensive β-sheet content, as indicated by far-UV circuar dichroism measurements (26,64-66).

Many other amyloid-forming proteins can form similar structures, including α-synuclein (67), poly-glutamine repeats (64), β2-microglobulin (68), the acylphosphatase from *Sulfolobus Solfataricus* (Sso AcP) (69), the N-terminal domain of the hydrogenase maturation factor from *Escherichia Coli* (HypF-N) (70). Protofibrils can be both on-pathway to fibrils (24,53) or off-pathway species that simply accumulate during amyloidogenesis (68,71). There is evidence that not only mature fibrils but also the oligomeric protofibrils share a number of structural features, regardless of the protein that originate them. This finding has also lead to the creation of a specific antibody able to recognize

protofibrillar material from different sources (72). The interesting remark is that this antibody does not recognize the corresponding monomers nor the fully fibrillar state of the protein, suggesting that the oligomeric states share some common epitopes on the surface that are lost with fibril formation.

1.5 Native-like aggregation

Aggregation of globular proteins can occur either via partial unfolding across the major free energy barrier of (un)folding, as well established in past years (73-76), or via a mechanism that implies the early formation of native-like oligomers (56,69,77,78). The partial unfolding theory has been the most accreditated for a considerable length of time, following the observation that in most cases the experimental conditions or mutations that promote aggregation also induce a destabilization of the protein native structure, up to the partial or complete loss of the native folding. Such experimental conditions are usually chemical or physical alterations of the solution, like high temperature, high pressure, low or high pH, moderate concentrations of organic co-solvents as fluoroalcohols (38,68,79-81). This "conformational change" hypothesis was supported also by the evidence that in some proteins, as human lysozyme, human transthyretin (TTR) and the N-terminal domain of the HypF protein of *E. Coli* (HypF-N), the aggregation process is promoted by amino acid substitutions that destabilize the native fold allowing the population of partially folded states (82,83). The need for unfolding, at least partial, is also supported by studies that have demonstrated that the addition of antibodies or specific

ligands stabilizing the native state, monomeric or oligomeric, impair aggregation (84-88).

In the last years, a different hypothesis emerged to describe aggregation from initially folded proteins: a globular protein can aggregate from its native state without any need to unfold across the major free energy barrier, but just via structural fluctuations that normally occur under physiological conditions (Fig. 1.7) (17). For example, at low pH the S6 protein from *Thermus Thermophilus* converts very rapidly to a quasi-native state; a study performed on the kinetic properties of a set of mutants has shown that the longer lag phases in aggregation under these conditions correlate with faster rates of unfolding (56). This means that kinetic stability can control fibrillation, implying that the quasi-native state, rather than the denatured state, is the species that participates in nucleation (56). Human insulin forms fibrils at low pH that are preceeded by the formation of oligomers in which the protein retains the native conformation rich in α-helices (77). In this process the extensive β-sheet content typical of amyloid appears only in a later phase (77). Another example, that is particularly relevant in relation to the present thesis work, is the case of the acylphosphatase from *Sulfolobus Solfataricus* (Sso AcP). This protein aggregates under conditions in which a native-like state is initially populated and forms, as a first step, aggregates in which the monomers maintain their native-like topology (69,78). The unstructured 11-residue N-terminal segment and the edge β-strand-4 has been shown to play a major role in the process (89). Kinetic experiments on a set of Sso AcP variants revealed that the major event of the first aggregation step is the establishment of an inter-molecular interaction between the unstructured segment of one Sso AcP molecule and the globular unit of another (90).

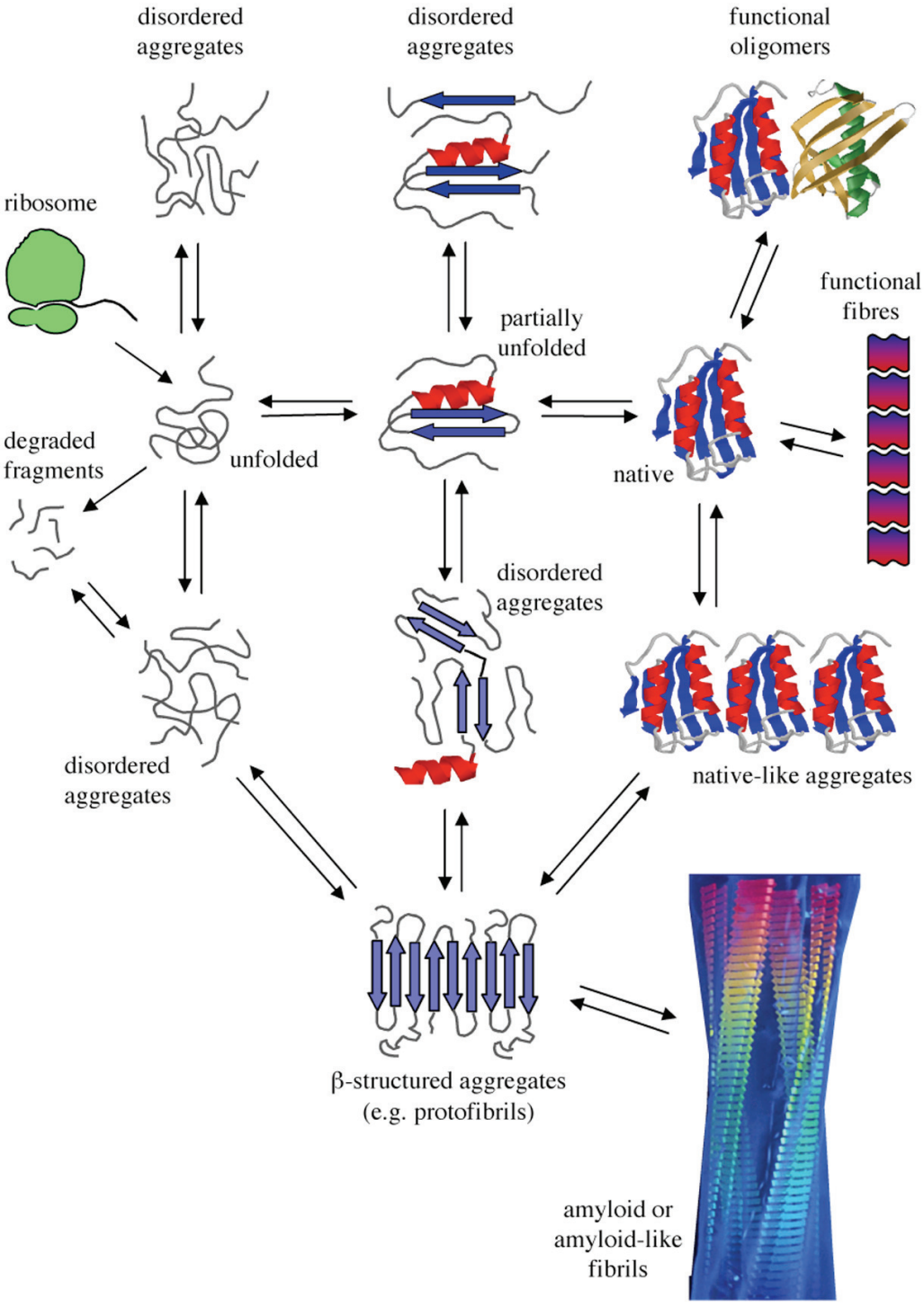


Figure 1.7 A schematic representation of some of the many conformational states that can be adopted by polypeptide chains and of the means by which they can be interconverted. The transition from β-structured aggregates to amyloid fibrils can occur by addition of either monomers or protofibrils (depending on protein) to preformed β-aggregates. Figure and caption adapted from (17).

In a subsequent step these initial aggregates convert into β-sheet containing and ThT-/Congo Red positive amyloid-like protofibrils, via a structural reorganization process in which the individual Sso AcP molecules undergo an intramolecular conformational rearrangement (69,78,89).

1.6 The Acylphosphatase-like structural family: structure and function

1.6.1 The acylphosphatase-like family

In this thesis I focus my attention on proteins belonging to the acylphosphatase-like structural family, mainly the acylphosphatase from *Sulfolobus solfataricus* (Sso AcP) and the acylphosphatase 2 from *Drosophila Melanogaster* (AcP Dro2). In chapter V, for comparison purposes, I will also analyze some features of the acylphosphatase from human muscle (mt AcP) and the N-terminal domain of the HypF protein from *E. Coli* (HypF-N).

Acylphosphatase (AcP) is a small, about 100 residue long, α+β protein belonging to the ferredoxin-like fold (Fig. 1.8). The structure of the protein is highly conserved throughout the family. All the AcPs so far characterised show the same βαββαβ topology that originates from an antiparallel β-sheet docking against two antiparallel α-helices (Fig. 1.8) (91-94).

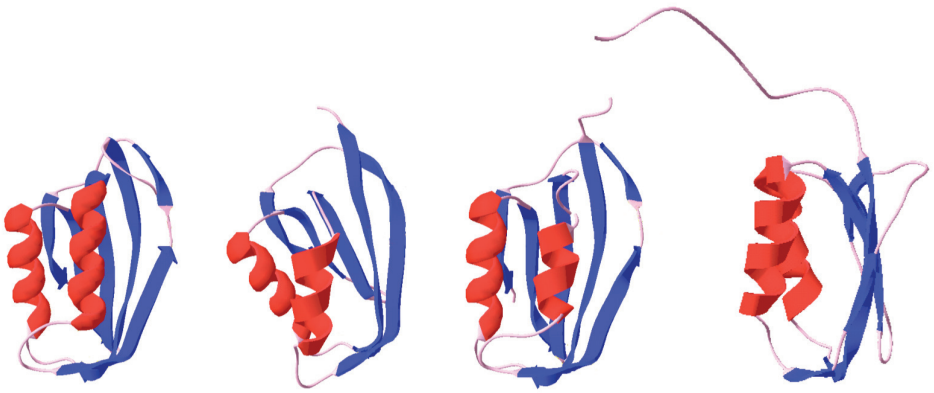


Figure 1.8 The structure of some members of the acylphosphatase-like structural family. From left to right: AcP from *Homo Sapiens* (muscular type, PDB entry 1APS), HypF-N from *Escherichia Coli* (PDB entry 1GXT), AcP 2 from *Drosophila Melanogaster* (PDB entry 1URR) and AcP from *Sulfolobus solfataricus* (PDB entry 1Y9O).

AcP is a widely diffused enzyme (enzyme commission 3.6.1.7) able to hydrolyze acylphosphates, with formation of a phosphate ion and a carboxylate group (95) . The mechanism of catalysis has been studied in detail in mt AcP (96-98). The catalytic residues are an Arginine and an Asparagine residues which are highly conserved within the family. The catalytic activity requires the presence of both residues cited above and a water molecule that is used to start the nucleophilic reaction that leads to the hydrolysis. In particular, the Arginine residue has been shown to be the main phosphate binding residue (98), while the Asparagine is known to bind the water molecule and catalyze the hydrolysis reaction (97). Analysis of the crystal structure of both the muscular (mt AcP) and common-type AcP (ctAcP) displayed that the side chain of this Arginine, namely Arg23, is in close spatial proximity with the side chain of the prevoiusly cited Asparagine (Asn41) and with the 15-21 loop, a small mobile region highly conserved among the different isoforms (92,99). A mutational analysis of the loop resulted in an impairment of the binding capacity and a

wrong positioning of the substrate, leading to a loss of catalytic activity (96).

The precise physiological function of AcP in the various organisms is not clear, although different functions have been suggested for this protein. Some authors proposed a role in regulation of membrane transport, relying on the ability of AcP to hydrolyse the β-aspartyl-phosphate forming during the action of membrane pumps (100); others showed the ability of AcP to induce apoptosis in HeLa cells (101) and also undergo a nuclear migration and interact with other DNAses in response to various apoptotic stimuli (102). Finally, since the AcP levels increase in several cell lines during differentiation and AcP is able to hydrolyse both γ and β phosphate groups of ATP, it was proposed a role for AcP in regulation of differentiation through a control of the ATP and ADP levels (103).

1.6.2 Folding and aggregation in the acylphosphatase-like family

In the last years AcPs have been largely used as models for both folding and amyloid-like aggregation studies. The folding process shows high variability in the acylphosphatase-like family. The first comparisons were made between the two human isoforms, mtAcP and ctAcP, ad it was shown that ctAcP bears a lower conformational stability than mt AcP, although its folding rate is greatly higher than that of mt AcP, suggesting no correlation between folding rate and conformational stability within the AcP family (104,105).

In addition, both the folding processes of Sso AcP, HypF-N from *E. coli* and AcP from *E. coli* have been characterised in some detail (106-108). These studies, carried out on structurally related but evolutionary distant proteins, allowed some important parameters for folding to be identified (106-108).

Indeed, correlations within the acylphosphatase-like family indicate hydrophobicity, relative contact order, α-helical propensity and disulfide bridges as important determinants of folding rate (106-109) .

The processes of amyloid aggregation for different AcPs have also been studied. Interestingly, in spite of the common structural features they share, different AcPs can undergo aggregation starting from different conformational states. In the case of mt AcP and HypF-N, it was shown that a partial unfolding is required to initiate the process (35,110), while Sso AcP can form amyloid fibrils starting from a native-like state (69,78,89,90). In the latter case, as mentioned in the previous section 1.4, two types of prefibrillar aggregates have been found. Aggregation results initially in the formation of oligomers with a small content of β-sheet structure, no ThT, Congo Red or ANS binding and, remarkably, enzymatic activity. Later on, oligomers with characteristics reminiscent of amyloid protofibrils will form, with the latter originating from the structural reorganization of the initial assemblies. These species consist in spherical or chain-like protofibrils with extensive β-sheet structure, ThT, Congo Red and ANS binding and no enzymatic activity.

1.7 Aim of the thesis

The scope of this thesis is to study in further detail the native-like aggregation process using Sso AcP and a novel AcP that has not been studied before, namely the isoform 2 of acylphosphatase from *Drosophila Melanogaster*.

The first part of the thesis describes the process of amyloid fibril formation of the acylphosphatase 2 of *D. Melanogaster* (AcPDro2), previously cloned and purified as described (111). We found that this protein is able to form amyloid-like fibrils under conditions in which the conformational stability of its native state is unchanged. The results of this study have been published (112).

The second part of the thesis was aimed at testing the effects of specific ligands binding to the active site on the aggregation behaviour of AcPDro2 and Sso AcP, i.e. proteins in which amyloid formation starts from a native-like state. For the study performed on SsoAcP I am greatly indebted with dr. Georgia Plakoutsi, who made the measurements. The results of this part of the work have been published (113).

Finally, the last part of the work presented in this thesis stemmed from a theoretical comparative inspection of the structures of the four AcPs for which the aggregation process has been studied in some detail, namely Sso AcP, AcPDro2, HypF-N and mtAcP. This analysis provided clues to the identification of the factors determining either aggregation via unfolding or native-like aggregation and was coupled with a specific mutational analysis of the edge β-strand-4 of SsoAcP to convert the aggregation pathway from one to other mechanism. The results of this study have been published (114).

CHAPTER II – MATERIALS AND METHODS

2.1 Protein cloning, mutagenesis and expression

2.1.1 Cloning of the wild-type AcPDro2 and Sso AcP gene in pGEX vectors

The genes coding for AcPDro2 and Sso AcP were cloned in pGEX-4T1 and pGEX-2T plasmids, respectively (GE Healthcare Life Sciences, formerly Amersham Biosciences, Uppsala, Sweden). The genes were amplified by polymerase chain reaction (PCR) using suitable primers containing the restriction sites for *Xho*I and *Eco*RI (for AcPDro2), and for *Bam*HI and *Eco*RI (for Sso AcP), respectively (New England Biolabs, Ipswich, Massachusetts). The fragments resulting from PCR amplification were digested with the corresponding restriction enzymes and then ligated into pGEX vectors (Figure 2.1) previously digested with the same endonucleases.

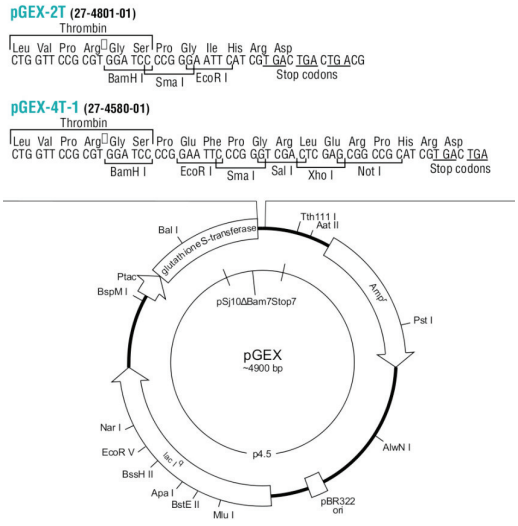


Figure 2.1. Map of the pGEX vectors. GST fusion proteins are constructed by inserting a gene or gene fragment into the multiple cloning site. Expression is under the control of the tac promoter, which is induced by the lactose analog isopropyl β-D thiogalactoside (IPTG). The vectors is also engineered with an internal $lacI^q$ gene. The $lacI^q$ gene product is a repressor protein that binds to the operator region of the tac promoter, preventing expression until induction by IPTG, thus maintaining tight control over expression of the insert.

In the ligated plasmid the protein sequence is inserted downstream and in frame with glutathione S-transferase (GST) resulting in a fusion protein of approximately 37 kDa. At the junction site of the two proteins there is the recognition and cleavage sequence of the protease Thrombin (Sigma-Aldrich, Saint Louis, Missouri). The plasmid resulting from the cloning procedure was checked by DNA sequencing.

2.1.2 Site-directed mutagenesis

The mutated genes of AcPDro2 and Sso AcP were created with the QuickChange site-directed mutagenesis kit from Stratagene (La Jolla, California), using specifically designed primers containing the desired mutations (purchased from MWG Biotech, Firenze). The PCR reaction generated a mutated plasmid that was then treated with *DpnI* endonuclease (1h at 37 °C). This enzyme specifically digesting methylated DNA selectively cuts the non-mutated parental DNA template. Finally, the digestion mix has been transformed into *E. coli* XL1-Blue supercompetent cells. The successfully transformed cells were then selected by growing the transformation mix in the presence of a specific antibiotic. The presence of the desired mutation was assessed by sequencing the entire gene.

2.1.3 Protein expression and purification from pGEX vectors

Cultures of *E. coli* XL1-Blue cells harbouring the plasmid were grown overnight at 37 °C in LB medium supplemented with 100 μg/ml ampicillin with vigorous shaking. These were then diluted (1:50) in fresh LB medium and grown until the OD600 reached approximately 0.6 under the same conditions. Protein

expression was induced for 2-3h at 37 °C using 0.2 mM isopropyl-β-D-thiogalactoside (IPTG). The induced cells were harvested by centrifugation and the pellet was resuspended in PBS buffer (20 mM NaH2PO4, 250 mM NaCl, pH 7.3, containing 1mM EDTA, 1 mM β-mercaptoethanol), and freezed to -20°C overnight or more. Cells were then lysed using 1.0 mg ml^{-1} lysozyme (30 min in ice under agitation) followed by 5-6 cycles of 30 s sonication at 20 kHz and 100W with 30 s cooling period between each burst. The supernatant obtained after a centrifugation step (40 min, 16000 rpm, 4 °C) that contains the soluble proteins, was applied on an affinity chromatography column (glutathione-agarose resin, from Sigma) equilibrated in the same buffer. Cleavage of the fusion protein bound to the resin was performed in 50 mM Tris-HCl buffer, pH8.0, containing 150 mM NaCl, by overnight incubation at 4 °C with bovine Thrombin (50 units). The Thrombin digestion yields a protein with an extra glycine-serine dipeptide at the N-terminal that originates from its cleavage site. The protein eluted from the column in Tris-HCl buffer was subsequently exchanged to the storage buffer (see section 2.1.5) by ultrafiltration (YM membrane, 3000 Da MW cut-off, Millipore).

2.1.4 Determination of protein purity and concentration

Protein purity was checked by SDS-PAGE. The final protein concentration was calculated from absorption measurements at 280 nm, using ε_{280} values of 1.09 and 1.24 ml mg^{-1} cm^{-1} for AcPDro2 and Sso AcP, respectively. ε_{280} values were obtained using the ProtParam application available on the ExPASy Proteomics server (www.expasy.org)

2.1.5 Protein storage

AcPDro2 was stored in 50 mM acetate buffer, 2 mM DTT, pH 5.5. Sso AcP was stored in 10 mM Tris-HCl, pH 8.0. Before starting each experiment, each protein solution was centrifuged at 18000 rpm for 3-5 minutes, filtered using 0.02 μm Anotop 10 filters (Whatman International Ltd, Maidstone, UK), and the protein concentration was re-determined spectrophotometrically.

2.2 Dye binding assays

2.2.1 Congo Red absorbance

Congo Red is a diazole dye that binds amyloid proteins with a relatively high specificity. This binding has been shown to depend on the secondary conformation of amyloid, and specifically the β-sheet conformation (115). When Congo Red binds to amyloid, its absorption peak undergo a marked red shift, from 490 to 540 nm (116).

AcPDro2 was incubated at a concentration of 0.4 mg ml^{-1} for approximately 60 hours in 50 mM acetate buffer, 2 mM DTT, pH 5.5, 25°C, in the presence of 0 or 5% (v/v) trifluoroethanol (TFE). In both cases a 60 μl aliquot was mixed with a 440 μl solution containing 20 μM CR, 5 mM phosphate buffer, 150 mM NaCl, pH 7.4, in a reduced volume, 5 mm path lenght quartz cuvette. Corresponding solutions of CR without protein and protein with no CR were used as controls. Optical absorption spectra were acquired from 400 to 700 nm by an Ultrospec 2000 spectrophotometer (Pharmacia Biotech, Cambridge, UK.). The spectrum obtained with buffer only was subtracted from all acquired spectra.

2.2.2 Thioflavin T fluorescence

The fluorescent dye Thioflavin T (ThT) is a benzothiazole dye able to bind selectively to amyloid aggregates, probably through ionic and hydrophobic interactions (117-119). Following this binding the fluorescence emission of ThT shows an increased maximum around 485 nm that can be easily monitored. Recent studies showed that ThT binds in a specific regular fashion to amyloid fibrils such that their long axes run parallel (118). It has been proposed that the binding occurs in "channels" along the length of the β-sheet structure of protofilaments (Fig. 2.2).

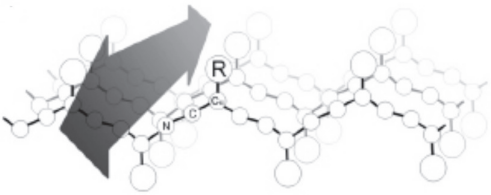


Figure 2.2 Schematic representation of the possible mechanism of interaction between Thioflavin T and a β-sheet. The major axis of the dye can fit in a possible "binding channel" of the β-sheet. From (118).

In a first set of experiments, described in Chapter III, AcPDro2 was incubated at a concentration of 0.4 mg ml^{-1} in 50 mM acetate buffer, 2 mM DTT, pH 5.5, 25°C in the presence of concentrations of TFE ranging from 0 to 20% (v/v). At regular times 60 μl aliquots of each sample were added to 440 μl of a solution containing 25 μM ThT, 25 mM phosphate buffer, pH 6.0. The steady-state fluorescence values of the resulting samples were measured at 25°C using a 2x10 mm path length quartz cuvette and a Perkin-Elmer LS 55 spectrofluorimeter (Wellesley, Massachusetts) equipped with a thermostated cell holder attached to a Haake F8 water bath (Karlsruhe, Germany). The excitation and emission wavelengths were 440 and 485 nm, respectively.

Another set of data was collected after incubating the protein in 50 mM formiate buffer, 5% (v/v) TFE, 25 mM NaCl, 2 mM DTT, pH 3.5, 25°C.

In the experiments described in Chapter IV, AcPDro2 and Sso AcP were incubated at concentrations of 0.4 mg ml^{-1} in 50 mM acetate buffer, 2 mM DTT, pH 5.5, 25 °C, containing P_i at a concentration ranging from 0 to 10 mM and TFE (10 and 20% (v/v) TFE for AcPDro2 and Sso AcP, respectively). At regular times 60 µl aliquots of each sample were added to 440 µl of a solution containing 25 µM ThT, 25 mM phosphate buffer, pH 6.0. The steady-state fluorescence values of the resulting samples were measured at 25°C using the same equipment described above. All measured fluorescence values are given after subtracting the fluorescence measured in the absence of protein and normalised so that the final fluorescence at the endpoint of the kinetic trace was 100%.

Finally, in the experiments reported in Chapter V, all variants of Sso AcP were incubated at concentrations of 0.4 mg ml^{-1} in 50 mM acetate buffer, pH 5.5, 25 °C, in the presence of 3%, 10% or 20% (v/v) TFE. Measurements were performed as described above.

2.2.3 ANS fluorescence

1-anilino-8-naphtalenesulfonic acid (ANS) is an amphiphilic dye that is able to bind to exposed clustered hydrophobic residues of a polypeptide chain (120). Binding is non-covalent and involves a combination of hydrophobic and electrostatic interactions, the latter depending on a strong ability to bind cationic groups (121) . The fluorescence emission of this molecule undergoes a blue shift and strong enhancement when its exposure to water is decreased. Consequently, fluorescence of ANS increases substantially when proteins to which it is bound undergo transitions from unfolded to fully or partially folded

states that provide shielding from water (122).

In the experiments performed during this thesis, AcPDro2 was diluted to a final concentration of 0.02 mg ml^{-1} in 55 μM ANS, 50 mM acetate, 2 mM DTT, pH 5.5, 25°C. We prepared two samples containing 0 and 5% TFE, respectively. Fluorescence spectra were acquired immediately after dilution using the same cell and equipment described under the *ThT assay* paragraph. Excitation wavelength was set at 380 nm and emission range was from 410 to 630 nm. Spectra were also acquired for the corresponding solutions in the absence of protein. Each spectrum was the average of two scans and was smoothed using a moving average function with the statistical mean of seven points. ANS concentration was determined spectrophotometrically using an ε_{375} value of 8000 ml $mmol^{-1}$ cm^{-1}.

2.3 Enzymatic activity assay

The enzymatic activity of the acylphosphatases used for these studies has been measured monitoring the absorbance of the reaction substrate. The absorbance decay was followed at 283 nm using benzoylphosphate as substrate, as previously well described for other acylphosphatases (97,123). Activity was measured with a Perkin-Elmer λ 4 B UV-Vis spectrophotometer (Wellesley, Massachusetts). The non-catalyzed spontaneous hydrolysis of benzoylphosphate was subtracted from all measurements.

In the experiments described in Chapter III, the enzymatic activity of AcPDro2 was tested in the presence of TFE concentrations ranging from 0 to 5%, using a

final protein concentration of 0.5 μg ml^{-1} and 4 mM benzoylphosphate. Other experimental conditions were 50 mM acetate buffer, 2 mM DTT, pH 5.5, 25°C. Measurements were performed immediately after addition of TFE, without any incubation time.

In the experiments reported in Chapter V, the enzymatic activity of Sso AcP and its mutants was tested using a final protein concentration of 0.5 μg ml^{-1} and 4 mM benzoylphosphate. Other experimental conditions were the same as in the previous set of experiments.

2.4 Intrinsic fluorescence measurements

2.4.1 Theory

Some atoms and molecules, called fluorochromes, are able to absorb light at a well defined wavelength and then to emit light of longer wavelength. For any molecule, several different electronic states exist, each subdivided into a number of vibrational and rotational energy levels (shown as S0, S1… Sn in Fig. 2.3). When common fluorophores are irradiated with ultraviolet or visible light, they absorb energy and are excited to a higher vibrational level. After energy absorption, molecules remain in the excited state for approximately 10^{-9} seconds, and then return to ground state with a process called relaxation. There are many different ways in which an excited molecule can return to ground state. If the photon emission occurs between two energy levels with the same spin, as illustrated in fig. 2.3 (e.g. S1 and S0) the phenomenon is called fluorescence. Instead, if the spin states of the levels are different (e.g. from T1 to

S0), the emission is called phosphorescence. Fluorescence is much more likely to occur than phosphorescence for most molecules, but the lifetime of fluorescent states is very short (from 10^{-5} to 10^{-8} seconds). The emitted photons have less energy with respect to exciting ones, and so the emission spectrum is shifted to longer wavelengths.

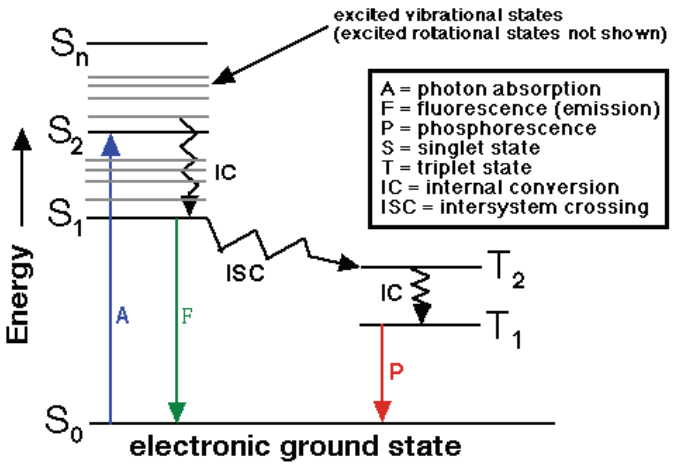


Figure 2.3 Jablonski energy diagram.

Fig.2.3 shows also some relaxation processes, called nonradiative, in which there isn't any light emission: internal conversion (IC) and intersystem crossing (ISC). Internal conversion is a transition between energy states of the same spin, similar to fluorescence but without light radiation, occurring in 10^{-12} seconds or less. Intersystem crossing is a transition between different spin states, comparable to phosphorescence but radiationless, occurring in 10^{-9} seconds.

Fluorescence is generally associated with highly conjugated polycyclic aromatic molecules. The aromatic side chain amino acids that are present in polypeptide chains (Trp, Tyr, Phe) confer them the ability to emit fluorescence when excited. In proteins, about the 90% of the fluorescence emission is given by Trp side chains. Considering that Trp fluorescence is highly influenced by its close environment, it is possible to follow changes in protein conformation,

stability, substrate binding, etc., monitoring the protein fluorescence emission. For example, Trp side chains normally resides on the hydrophobic interior of proteins where the environment is relatively non-polar. Upon protein destabilization or even denaturation with heat or a chemical agent, the environment of the Trp aromatic ring changes, becoming more polar as the residue get in contact with the aqueous solution. This leads both to an increase in intensity of fluorescence emission and to a shift of the maximum peak from approximately 330 to 365 nm.

In some cases, it can be difficult to follow conformational transitions that reach the equilibrium very rapidly, because conventional fluorimeters require manual mixing that causes a delay of many seconds in measurements, while the transition process ends in few milliseconds. To follow a process from the very beginning one should use a fluorimeter coupled with a stopped-flow apparatus. Using this instrument, small volumes of solutions are pushed from two ore more syringes to a high efficiency mixer just a few milliseconds before passing into a measurement flow cell. The steady state equilibrium is reached during the solution flow. At a certain moment the flow is stopped and some of the flowing mixed solution remains trapped in the flow cell. As the reaction proceeds, the kinetics can be followed by monitoring the fluorescence variations. Not even in this case the observation can start at the time zero of the reaction, but there is a short dead time that is the time needed from the solution to reach the middle of the cell from the last mixing point. The dead time is dependent on the volume from the last mixing point to the center of the observation area (dead volume) and on the total flow rate through the cell corresponding to the sum of the flow rate set on the syringes.

For the rapid kinetic experiments described in this thesis I used a fluorimeter equipped with the Bio-logic SFM-3 stopped-flow apparatus (Claix, France). The latter handles three stainless steel syringes driven by independent stepping motors and two mixing chambers (Fig. 2.4).

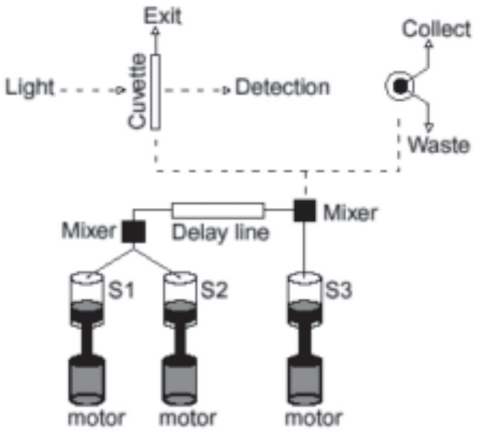


Figure 2.4 Schematic representation of a stopped-flow apparatus

The first mixing chamber is situated after the syringes 1 and 2. After this comes a delay line whose volume determines the time constant. The second mixing chamber is found at the end of the delay line, mixing the solution coming from the mixer 1 with that from syringe 3. Then, the liquid is filled in the observation cell. Volumes and mixing times are computer controlled.

2.4.2 Intrinsic fluorescence spectra

To follow the effect of chemical denaturants on the structural packing, AcPDro2 was diluted to a final concentration of 0.02 mg ml^{-1} in 5 mM acetate, 2 mM DTT, pH 5.5, 25°C and in the presence of various denaturants. Fluorescence spectra were acquired immediately after dilution using the cuvette and instrumental apparatus described under the *ThT assay* paragraph. Excitation wavelength was set at 280 nm and emission range was from 300 to 450 nm. Each spectrum was recorded as the average of two scans and smoothed using a moving average function with the statistical mean of seven points.

2.4.3 Equilibrium unfolding

Unfolding of AcPDro2 was studied at equilibrium in the presence of different concentrations of TFE, ranging from 0 to 5% (v/v). For each TFE concentration 30 samples were prepared containing 0.02 mg ml^{-1} protein in 50 mM acetate buffer, 2 mM DTT, pH 5.5, and various urea concentrations ranging from 0.1 to 8 M. The samples were left to equilibrate for a few minutes at 25°C. Fluorescence spectra were acquired at 25°C using the cuvette and instrumental apparatus described under the *ThT assay* paragraph, with a wavelength range between 300 and 450 nm and an excitation wavelength at 280 nm. Each spectrum was recorded as the average of two scans and smoothed using a moving average function with the statistical mean of seven points. The fluorescence at 348 nm was plotted *versus* urea concentration and the resulting plot was fitted to the equation described by Santoro and Bolen (124) to determine the free energy difference between the unfolded and folded states of the protein in the absence of urea ($\Delta G_{U-F}^{H_2O}$), the dependence of ΔG_{U-F} on urea concentration (*m* value) and the midpoint of unfolding (C_m).

The monitoring of the degree of unfolding at equilibrium has been also performed using the far-UV circular dichroism technique (see section 2.5).

2.4.4 Folding and unfolding kinetics

For the experiments described in Chapter III, unfolding and refolding of AcPDro2 were followed with the Bio-logic SFM-3 stopped-flow device, coupled to a fluorescence detection system, as described in section 2.4.1. The flow cell was thermostated at 25°C with a Neslab RTE-200 water-circulating bath

(Newington, New Hampshire). The excitation wavelength was 280 nm, and the fluorescence emitted above 320 nm was monitored using a bandpass filter. All the experiments were performed in 50 mM acetate, 2 mM DTT, pH 5.5, at a final protein concentration of 0.02 mg ml^{-1}. In a first set of unfolding and folding experiments the protein was initially in 0 and 6 M urea, respectively. Both reactions were initiated by ten fold dilutions. Unfolding and folding kinetics were studied at final urea concentrations of 5.4 and 0.6 M, respectively, in the presence of TFE concentrations ranging from 0 to 8.25% (v/v). The dead-time was 10.44 ms. The unfolding and refolding traces were fitted to single and double exponential functions, respectively (see Equation A in chapter III). In a second set of experiments, the unfolding and refolding kinetics were monitored at final urea concentrations ranging from 1.80 M to 7.20 M for unfolding and from 0.50 M to 3.20 M for folding. Finally, in a third set of experiments the protein was unfolded in 2 M urea, 10 mM trifluoroacetic acid (TFA), pH 2.0. Folding was initiated by a 20 fold dilution in 50 mm acetate buffer, pH 5.5, 2 mm DTT, 25°C, at final urea concentrations ranging from 0.1 to 1.8 M. These experiments were performed both in 0% and 5% (v/v) TFE. In these second and third sets of experiments the analysis was carried out at final TFE concentrations of 0 and 5% (v/v). The unfolding and refolding traces were fitted to an exponential function (Equation E) and the natural logarithm of k_i was plotted *versus* urea concentration for both the analyses carried out in 0 and 5% TFE. The resulting plots were fitted using the equation edited by Jackson and Fersht (125). (see Equation B in Chapter III).

For the experiments described in Chapter V, unfolding and refolding kinetics of Sso AcP and its mutants were studied with the same device but

thermostated at 37 °C. Unfolding kinetics were monitored by preparing a protein sample at 0.4 mg ml^{-1} in 10 mM Tris/HCl, pH 8.0 and by a twenty fold dilution into solutions containing 50 mM acetate buffer, pH 5.5 and various guanidine hydrochloride (GdnHCl) concentrations. In refolding experiments all protein variants, except wild-type, were initially unfolded in 5 M GdnHCl, 10 mM Tris/HCl pH 8, whereas wild type was unfolded in 6 M GdnHCl, 10 mM Tris/HCl pH 8. Reactions were started by a twenty fold dilution into solutions containing 50 mM acetate buffer, pH 5.5 and small GdnHCl concentrations. Unfolding and refolding kinetics were monitored at final protein concentration of 0.02 mg ml^{-1}, final GdnHCl concentrations ranging from 1.90 M to 7.60 M for unfolding and from 0.30 M to 3.90 M for refolding. These experiments were performed both in the absence and in the presence of 3% (v/v) TFE for all variants, except Y86E, for which a 10 % (v/v) TFE concentration was also used. The dead-time was 10.44 ms. The recorded traces were fitted to exponential functions (Equation A) and the natural logarithm of k_i was plotted *versus* GdnHCl concentration for each variant and for each TFE concentration used. In each case, the unfolding and refolding data from these plots were analyzed separately with a best fitting procedure using a linear and a 2nd order polynomial function, respectively. This allowed the values of folding rate constant from a partially folded state ($k_{I-F}^{H_2O}$) and unfolding rate constant ($k_{F-U}^{H_2O}$) to be extrapolated in the absence of GdnHCl from the lines of best fit.

2.5 Circular dichroism measurements

2.5.1 Theory

Circular dichroism (CD) spectroscopy is a form of light absorption spectroscopy that measures the difference in absorbance of right- and left-circularly polarized light by a substance. Unlike in isotropic light, in which the electric field vector oscillates uniformly in all directions, in linearly polarized light the magnitude of its electric field vector oscillates only in one plane, similar to a sine wave. In circularly polarized light, the electric field vector has a constant length, but rotates about its propagation direction. Hence, it forms a helix in space while propagating (Fig. 2.5). If this is a left-handed helix, the light is called "left circularly polarized", and *viceversa* for a right-handed helix.

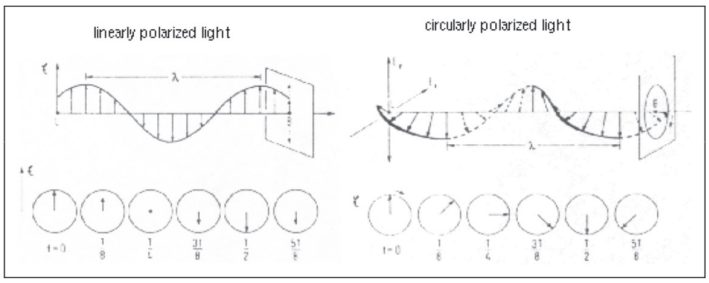
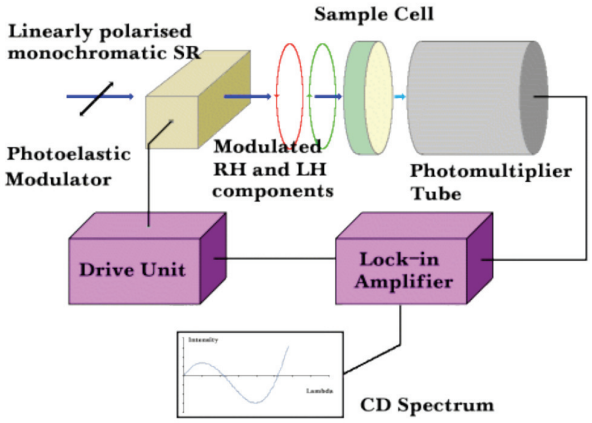


Figure 2.5 Linearly and circularly polarized light

Circularly polarized light interacts differently with chiral molecules, because the two types of circularly polarized light are absorbed to different extents. In a CD experiment, equal amounts of left and right circularly polarized light of a selected wavelength are alternately radiated into a sample (Fig. 2.6). If the sample is optically active, i.e. chiral, one of the two polarized beams is absorbed more than the other, and this difference of absorption is measured in dependence of wavelength, giving the CD spectrum of the sample. For the experiments reported in this thesis, a Jasco J-810 spectropolarimeter (Tokyo,



Japan) equipped with a thermostated cell holder attached to a Thermo Haake C25P water bath (Karlsruhe, Germany) was used.

Figure 2.6 Schematic drawing of a spectropolarimeter

Chromophores are optically active groups within the protein: backbone amide bonds, disulphide bonds, and aromatic side chains such as Phe, Trp, and, to a lesser extent, Tyr. In secondary structure conformations, the backbone and hence the amide bond chromophores are arranged in regular organized patterns. CD spectroscopy is extremely sensitive to these patterns and each conformation gives rise to characteristic spectral features. Random coil has a positive band at 212 nm and a negative one at around 195 nm; β-sheet shows a negative band at 218 nm ($\pi \rightarrow \pi^*$ transition) and a positive one at 196 nm ($n \rightarrow \pi^*$); α-helix has two negative bands at 222 nm ($n \rightarrow \pi^*$) and 206 nm ($\pi \rightarrow \pi^*$) and a positive band ($\pi \rightarrow \pi^*$) at 190 nm (126).

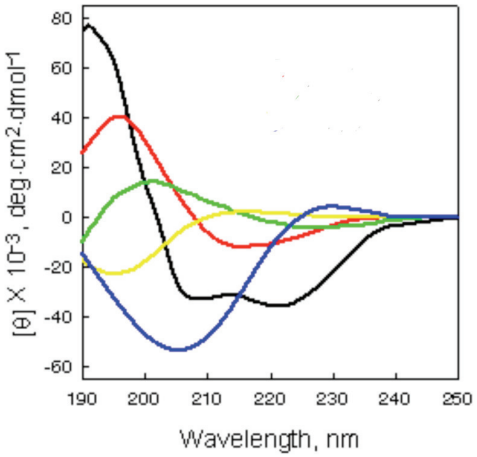


Figure 2.7 Circular dichroism spectra of various "pure" secondary structure: α-helix (black), β-sheet (red), type-1 turn (green), random coil (yellow) and poly-L-proline (blue).

The far-UV CD spectrum of proteins can reveal important characteristics of their secondary structure. CD spectra can be used to estimate the fraction of a protein that is in each secondary structure conformation: α-helix, β-sheet, β-turn, random coil etc. (127). It can be also used to study how the secondary structure of a protein changes as a function of pH, temperature or concentration of denaturing agents, e.g. trifluoroethanol, urea or guanidinium hydrochloride. The near-UV CD spectrum (>250 nm) of proteins provides information on the tertiary structure. The signals obtained in the 250-350 nm region are due to the absorption, dipole orientation and the nature of the surrounding environment of the phenylalanine, tyrosine, cysteine (or S-S disulfide bridges) and tryptophan amino acids. Unlike in far-UV CD, the near-UV CD spectrum cannot be assigned to any particular 3D structure.

2.5.2 Far-UV

For the experiments described in Chapter III , samples of AcPDro2 were prepared at final concentrations ranging from 0.01 to 0.8 mg ml^{-1} in 5 mM acetate buffer, 2 mM DTT, pH 5.5, 25°C, in the presence of 0%-5% and 80% TFE, or 5 M urea with no TFE. Far-UV CD spectra were acquired at 25°C immediately after diluting the stocked protein sample into these solutions. Quartz cuvettes with path lengths ranging from 1 to 10 mm were used depending on protein concentrations. Each spectrum of the protein was recorded from 190 to 260 nm as the average of four scans, blank-subtracted, processed using the adaptive smoothing method (128) and normalised to mean residue ellipticity.

For the experiments described in Chapter IV, the aggregation process of Sso AcP was started by incubating the protein, at a concentration of 0.4 mg ml^{-1}, in 20% (v/v) TFE, 50 mM acetate buffer, pH 5.5, 25 °C, in the presence of 0-10 mM P_i. Each kinetic trace was processed as described above for CD spectra.

For the experiments described in Chapter V, far-UV spectra were first acquired, for all variants, under conditions in which Sso AcP has a native structure, i.e. in 10 mM Tris/HCl buffer, pH 8.0, 25 °C, at a protein concentration of 0.2 mg ml^{-1}, using a 1 mm path length quartz cuvette. Spectra result from 10 accumulations. Far-UV CD spectra were also acquired under conditions promoting aggregation, in 50 mM acetate buffer, pH 5.5, 25 °C, in the presence of the TFE concentration reported in the text, at a protein concentration of 0.4 mg ml^{-1}. In this case spectra were acquired immediately after addition of TFE. In another experimental set, the far-UV CD signal at 208 nm was followed to monitor the first phase of aggregation. Samples were prepared by diluting protein stock solution in 50 mM acetate, pH 5.5, 25 °C, 3-20% (v/v) TFE, at a final protein concentration of 0.4 mg ml^{-1}. The same cuvette described above was used.

Circular dichroism was also used to study the degree of unfolding at equilibrium in presence of denaturants. For the experiments described in Chapter IV, unfolding of AcPDro2 was studied at equilibrium in the presence of different concentrations of P_i, ranging from 0 to 10 mM. For each P_i concentration, 25-30 samples were prepared containing 0.02 mg ml^{-1} protein in 50 mM acetate buffer, 2 mM DTT, pH 5.5, and various urea concentrations ranging from 0.1 to 8 M. The samples were left to equilibrate for a few minutes at 25 °C. Far-UV CD spectra were acquired at 25 °C using the apparatus

described in the previous paragraph and processed using the adaptive smoothing method (128). The ellipticity at 222 nm was plotted *versus* urea concentration and the resulting plot was fitted to the equation described in (124).

For the experiments described in Chapter V, the conformational stability of each Sso AcP variant was studied at equilibrium by preparing 25-30 samples containing 0.02 mg ml^{-1} protein, 50 mM acetate buffer, pH 5.5, and various GdnHCl concentrations ranging from 0 to 8 M. The samples were left to equilibrate for 10 minutes at 37 °C. The ellipticity at 222 nm was acquired using the CD apparatus described above. Ellipticity values were plotted *versus* GdnHCl concentration. The resulting plot was fitted to the equation edited by Santoro and Bolen (124) to obtain C_M and m values. The $\Delta G_{F-U}^{H_2O}$ was calculated according to the equation $\Delta G_{F-U}^{H_2O} = C_M \cdot \underline{m}$, where $\underline{m}$ is the average m value over a set of ca. 30 Sso AcP variants (data not shown).

2.5.3 Near-UV

Samples of AcPDro2 were prepared at final protein concentrations of 0.2 mg ml^{-1} in 5 mM acetate buffer, 2 mM DTT, pH 5.5, 25°C, in the presence of 0% and 5% (v/v) TFE, or 4.5 M urea. Near-UV CD spectra were acquired at 25°C from 250 to 350 nm, immediately after diluting the stocked protein sample into these solutions. A 10 mm path length quartz cuvette and the CD instrument described in the *far-UV CD* paragraph were used. Due to the low signal/noise ratio, spectra for the protein samples in 0% TFE and in 4.5 M urea were recorded as the average of 40 separate scans, while 10 scans were averaged for the protein sample in 5% TFE. Each spectrum was blank-subtracted, processed

using the adaptive smoothing method (128) and normalised to mean residue ellipticity.

2.6 Dynamic Light Scattering

2.6.1 Theory

DLS measures the fluctuation of intensity of the light scattered from a sample containing particles subjected to Brownian motion. When a laser beam illuminates the sample, the particles in solution deviate some of the light waves in all directions and some of the scattered light hits a detector screen placed at an angle that can vary but is always different from 180° (Fig. 2.8).

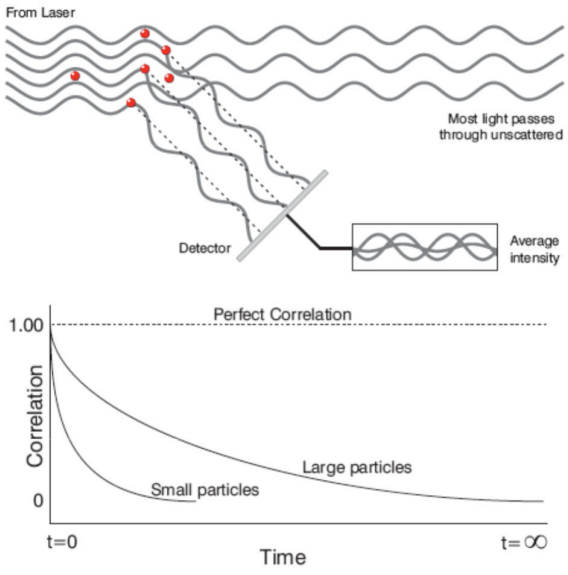


Figure 2.8 Schematic representation of a DLS measurement, showing the mechanism of scattering detection (upper panel) and examples of correlation functions.

At any time the screen detector will show a so-called "interference pattern". The name is due to the fact that the light intensity detected by the screen varies with relation to the phase interferences of the particles hitting it at a certain time: particles with the same phase interfere constructively and give a more intense light on the detector. The instrument determines the change in intensity of the interference pattern during time, using a digital correlator that measures

the degree of similarity (or correlation) between the interference pattern at time 0 and that at a following time t. This is translated into a normalised correlation function, i.e. the correlation versus time. Figure 2.8 shows the correlation function for large and small particles. Large particles move slowly while small particles move quickly, such that the intensity of the speckle pattern will also fluctuate slowly and quickly, respectively. Therefore, the rate of decay for the correlation function will be related to particle size, as it is much faster for small particles than it is for large ones.

All the experiments shown in this thesis were performed with a Zetasizer Nano S DLS device from Malvern Instruments (Malvern, Worcestershire, UK) thermostated with a Peltier system.

2.6.2 Determination of the apparent hydrodynamic diameter

For the experiments shown in Chapter III, samples were prepared at final AcPDro2 concentrations ranging from 0.3 to 1.0 mg ml^{-1} in 50 mM acetate buffer, 2 mM DTT, pH 5.5, 25°C, in the presence of 0% and 5% TFE. Each protein sample was prepared by diluting a stocked protein sample and was immediately centrifuged (18000 rpm for 5 min) and filtered (0.02 μm) shortly before acquiring its size distribution to remove any pre-existing aggregates and dust particles. A low volume (45 μl) black quartz cuvette with 10 mm light path was used.

For the experiments shown in Chapter V, DLS measurements were performed at a protein concentration of 0.3 mg ml^{-1}, in 50 mM acetate buffer, pH 5.5, 25 °C. In another experimental set the measurements were performed at 0.4 mg ml^{-1} protein, after 2-3 hr incubation in 50 mM acetate buffer, pH 5.5, 25

°C, in the presence of 3-20% (v/v) TFE. Disposable plastic cuvettes with a 10 mm light path (Plastibrand, Brand GMBH, Wertheim, Germany) were used.

For all sets of measurements, data were obtained with the apparatus described in section 2.6.1, setting the appropriate viscosity and refractive index parameters for each solution and keeping the temperature at 25°C during the measurements by means of the thermostating system.

2.7 Tapping Mode Atomic Force Microscopy

The aggregate sizes reported in Chapter III were obtained from AcPDro2 incubated at a concentration of 0.4 mg ml^{-1} in 50 mM acetate buffer, 5% (v/v) TFE, 2 mM DTT, pH 5.5, 25°C. Aliquots of 20 µl were withdrawn at various times, deposited on freshly cleaved mica substrates and dried under mild vacuum for 30 min. TM-AFM images were acquired in air using a Dimension 3000 microscope (Digital Instruments, Veeco, Santa Barbara, California) equipped with a 'G' scanning head (maximum scan size 100 µm). Single beam uncoated silicon cantilevers were used (type RTESP, Veeco, Santa Barbara, California). The drive frequency was around 300 kHz; the scan rate was between 0.3 and 0.7 Hz. The aggregate sizes were obtained by measuring the aggregate height in cross section in the TM-AFM images. The apparent size deduced visually from the image turns out to be overestimated because of the broadening effect caused by the AFM tip in the scan plane (129).

The AFM measurements included in Chapter III of this thesis have been performed at the Department of Physics of the University of Genoa by Silvia Torrassa and Dr. Annalisa Relini.

CHAPTER III - RESULTS

Characterization of a new isoform of acylphosphatase from D. Melanogaster (AcPDro2) and evaluation of its aggregation pathway.

3.1 Introduction

In some of the pathological conditions associated with formation of amyloid-like structures the precursor protein is normally folded, in its soluble state, into a globular unit with persistent secondary structure and long-range interactions. Examples include β2-microglobulin, transthyretin, lysozyme, immunoglobulin light chain, superoxide dismutase, ataxin-3, cystatin c and gelsolin (21,130,131). Such proteins spend most of their lifetime in a folded state, but can transiently adopt a partially or totally unfolded state, for example during biosynthesis, translocation or stress conditions. Partially or totally unfolded states can also be in equilibrium with the native state even under non-stress conditions; this may provide a pool of protein molecules that can initiate and propagate the aggregation process, especially when the native structure is destabilised by mutations that shift the equilibrium towards the unfolded state (82,131).

A major question that needs to be answered is whether amyloid formation is initiated by such partially or totally unfolded states that are only occasionally populated, or rather involve directly the native state. The observation that unfolded states have a greater propensity than folded states to aggregate and that amyloid formation is often associated with genetic mutations that can effectively destabilise the native state has led to the proposal that a conformational change is the first essential step in amyloidogenesis, at least for those diseases associated with such globular proteins (21,75,132). However, recent observations have questioned the generality of this mechanism. The pathogenic variant of ataxin-3, the protein associated with Spinocerebellar ataxia type-3, does not appear to have a destabilised native structure leading to the proposal that the pathway for fibril formation is distinct from that of unfolding (133). Amyloid formation by insulin is preceded by an oligomerisation of the protein in which a native-like content of α-helical structure is retained almost completely (77). Within a group of variants of the protein S6 from *Thermus thermophilus*, no significant correlation was found between the rate of fibril formation, under conditions in which a quasi-native state was populated prior to aggregation, and the unfolding rate or conformational stability (56). Aggregation of the acylphosphatase from *Sulfolobus solfataricus* into amyloid protofibrils is more rapid than unfolding under conditions in which the native state is thermodynamically more stable than the major partially unfolded state (69). All of these observations suggest that an association of protein molecules in their native-like states can be the first event in the aggregation of some globular proteins, with the structural conversion into an amyloid conformation occurring subsequently.

The study of amyloid formation under solution media in which the protein is initially in a native-like conformation is therefore of vital important to understand aggregation pathways that may potentially occur *in vivo*. Our ability to distinguish between mechanisms in which partial or full unfolding is a first step and others in which oligomerisation precedes a structural reorganisation is important not just to clarify the pathogenesis of specific protein deposition diseases, but also to recognise the dangerous pathways from which proteins generally need to escape to remain soluble *in vivo*.

As previously described in Chapter I, all the model proteins used in this study belong to the Acylphosphatase superfamily; in particular, the first project was focused on the characterization of a novel acylphosphatase from *Drosophila Melanogaster* (fruit fly), a 102 residue protein domain having a ferredoxin-like topology in its native state, prevoiusly identified and cloned (111). Some years before, another *Drosophila Melanogaster* acylphosphatase (named AcPDro) was also characterized, as the product of a single exon in the Acyp gene located in 2L chromosome. Since it was the most phylogenetically distant acylphosphatase that was characterized to the moment, it was considered as the possible common ancestor for MT and CT vertebrate isoenzymes (134). That hypotesis has been ruled out by the discovering of AcPDro2 and its characteristics. Coded by the CG18505 gene, namely Acyp2, this protein has been shown to be the product of a four-exon gene (while AcPDro is coded by a single exon), and is located on a different chromosome. The expressed and purified protein has been demonstrated to be a well folded enzyme, stable and with full catalytic activity, although the catalytic efficiency is lower than that of the vertebrate isoenzymes (111). The critical residues of the bovine CT AcP active site

(particularly Arg23 and Asn41) are conserved in AcPDro2, together with two active-site water molecules located in the region defined by the L1 and L3 loops and by the N-terminal region of the α1 helix (94).

3.2 Formation of amyloid-like structures at low concentrations of TFE

At the beginning of the project, the very first thing to evaluate was the potential tendency of the protein to aggregate. Since for other acylphosphatases the fluoroalcohol trifluoroethanol (TFE) have been shown to be effective in inducing aggregation, AcPDro2 was incubated at a concentration of 0.4 mg ml^{-1} in 50 mM acetate buffer, 2 mM DTT, pH 5.5, 25°C at TFE concentrations ranging from 0 to 20 % (v/v). After 2 and 24 hours the protein samples were subjected to the ThT assay (119). All protein samples incubated at TFE concentrations higher than 4-5% (v/v) were able to increase significantly the fluorescence of ThT, while no significant aggregation was found to occur at lower TFE concentrations after 24 hr (Fig. 3.1). In 50% (v/v) TFE aggregation appears to be slightly lower than in 20%, with no significant differences in ThT binding after 2 hours and 24 hours (data not shown).

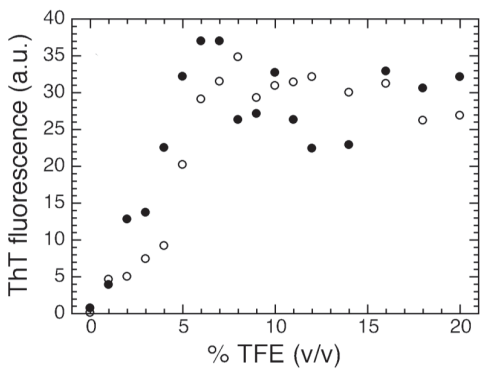


Figure 3.1 AcPDro aggregation measured by ThT fluorescence, after 2h (empty circles) and after 24h (full circles) of incubation in 0-20% TFE.

The protein samples incubated for 60 hours in 0 and 5% (v/v) TFE were also subjected to the CR assay (115). The spectra acquired for CR in the presence and absence of protein pre-incubated in buffer without TFE are superimposable (Fig. 3.2a). The difference spectrum, obtained by subtracting the spectra of protein only and CR only from the spectrum of CR with protein, does not show a peak at 540 nm, indicating that the protein is not able to bind to the dye and cause the characteristic red-shift from 490 to 540 nm (Fig. 3.2a, inset). By contrast, the spectrum acquired for CR in the presence of protein pre-incubated in 5% (v/v) TFE exhibited a significant red-shift relative to the spectrum of CR alone (Fig. 3.2b). The higher absorbance values found in the former at all wavelengths arises from the ability of the protein aggregates to scatter the incident light, as shown by the spectrum acquired for the protein only in the absence of CR (Fig. 3.2b). The difference spectrum, obtained as described above for the analysis in the absence of TFE, shows a peak at 540 nm, as expected for CR bound to ordered aggregates (Fig. 3.2b, inset).

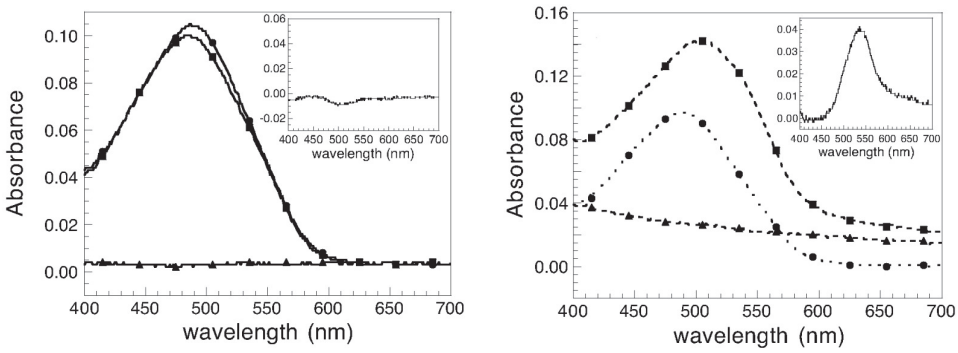


Figure 3.2 Congo Red assay on AcPDro2 aggregates. The protein was preincubated for ~60h in two separate samples containing no TFE (a) and 5% TFE (b). Spectra were acquired for protein in phosphate buffer (triangles, sp. 1), Congo Red (circles, sp.2) and protein in CR-phosphate buffer (squares, sp.3). The insets shows the respective difference spectra, obtained subtracting spectra 1 and 2 from spctrum 3.

Secondary structure of the aggregates was monitored using far-UV CD (Fig. 3.3). Spectra were acquired using two sets of protein samples, each at 0.1, 0.2 and 0.4 mg ml^{-1} in 5% (v/v) TFE. While the first set was pre-incubated for 48 hours, the other was tested immediately after the addition of TFE. The CD spectra of the aggregated protein are significantly different from those obtained immediately after the addition of TFE when the aggregates are not yet present (Fig. 3). This indicates that a conformational change is associated with AcPDro2 aggregation. The CD spectra of the aggregated protein display a single minimum at ca. 220-224 nm, typical of aggregates having β-sheet structure (135). The peak is generally at 215-218 nm, but undergoes a red-shift due to light scattering phenomena arising from the aggregates.

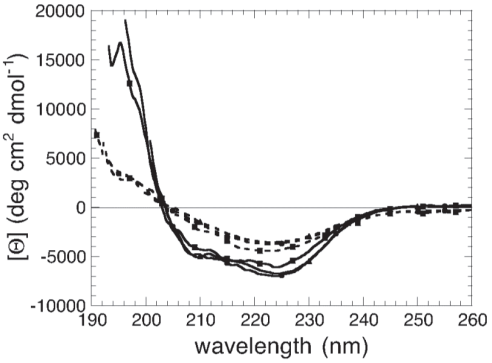


Figure 3.3 Far-UV CD spectra of AcPDro2 in 5% TFE at protein concentrations of 0.4 (triangles), 0.2 (circles) and 0.1 (squares) mg ml^{-1}. Spectra were acquired immediately after (solid lines) and 48h after (dashed lines) incubation under these conditions.

Protein samples incubated in 5% TFE were analysed using tapping mode atomic force microscopy (TM-AFM). Since the sample was dried to facilitate its adhesion to the mica substrate, the measured aggregate sizes reported below are reduced with respect to fully hydrated conditions, the shrinking factor being 2.0-2.5 as previously evaluated (129). After 24 h of incubation, aggregates with a twisted, fibrillar appearance are observed (Fig. 3.4a). The fibrils have a height of 2.8±0.2 nm, while at their ends the height is reduced to 1.8±0.1 nm. The typical fibril length is 50-200 nm. After 48 h of incubation, longer fibrils

(typically from 300 nm to 1 μm) are found (Fig. 3.4b). Twisted fibrils with a height of 3.6±0.1 nm result from the intertwining of thinner filaments whose diameter exhibits a bimodal distribution with peaks around 1.2 and 2.3 nm and mean value 1.9±0.1 nm. Given the shrinking factor of 2.0-2.5, the 3.6 nm high fibrils and their constituent 1.9 nm high filaments have widths consistent of amyloid fibrils and protofilaments, respectively. Overall, the aggregates formed from AcPDro2 analysed here appear to have morphological, structural and tinctorial properties typical of amyloid-like structures.

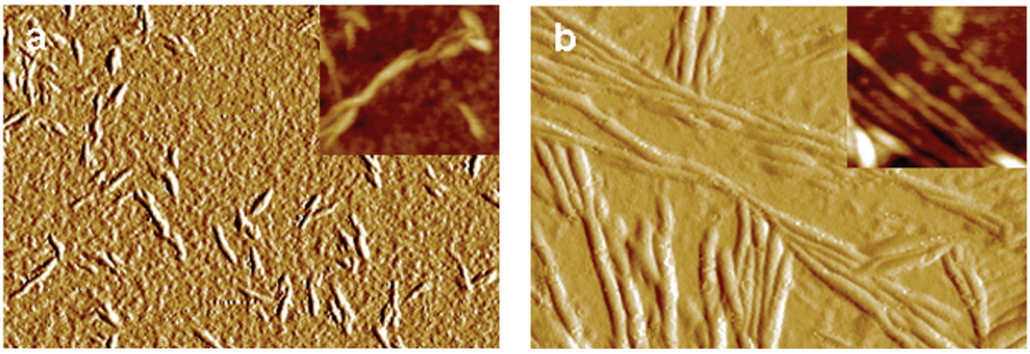


Figure 3.4 TM-AFM images (amplitude data; the insets show height data) of AcPDro2 aggregates obtained in 5% TFE after 24h (a) and 48h (b). Image sizes are (a) 1.0x0.8 μm, inset 380x320 nm; (b) 1.9x1.3 μm, inset 300x270 nm.The Z ranges for the height data reported in the insets are (a) 15 nm and (b) 10 nm.

3.3 AcPDro2 is initially in its native-like state

Next we investigated the conformational state adopted by AcPDro2 in 5% (v/v) TFE before the aggregates form. This was technically feasible due to the relatively slow kinetics of the process that allowed a biophysical investigation to be carried out before the aggregates accumulate significantly.

3.3.1 Early aggregation process monitored by Dynamic Light Scattering

Dynamic light scattering (DLS) measurements were performed on protein samples with and without 5% (v/v) TFE added immediately before the measurements. Fig. 3.5 shows two representative size distributions by light scattering intensity obtained at a protein concentration of 0.5 mg ml^{-1}, in 0% (solid line) and 5% TFE (dashed line). The peaks at 3.77±0.13 and 4.10±0.13 nm observed in 0 and 5% (v/v) TFE, respectively, are consistent with the hydrodynamic diameter of native and monomeric AcPDro2, as determined by the x-ray crystallographic structure (94) (errors reported in the manuscript indicate std. dev.). All pairs of measurements, carried out at different protein concentrations, show that the apparent hydrodynamic diameters measured in 5% (v/v) TFE are within error to those measured in the absence of TFE (Fig. 3.5, inset). This result led us to the conclusion that a monomeric form of AcPDro2, with a compactness similar to that of the native protein, is the most represented in solution before the aggregation process starts.

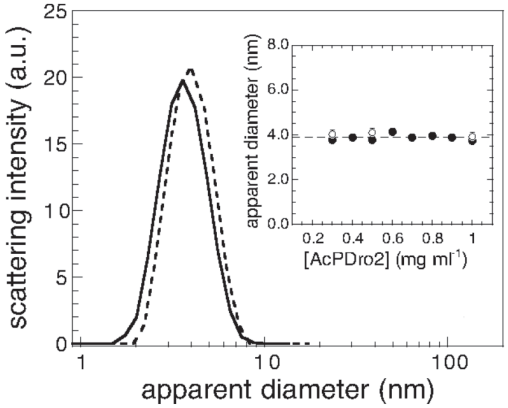


Figure 3.5 Size distibutions by light scattering intensity of AcPDro2 in 0% (solid line) and 5% (dashed line) TFE. The inset shows the apparent diameters measured at different protein oncentrations in 0% (filled circles) and 5% (empty circles) TFE.

The secondary structure content of AcPDro2 prior to aggregation was investigated using far-UV CD. The CD spectra of AcPDro2, recorded immediately after the addition of 1, 2, 3, 4 and 5% (v/v) TFE, are very similar to that obtained for the native protein in the absence of TFE. Fig. 3.6 shows that