UNIVERSITY OF MILANO - BICOCCA

Faculty of Mathematics, Physics and Natural Sciences
Department of Materials Science

PhD school in Nanostructures and Nanotechnologies – XXIII cycle

The aggregation mechanisms of amyloid proteins studied by mass spectrometry and other biophysical techniques

Author: Carlo SANTAMBROGIO
Tutor: Professor Rita GRANDORI

SUMMARY

INTRODUCTION

1. WHY BIONANOTECHNOLOGY ?

The field of nanotechnology deals with the control of matter at the molecular level to produce materials and devices organized on the nanometrical scale (10^{-9} m). In this dimensional range the statistical properties of an ensemble of molecules become less important, while quantum mechanical effects start to be relevant. This confers to the nanomaterial several properties (mechanical, electrical, optical, etc.) that are completely different with respect to conventional systems, opening new insights for the construction of useful devices.

Among several branches afferent to nanotechnology, the application of nanodevices for biological purposes (nanobiotechnology) and the use of biomolecules for the assembly of nanomachines (bionanotechnology) have been highly developed in the last ten years (Gazit 2007b) (Figure 1.1). Indeed, each part of a living system is organized at the nanoscale to carry on its function, suggesting the idea to exploit the opportunities offered by "Mother Nature" for the creation of nanodevices.
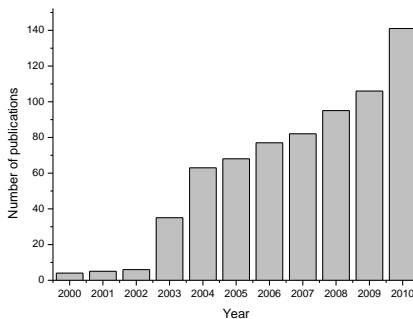

Figure 1.1 Number of papers regarding the word "bionanotechnology" as a function of the publication year (www.pubmed.org).

Molecules composing biological systems explicate their work thanks to a series of features that offer important advantages for applications in nanotechnology.

Biomolecules are designed through the genetic code: in living systems genes provide the necessary information for the biosynthesis of the entire organism. Since we know the code used by nature to create different proteins and enzymes, it is possible to exploit recombinant-DNA technology for several purposes. The most immediate advantage is represented by the ability to express in the ribosome the desired macromolecule, or to easily introduce a particular modification in a specific position (as a cysteine in a protein) to obtain a functionalized system. Moreover, it is possible to use genetic engineering in order to fuse a wide array of proteins with different features, in order to create with complete control a multi-active device (Scheibel et al. 2003). Finally, the specificity of complementary base-pairing renders DNA sequences highly programmable, suggesting the opportunity to be exploited for algorithmic nano-assemblies (Seeman 2010).

Biomolecules present sensitive molecular recognition: several activities in living organisms are regulated by specific interactions between molecules. Enzymes catalyze chemical reactions binding substrates in the active site, and their performances are affected by the interaction with cofactors; in the immune system antibodies identify their unique antigens among millions of molecules thanks to highly specific interactions; membrane proteins, through the mediation of ligand molecules, can connect internal and external environments of a cell. This deep affinity and high specificity between biomolecules can be employed for the construction of nanosensors with increased sensitivity, to improve targeted drug

delivery techniques, or to project diagnostic devices able to detect even small traces of a particular analyte (Cui and Gao 2003).

Biomolecules self assemble in ordered structures: proteins and peptides are often the building blocks of organized and extended structures used by living organisms to perform a protective function (like collagen and keratin in skin and nails), to assemble a matrix for mineral components (like collagen in bones) or to carry out cellular mobility (like actin in cytoskeleton) (Gazit 2007a). The ability of these molecules to spontaneously arrange in a precise and iterative fashion represents the heart of the bottom-up approach for nanotechnology (Shah et al. 2007). Periodic arrangements of proteins, peptides or nucleic acids present disparate features and shapes according to their particular composition, leading to countless applications that span from scaffold materials for electronic apparatuses to nanotubes or cages for storing devices (Figure 1.2).

Previous arguments can give an idea of the enormous opportunities offered by the development of this new discipline, justifying the increasing interest risen on the field both for pure knowledge and technological applications. Furthermore, the possibility to design molecular machineries with an atomic resolution promises a radical change in the concept of medicine, since many diseases could be approached at the molecular level with the employment of nanodevices for diagnostics and therapy. In this new branch, called nanomedicine, a prominent topic is represented by the study of protein amyloid fibrils, that are self-assembling nanostructures implicated in many serious disorders.
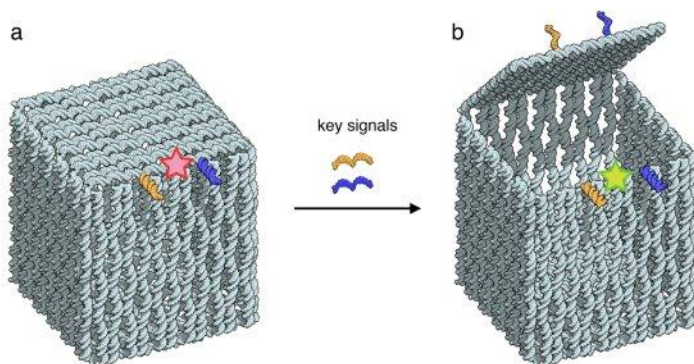
Figure 1.2 a) DNA box 42 x 36 x 36 nm with closed lid; b) the presence of DNA "keys" can open the lid (Andersen et al. 2009).

2. PROTEIN AMYLOID FIBRILS

Among several research-lines related to bionanotechnology, the study of protein amyloid fibrils is one of the most developed and promising, being a central topic in many scientific disciplines. This field, indeed, is extensively explored in medicine, due to the involvement of fibrillation in several diffuse diseases, like Parkinson and Alzheimer (Rahimi et al. 2008). Amyloid aggregation propensity is an intriguing research topic also for protein sciences, since it is emerging as a general and predictable property of polypeptide chains (Monsellier et al. 2008). More recently, protein fibrils have become interesting also for technological applications, thanks to particular features like their complex structure or their mechanism of assembly (Hamada et al. 2004).

General features of amyloid fibrils

Amyloid fibrils are protein aggregates (several micrometers long and a few nanometers wide) characterized by a well ordered and periodic architecture. Each fibril is usually composed of several protofilaments where proteins are disposed in cross β-strands oriented perpendicularly to fibril axis (Figure 2.1). These protofilaments typically assembly in a rope-like fashion or associate laterally in ribbons (Makin and Serpell 2005; Saiki et al. 2005). Even if many living organisms exploit these structures to perform biological functions (Kenney et al. 2002), amyloid fibrils are notorious for their intra- or extracellular accumulation that causes several human diseases (Sideras and Gertz 2009). Proteins related to these pathologies are extremely heterogeneous in their length,

sequence and structure (Table I), and they can produce fibrils *in vitro* under mildly denaturing conditions, suggesting the idea that the ability to form amyloid aggregates under appropriate circumstances is a property shared among all proteins. An increasing number of non pathological proteins, indeed, has been shown to assembly into fibrils *in vitro*, even if the propensity of the polypeptide chain to assume the fibrillar form is highly dependent to the sequence and the conditions employed (Stefani and Dobson 2003). Moreover, each protein can give origin to fibrils with distinct morphologies, depending on the environmental circumstances where the process occurs (Pedersen et al. 2006). These considerations underscore the dramatic complexity of amyloid aggregation and the strong dependence of the entire process to a large number of factors.
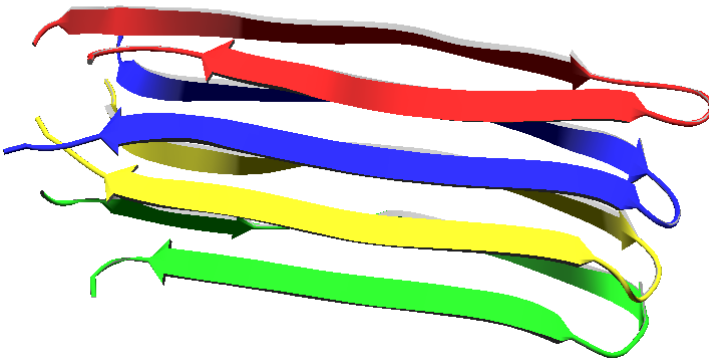

Figure 2.1 General structural motif of human CA150.WW2 amyloid protofilament (PDB ID code 2NNT). The repeat unit of 26 residues in a non-native beta hairpin was revealed by alanine scanning, NMR spectroscopy and electron microscopy data (Ferguson et al. 2006).

Factors affecting amyloid fibril formation

Some intrinsic properties of the polypeptide chain are determinant for the propensity, the rate and the stability of amyloid formation. It was shown by mutagenesis experiments that a high content of hydrophobic residues, a low net charge and the tendency to form β-sheet secondary structures encourage the transition to the fibrillar form (Chiti et al. 2002a; Chiti et al. 2002b). The knowledge of these parameters allowed the prediction of regions promoting aggregation and the ones involved in the fibrillar core for proteins like α-synuclein and Aβ (Chiti et al. 2003). Moreover, specific sequences can promote aggregation, like in the class of poly-Q proteins where a series of consecutive glutamines is the principal agent that induce fibrillation (Saunders and Bottomley 2009).

protein	residues	native fold	disease
Aβ-peptide	40-42	Unfolded	Alzheimer
α-synuclein	140	Unfolded	Parkinson
prion protein	253	Unfolded + α-helical	Spongiform encephalopathies
superoxide dismutase 1	153	All-β, Ig like	ALS
huntingtin	3144	Largely unfolded	Huntington
fragments of serum amyloid A protein	76-104	All-α, unknown fold	Mediterranean fever
mutants of lysozyme	130	α+β, lysozyme fold	Lysozyme amyloidosis
amylin	37	Unfolded	Type II diabetes
insulin	21+30 (2 chains)	All-α, insulin like	Injection localized amyloidosis
lactoferrin	692	α+β, periplasmic-binding protein like II	Corneal amylodosis
γ-crystallins	Variable	All-β, γ-crystallin like	Cataract

Table I List of some proteins associated with human amyloid diseases (Chiti and Dobson 2006).

Besides the physico-chemical properties of the amino acidic chain, amyloid assembly is crucially regulated by the structure assumed by the whole protein, or at least by some of its sections. Even if the aggregation from the folded state was recently discussed (Chiti and Dobson 2009), a significant degree of disorder seems to be necessary for fibril formation. Indeed, in different prion proteins containing both unstructured and globular domains, the former were shown to be responsible for aggregation (Baxa et al. 2006). Natively unfolded proteins have evolved in sequences less hydrophobic and more charged than globular proteins to contrast their high tendency to form fibrils (Uversky 2002); nevertheless, they are often involved in amyloidosis (Uversky 2009). Many (if not all) globular proteins with a permanent fold under non denaturing conditions can be converted to the aggregated state by a partial destabilization of their conformation (Uversky and Fink 2004). The change in protein structure necessary for the process can be obtained employing particular environmental factors – like organic solvents, high temperature, high pressure and acidic pH – or by the interaction with small molecules, especially metal ions (Eakin and Miranker 2005). On the other hand, when the binding of the protein with some macromolecules stabilizes its fold, like in the case of the complex with particular antibodies (Dumoulin et al. 2003) or chaperones (Nerelius et al. 2010), the protein is also inhibited in the formation of the fibrils.

Mechanism of amyloid fibrillation

In spite of the particular details typical of each protein, the most appropriate way to describe the amyloid aggregation process is by a nucleation-elongation model (Bhak et al. 2009). During nucleation, the protein undergoes a series of conformational changes and

oligomerizations in order to form the nuclei – the oligomers that act as templates for the fibrils growth – and the "active species" – the species that represent the building blocks of the aggregate. During the elongation, these species rapidly assembly on the nuclei until the fraction of protein in the soluble phase is consumed to a certain level. At this point the fibrils' growth decelerates and the amount of aggregated protein remains constant.

The first part of the mechanism is usually called lag phase, since no fibrillar material is formed till that point. The lag phase can be substantially reduced or even abolished by adding preformed aggregates ("seeding") to the sample under fibrillation conditions; the proteins that compose the seeds can be different from the one involved in the aggregation, even if the efficiency of the seeding is strongly dependent on the similarity between their sequences (Krebs et al. 2004). The features of nuclei and active species can vary dramatically according to the proteins and the conditions adopted. As mentioned in the previous section, a typical situation develops with the partial unfolding of the monomers, the subsequent growth of non-native oligomers till the dimensions of the nucleus are reached, and the elongation of the fibril by monomers or oligomers. However, there are several exceptions to this scenario: in some cases the aggregating protein does not need to unfold (Chow et al. 2004a), the active species is a native-like oligomer with enzymatic activity (Plakoutsi et al. 2005), the conformational changes occur only after the elongation (Kumar and Udgaonkar 2009), or the native monomer structure is retained in the fibril (Bousset et al. 2002).

These differences in the features of the aggregation mechanism are not surprising if the heterogeneous ensemble of protein structural states and the relative network of equilibria are considered (Figure 2.2). The normal pathway followed by a protein – starting from biosynthesis, passing to the functional conformation and finishing with its degradation – is usually composed by several transient

intermediate states that are necessary steps to reach the desired result but are occasionally prone to amyloid aggregation. Even if the process is highly controlled, a small change in the properties of the system can be sufficient for the accumulation of these species and the deflection of the pathway versus the formation of fibrils. The polymorphism of the amyloids and the differences in their assembly mechanism are probably due to the broad ensemble of conformers populated by the protein in each condition, and to the particular intermediate states responsible of the aggregation (Kodali and Wetzel 2007). The deep understanding of the fibrillation process, therefore, passes through the identification, among the labyrinth of the possible pathways and species, of the specific conformational and oligomeric states involved in the assembly, as well as through a kinetic and thermodynamic characterization of the different aggregation steps. Moreover, these intermediate species are also interesting because, for proteins involved in neurodegenerative diseases, they are emerging as the toxic agents of the pathology rather than mature fibrils (Caughey and Lansbury 2003). It is therefore evident that the characterization of these intermediates is crucial for the medical therapy: identifying the environmental factors that trigger their accumulation and understanding their structural properties is fundamental for the design of new specific drugs or for the development of novel therapeutic strategies.
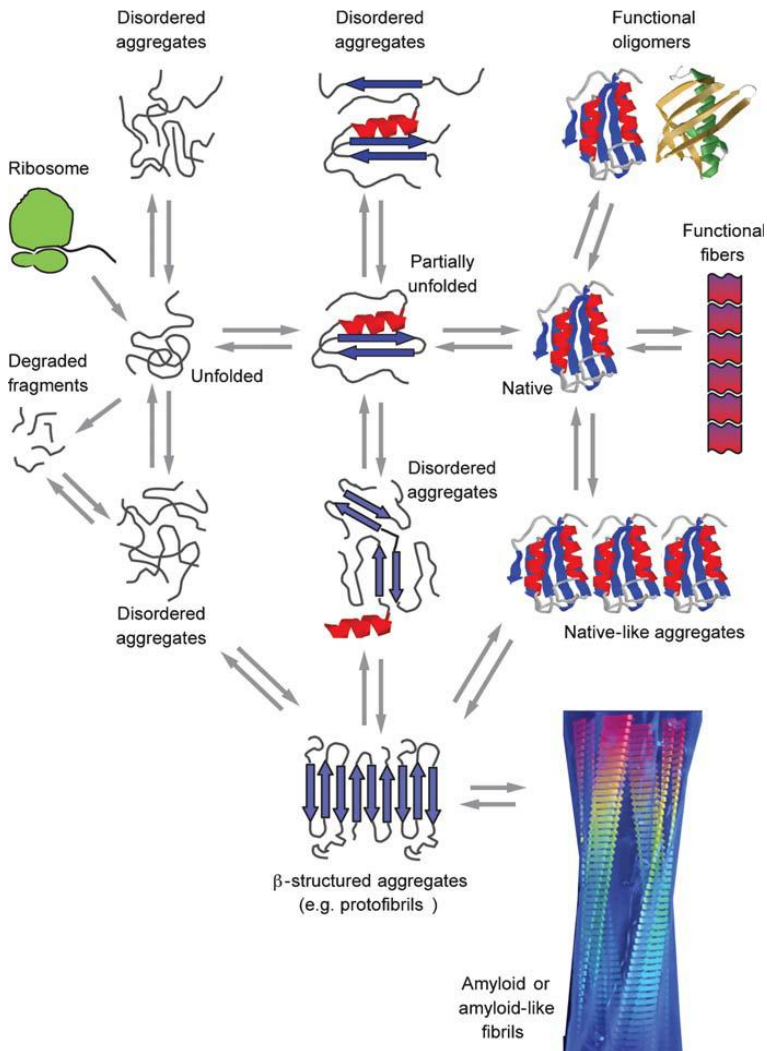
Figure 2.2 A schematic view of the possible states adopted by a polypeptide chain and their relative interconversions (Chiti and Dobson 2006).

Nanotechnological applications

Even if the field of amyloid fibrils is relatively young and several aspects of the assembly process are not known, the employment of these structures for technological purposes has already led to some important results. As previously mentioned, the self-assembly in uniform architectures is a fundamental requirement for the bottom-up approach, and the possibility to control fibril formation simply by changing the reaction conditions is therefore a crucial and challenging point for their technological applications.

First of all, amyloid fibrils were shown to possess mechanical properties comparable to that of steel and silk (Smith et al. 2006b), making them highly attractive components for the construction of resistant materials with high biocompatibility. Another opportunity offered by amyloids is their employment as templates for conducting molecules in order to form electronic nanowires (Scheibel 2005): the fibril acts as a scaffold where inorganic components – like golden atoms – are guided to assembly; the site of metallization can be created, for example, with the insertion by genetic engineering of a cysteine residue in the protein sequence. It is also possible to use amyloid templates on which to immobilize enzymes, without the lost of activity (Pilkington et al. 2010), or to attach on them a fluorescent protein (Baxa et al. 2002) with the intent to create an optical nanosensor. Finally, the ability of amyloid aggregates to form liquid crystalline phases (Corrigan et al. 2006) can open new routes in the assembly of novel biomaterials.

In this thesis, the study of the amyloidogenic mechanism for three different proteins – β2-microglobulin, ataxin-3 and β-lactoglobulin – will be presented. In particular, the early stages of the process will be analyzed by several biophysical methods in order to shed light on the conformational changes and the oligomeric assemblies involved.

3. TECHNIQUES EMPLOYED

Different experimental tools have been used in the present work to analyze protein structural features and to investigate their aggregation processes. Besides the intrinsic power of each biophysical method used, interesting information is often obtained by a synergic employment of different techniques, thanks to the complementarity of their results. For instance, the possibility to achieve both a local and a global information, or to test both average or species-specific properties of a molecular ensemble can contribute significantly to the development of a deeper knowledge of complex molecular systems.

Nano-electrospray ionization mass spectrometry (nano-ESI-MS)

Nano-ESI-MS (Mann and Wilm 1995) is becoming a widely employed technique in biochemistry and structural biology (Grandori et al. 2009), thanks to the opportunity to preserve both covalent and non-covalent interactions while generating gas-phase ions, avoiding in this way dissociation and denaturation of biopolymers (Chowdhury et al. 1990; Katta and Chait 1991). In a nano-ESI-MS experiment, ions of the analyte are transferred to the gas-phase by spraying a liquid sample from a metal-coated capillary in the presence of a strong electric field. Charged droplets, ions and ion clusters are driven through the instrument interface, where they are focused, desolvated and declustered. The process that leads from the liquid sample to gas-phase ions consists of several cycles of solvent evaporation and explosion of the droplets generated by electrospray. Indeed, the stability of a charged droplet is the result of two

contrasting forces: the surface tension, that promotes droplet cohesion, and Coulomb repulsions that tend to dissociate it. During desolvation, the latter force increases due to the increase of surface charge density. When the surface tension cannot balance electrostatic repulsions any more (the Rayleigh limit), the droplet explodes (Coulomb fission) in smaller droplets. Each daughter droplet is then below the Rayleigh limit and can undergo a new cycle of evaporation and explosion. The exact mechanism that brings to the final gas-phase analyte ion is not completely understood. In the "Charged-residue model" the evaporation-dissociation cycle continues until the droplet eventually contains only one analyte ion (Dole et al. 1968). Complete solvent evaporation from such a droplet would produce a gas-phase analyte ion On the other hand, in the "Ion evaporation model" a charged molecule is ejected from the droplet before its complete desolvatation (Iribarne and Thomson 1976). It is thought that the charged-residue model more closely describes the process in the case of proteins (de la Mora et al. 2000; Verkerk and Kebarle 2005; Kebarle and Verkerk 2009).

Numerous instrumental parameters control the ionization/desolvation process and have a critical importance in determining spray stability and spectrum quality. Most importantly, they can also affect the structure of proteins and protein complexes, and should be carefully optimized to avoid artefactual results. Among the most critical parameters are curtain gas flow rate, interface temperature and declustering potential. The curtain gas is a nitrogen stream that prevents entrance of neutral contaminants into the spectrometer and assists desolvation of ions. The collisions with gas molecules can result in protein unfolding and complex dissociation (Schneider et al. 2001; Sanglier et al. 2008). The instrument interface (IHT) can be heated to improve desolvation. Although it cannot be assumed that thermal equilibration is reached, by increasing the value of this parameter above room temperature it is possible to induce protein

unfolding and/or complex dissociation (Invernizzi et al. 2006; Natalello et al. 2007). Finally, the declustering potential determines the acceleration of ions at the entrance of the instrument, where collisions with gas molecules are still possible. This parameter is important for the removal of residual solvent molecules and disaggregation of clusters. Also in this case, too high values can result in protein unfolding and dissociation (Careri et al. 2004; Sanglier et al. 2008), and in the alteration of the protonation states (Fligge et al. 1998). As a general criterion, it is necessary to identify ranges of values in which none of these parameters becomes the limiting factor for protein stability during the electrospray process. After the desolvation and ionization steps, ions are transmitted to the mass analyzer, where they are sorted according to their mass-to-charge ratios (m/z). Quadrupoles, ion traps, cyclotron resonance traps and time-of-flight (TOF) tubes are the most employed analyzers in MS instruments. The spectrometer used here is a hybrid quadrupole-TOF analyzer: a first quadrupole, Q_0, focuses the ion beam in a second quadrupole, Q_1, where ions within a given m/z range are selected and transmitted, through the last quadrupole Q_2, to the TOF module. Even if not employed in this work, MS/MS measurements are possible by selecting a specific m/z value in Q_1 and performing its fragmentation in Q_2, which can operate as a collision cell. After the final quadrupole, the ions are pushed into the field-free TOF tube by an electrostatic plate that confers a specific kinetic energy to all of them, depending only on the applied potential and the ion charge. Then, the velocity of an ion inside the drift tube is determined only by the applied potential and its m/z value. Therefore, the ions are separated and reach the detector at different times after the pulse. To increase resolution, an electrostatic mirror (a reflectron) is employed in order to double the ion path (Benesch et al. 2007). The final result is presented in a mass spectrum, that

reports the (m/z) values of the ions on the *x*-axis and the relative intensities (ion counts) on the *y*-axis (Figure 3.1).

Proteins give rise to a bell-shaped series of peaks in the MS spectrum, called a charge-state distribution (CSD) (Kaltashov and Abzalimov 2008). Each peak of such envelopes represents a distinct charge state and differs from the neighboring ones by one charge. The CSD of a given protein is affected by several factors, including instrumental setting and solvent properties (Mirza and Chait 1994), but the most dramatic influence is exerted by the protein's conformational features (Chowdhury et al. 1990; Kaltashov and Abzalimov 2008). Indeed, the extent of ionization depends on the compactness of the three-dimensional structure: while folded proteins are characterized by narrow CSDs (a few peaks) at relatively high m/z values (i.e. low charge states), unfolded proteins give rise to broad CSDs (tens of peaks) shifted towards much smaller m/z values (i.e. high charge states), relative to the same protein under non-denaturing conditions.
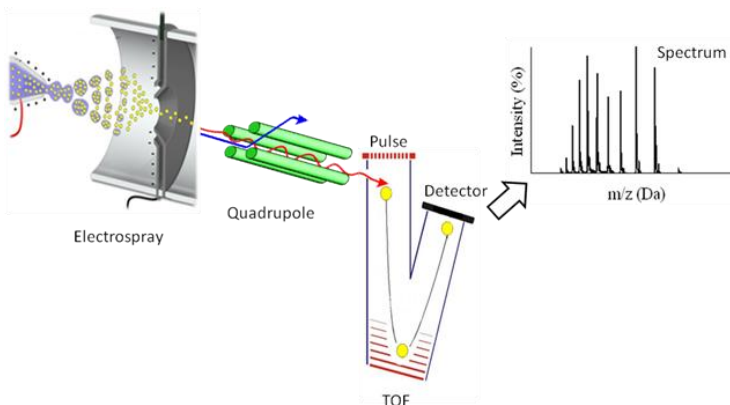

Figure 3.1 A representation of ion travel inside the spectrometer. Molecules are desolvated and charged during the electrospray, then focused and directed in the quadrupole. Here the ions are sorted and driven towards the TOF, where they are separated and then detected. The result is presented as plot of relative intensity (normalized ion counts) versus m/z values.

The lower average charge state of folded proteins reflects their higher structural compactness. On the other hand, the narrower CSDs of folded proteins reflect their higher conformational homogeneity. (Konermann and Douglas 1997; Grandori 2003a; Kaltashov and Mohimen 2005).

Thanks to the fact that protein CSDs are sensitive to the global compactness of the protein structure, nano-ESI-MS also offers the possibility to detect folding intermediates (Konermann and Douglas 1998; Kaltashov and Eyles 2002). Compactness represents a good coordinate for folding reactions, while secondary structure or local tertiary features do not always discriminate effectively partially folded states from both native and denatured states. Multimodal distributions are generated by the coexistence of different conformational states, reflecting distinct molecular species. While the CSDs of a given protein in its folded and fully unfolded states are generally well resolved, partial overlap of CSDs can be observed in the presence of folding intermediates. Deconvolution algorithms have been developed in order to assess species distributions in conformationally complex mixtures (Dobo and Kaltashov 2001; Borysik et al. 2004). This possibility to detect distinct conformational states, free from averaging over the molecular populations, is particularly attractive. Indeed, integration of signal intensity over all the charge states of a given envelope offers a quantitative estimation of the apparent fraction of molecules in the corresponding conformational state. Unfortunately, depending on the experimental conditions, unfolded proteins can sometimes display stronger signal responses – called "relative transmission factor" – in nano-ESI-MS than folded ones, making quantitative analysis of species distributions by this method not generally applicable (Kuprowski and Konermann 2007).

Binding properties can also be studied by nano-ESI-MS under non-denaturing conditions (Heck and Van Den Heuvel 2004; Akashi

2006). For such applications, it is required that non-covalent interactions survive not only the desolvation process but also ion beam focusing and transfer to the flight tube. Fast growing evidence shows that this is the case for supramolecular complexes of very different nature, including protein-protein, protein-nucleic acid and protein-small ligand interactions (Ilag et al. 2005; Kaltashov et al. 2006; Gordiyenko and Robinson 2008). Non-covalent complexes can be detected by nano-ESI-MS thanks to the mass shift related to the transition from the free to the bound species. Protein oligomerization represents a peculiar case (Šamalikova et al. 2005; Mendes et al. 2005). Since this process involves association of identical subunits, different oligomer sizes could give rise to the same m/z values. Nevertheless, some peaks will be specific for a given stoichiometry, like for instance odd charge states of dimers, which would imply a non-integer number of charges for the monomer. Furthermore, CSDs are progressively shifted to the high m/z region of the spectrum as oligomer size increases. This is because the same subunit acquires fewer charges as oligomerization proceeds, likely by the same mechanisms that reduce charging as a consequence of folding. Since charges of free subunits report on folding and the mass of ions reports on binding, nano-ESI-MS provides useful, combined information on the association and conformational state of proteins. It is possible to evaluate whether the free form coexists with the bound form as a folded subunit, or whether folding is induced by binding. This kind of analysis is particularly useful for protein-small ligand interactions, where the CSDs of the bound forms can also be easily interpreted in terms of conformational properties (Osz et al. 2007). Thanks to the capability to detect the distinct molecular and supramolecular species in heterogeneous samples, nano-ESI-MS offers interesting possibilities for the analysis of complex systems like those governed by linked binding equilibria. The invaluable contribution to binding analysis that can be derived in these cases is

the identification of the different species involved, which is given as an assumption in the model-dependent studies based on conventional methods. Examples described in the literature include protein oligomerization induced by cofactor binding (Ayed et al. 1998; Natalello et al. 2007) and the stepwise assembly of large protein complexes like chaperones (Donald et al. 2005). Unfortunately, several problems prevent the general applicability of nano-ESI-MS to quantitative analysis of the pre-existing equilibrium in solution. Dissociation under electrospray conditions can occur to some extent, particularly for complexes that are mainly stabilized by hydrophobic interactions in solution (Clark and Konermann 2004; Peschke et al. 2004). On the other hand, non-specific complexes can form (Wang et al. 2005; Sun et al. 2007) particularly at analyte concentrations above 10^{-4} M. Despite these problems, nano-ESI-MS offers highly detailed qualitative information on the molecular and supramolecular species populating heterogeneous samples.

Circular dichroism (CD)

Circular dichroism is widely recognized as a useful spectroscopic tool to characterize protein structure and to study protein stability. It is based on the difference in absorption by the sample of circular polarized radiations with opposite signs – clockwise and counter-clockwise. This effect occurs when the chromophores populating the solution are intrinsically chiral or are bound to a chiral centre. The sample presents a CD activity also when the chromophores are placed in an asymmetric environment as a result of the structure adopted by the molecules or of a binding reaction. In a classical CD instrument, the light coming from a xenon lamp is linearly polarized by the passage through two prisms and is focused on a piezoelectric quartz crystal subjected to an alternating electric field. In this way

the crystal splits the light into its circular polarized components that are alternatively transmitted through the sample before the detection by a photomultiplier. A CD spectrum is finally obtained by plotting the dichroism signal as a function of the radiation wavelength (Kelly et al. 2005). In order to remove oxygen molecules from the different compartments, a stream of nitrogen gas flows inside the instrument. This stream prevents the damage of the optical system by ozone and allows spectral measurement in the deeper far-UV region. By changing the optical path it is possible to test a very wide range of protein concentrations (from μM to mM), providing that buffer absorbance is not too high. Moreover, CD measurements can be easily performed in extreme experimental conditions, such as very acidic or basic pH, and high temperatures, making CD a fast and convenient technique for protein studies, especially in testing conformational changes (Kelly and Price 2000).

Each region of a CD spectrum reflects prevalent absorbance of a particular kind of chromophore and thus gives specific information on the composition of the sample. For this reason, CD data are usually presented by different normalizations of a common unit, according with the fundamental absorption element in the wavelength range studied. The most adopted unit is the ellipticity *(θ)* – defined as $tan^{-1}(b/a)$, where *b* and *a* are the minor and major axes of the elliptical polarized light resulting by the combination of transmitted clockwise and counter-clockwise radiations. CD measurements are usually presented in molar ellipticity, *[θ]*, expressed in the units $deg \cdot cm^2 \cdot dmol^{-1}$. In far-UV (170-260 nm) the peptide bond is the main responsible of absorption, due to a weak but broad n→π* transition between 200 and 220 nm and a more intense π→π* transition around 190 nm (Greenfield 1999). This comports the presentation of the data in mean residue ellipticity, *[θ]*$_{mr}$ – the ellipticity divided by the molar concentration, the optical path and the number of amino acid residues, all expressed in $deg \cdot cm^2 \cdot dmol^{-1}$.

The dependence of far-UV optical activity by the peptide bond geometry leads to characteristic CD spectra for the different types of regular secondary structure present in proteins (Figure 3.2): the α-helix motif results in an intense negative band with two peaks (at 208 and 222 nm) and in a strong positive band (at 191 - 193 nm); the β-sheets present a broad negative band at 210-225 nm and a positive peak around 190-200 nm; random coils have a negative band at 195 - 200 nm (Martin and Schilstra 2008). The knowledge of the CD profile for each motif of secondary structure permits the estimation of the secondary structure content in a protein by different algorithms – like CONTILL, CDSSTR or SELCON – that are programmed to compare the measured spectrum with a database of protein CD spectra (Sreerama and Woody 2000). The database is usually composed of distinct groups of globular, membrane and disordered proteins.
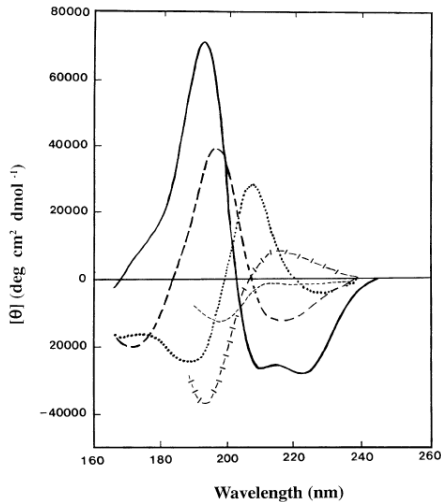
Figure 3.2 Far-UV CD spectra associated with various types of secondary structure. Solid line, α-helix; long dashed line, anti-parallel β-sheet; dotted line, type I β-turn; cross dashed line, extended 3_1-helix or poly (Pro) II helix; short dashed line, irregular structure (Kelly et al. 2005).

Unfortunately, the characterization of protein secondary structure with these methods is not highly reliable. First of all, it must be observed that CD is a low resolution technique: only the average content of secondary structure motifs can be estimated, without any indication on the specific structure of different regions. Moreover, other chromophores distinct from the peptide bound can interfere with the signal in far-UV, as demonstrated for aromatic residues (Woody and Dunker 1996; Krittanai and Johnson 1997). Finally, a spectrum including the deep far-UV region (170-190 nm) is necessary for an accurate secondary structure estimation, but the significant absorption of several buffers at these wavelengths makes difficult to measure much below 200 nm, unless by non-conventional instruments (Wallace 2000).

In the near-UV (250-320 nm), protein CD signal is principally due to aromatic amino acids, with a weak contribution of disulphide bond. Phenylalanine presents a peak between 255 and 270 nm, tyrosine between 275 and 282 nm, and tryptophan between 280 and 290 nm (Strickland 1974). The side chain of aromatic residues is particularly sensitive to the asymmetric environments characteristic of the protein tertiary structure. Nonetheless, it is not generally possible to extract any detailed information on the folded state by a near-UV CD spectrum, since several factors influence the shape and the intensity of the signal, like the number and the relative positions of each type of aromatic amino acid, their mobility, and the nature of the environment in terms of hydrogen bonding and polarization. The near-UV CD bands from individual residues can be either positive or negative, thus a near-zero spectrum, usually typical of proteins in the unfolded state, can also be the result of mutual cancelling contributions (Griffin et al. 1972). On the other hand, the near-UV CD spectrum under non-denaturing conditions represents a fingerprint of protein native tertiary structure, useful to be compared with the spectra obtained in other circumstances, like under

unfolding conditions or after mutagenesis (Krell et al. 1996). In particular, the presence of molten globule conformations can be easily confirmed by near-UV CD, since they are characterized by very weak signals reflecting the high mobility of aromatic side chains (Ptitsyn 1995). Near-UV and visible are also interesting regions to monitor protein-ligand interactions. Indeed, many cofactors, like pyridoxal-5'-phosphate, heme or flavins, present no signal when free in solution, but show a band when bound to their protein partner, as a result of the acquired asymmetric environment (Price and Stevens 1983). This feature allows one to monitor the interaction as a function of ligand concentration, or to verify the integrity of the binding site in several experimental conditions. Unfortunately, CD has no specific sensitivity to protein quaternary structure. However, protein-protein interactions can be studied by CD in some particular cases, such as when the binding comports a modification in the structure of the protein subunits, or when the oligomeric interface affects the environmental symmetry of some aromatic residues (Greenfield 2004).

Fluorescence spectroscopy

Fluorescence spectroscopy refers to the detection of the light radiated by a sample after its excitation by electromagnetic radiation. A schematic representation of electronic transitions involved in the fluorescence emission process is given in Figure 3.3 by traditional Jablonski diagrams (Lakowicz 1987). The absorption of light by a molecule excites some of its electrons from the ground state to one of the vibrational level in a higher singlet state. This transition is estimated to be completed in $\sim 10^{-15}$ s. Then, internal conversion – that is the jump from an excited state to a less energetic one – and vibrational relaxation – that is the shift, within an excited state, to the

lowest vibrational level – occur in picoseconds, leading to the population of the lowest level in the first excited state. This is the starting level for the back transition to the ground state, that can be obtained by emission of fluorescence or by the loss of excitation through different dissipating phenomena, like intersystem crossing to the triplet state or other non-radiative mechanisms. The partial loss of energy between absorption and emission comports a shift towards higher wavelengths of the fluorescence signal relative to the absorption peak (Stokes shift). Besides vibrational relaxation and internal conversion, solvent relaxation is another process that comports energy loss before the emission of light, contributing to the red shift of the fluorescence signal relative to absorption (Lakowicz 1987).
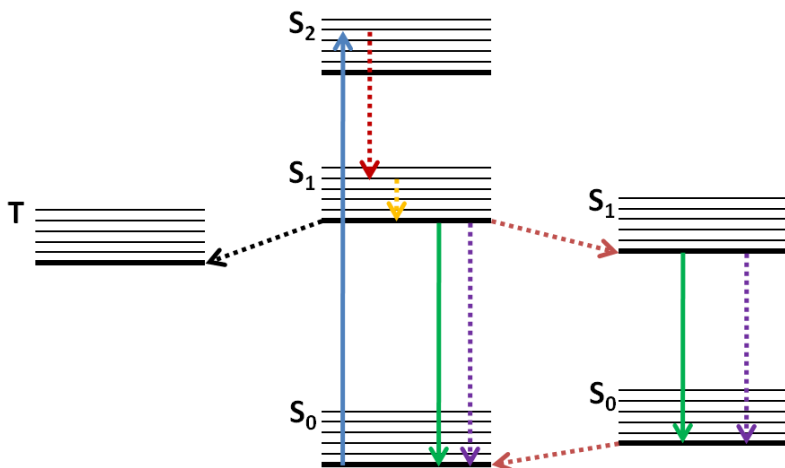

Figure 3.3 Jablonski diagrams of the electronic transitions involved in fluorescence emission. Absorption (blue line); internal conversion (red); vibrational relaxation (yellow); fluorescence (green); solvent relaxation (pink); intersystem crossing (black); non-radiative dissipation of energy (violet); ground state (S_0); first excited state (S_1); second excited state (S_2); triplet state (T).

Solvent relaxation consists in a reorganization of solvent molecules around the excited site in order to reach a more favorable energy configuration. This leads to a smaller gap between the energy levels and consequently to an increase of the emission wavelength. Since the effect is highly affected by solvent polarity, the emission maximum in a fluorescence spectrum gives information on the environment surrounding the probe (Lakowicz 1987).

As mentioned before, only part of the absorbed energy is released by fluorescence. The fraction of the total energy that a fluorophore emits by radiation is called quantum yield. This parameter is strictly dependent to the quenching or energy transfer effects exerted by particular molecules or groups close to the fluorescent site. Therefore, the intensity of the fluorescence signal is another indicator of the fluorophore neighborhood (Lakowicz 1987).

In the study of proteins, fluorescent probes can be divided in the two main classes of intrinsic and extrinsic dyes. Aromatic amino acids are the only intrinsically fluorescent residues of the polypeptide chain. Among them, tryptophan is widely the most useful and employed in fluorescence spectroscopy. Indeed, phenylalanine and tyrosine present a much lower quantum yield and are less sensitive to environmental changes. For this reason, the excitation wavelength is usually fixed at 295 nm, in order to selectively excite tryptophan residues and obtain a specific local information, exploiting the shortage of this amino acid in protein sequences (Royer 2006). In the absence of tryptophan residues in protein primary structure, a common approach is to make a tyrosine → tryptophan substitution, in order to confer a fluorescence activity with a minimal change in protein fold and stability (which must be confirmed experimentally) (Scalley et al. 1997). The same strategy can be employed to monitor the unfolding in distinct parts of a multi-domains protein by tryptophan mutagenesis at various positions (Hannemann et al. 2002).

Due to its hydrophobic character, tryptophan tends to be buried in the protein core or to be placed in the interface between domains. In the native fold tryptophan emission maximum is therefore relatively blue-shifted, ranging from 309 nm in an extremely protected environment (Szabo et al. 1983) to the more common 335 nm in a partially exposed state (Royer et al. 1993a). On the other hand, a tryptophan fully exposed to water has a fluorescence peak at 355 nm (Declerck et al. 2001). An increment in the tryptophan emission position under particular experimental conditions – as in the presence of denaturing agents – is therefore usually interpreted as a destabilization of protein structure.

Unlike the emission maximum, the unfolding process does not give a unidirectional response in the tryptophan fluorescence intensity. Due to solvent quenching effects, the exposition to the environment of tryptophan residues typically leads to a decrease in their emission signal (Altekar 1977). However, several amino acid side chains, the peptide bond, and some prosthetic groups – like heme groups or nicotinamide adenine dinucleotide (NADH) – are often similarly efficient quenchers of tryptophan fluorescence (Chen and Barkley 1998; Adams et al. 2002). In the case where a strong quenching acts in the folded state, the displacement of tryptophan residues from the protein core during unfolding can result in an increase of fluorescence intensity (Royer et al. 1993b). Otherwise, solvent quenching dominates and the common decrease in intensity upon denaturation occurs (Roumestand et al. 2001). It is therefore important to establish the behavior of proteins case by case, performing experiments of thermal or chemical denaturation (Di Stasio et al. 2004).

A different approach consists in the covalent attachment of an extrinsic dye to the protein chain, typically by the ε-amino group of lysine, the α-amino group of the N-terminus, or the thiol group of cysteine. However, non-covalent extrinsic dyes are match more used,

exploiting the change of the fluorescence signal upon hydrophobic or electrostatic interactions with the protein occur (Hawe et al. 2008). Among the different available probes, 1-anilinonaphthalene-8-sulfonate (ANS) and Thioflavin T (ThT) are two of the most widely employed for the characterization of protein structures and assemblies. In polar solvents, free ANS exhibits minimal emission of fluorescence, as a consequence of a charge transfer reaction in the excited state (Das et al. 1992). On the other hand, hydrophobic and electrostatic interactions can promote binding of ANS to proteins. Decrease in polarity of the ANS environment results in an increase of the fluorescence signal and a blue-shift of the emission maximum. For this reason, ANS has been widely employed to study partially folded structures (Fink et al. 1994) and to monitor protein-ligand interactions (Santos and Spector 1972).

ThT fluorescence intensity is principally affected by solvent viscosity and rigidity of its surrounding environment (Voropai et al. 2003). In the presence of amyloid fibrils ThT becomes highly fluorescent, with a maximum emission around 480 nm. Even if the mechanism of interaction is controversial (Khurana et al. 2005; Krebs et al. 2005), this feature is extensively used to detect the progressive accumulation of protein fibrils by monitoring ThT fluorescence signal (Ban et al. 2003). However, to induce ThT fluorescence is not an exclusive property either of amyloid aggregates or of β-sheet structures (Groenning et al. 2007). Therefore, the presence of fibrils cannot be assumed only on the basis of ThT experiments.

In conclusion, the high sensitivity to their environment of both intrinsic and extrinsic fluorescent probes permits a fast characterization of protein conformational properties in a wide range of experimental conditions and with a minimal consumption of material.

4. BACKGROUND OF ANALYZED PROTEINS

Three proteins with different features have been analyzed in this work. The first one, human β2-microglobulin, is a protein related to a misfolding disease affecting people undergoing hemodialysis (Dialysis Related Amyloidosis). The second protein, human ataxin-3, was chosen among the poly-glutamine family and its amyloidosis is responsible of a neurodegenerative pathology, the spinocerebellar ataxia type-3 (SCA3). The last one, bovine β-lactoglobulin, is a protein that is not related to any known amyloidosis, but can be induced to aggregate *in vitro* under appropriate conditions.

Human β2-microglobulin

β2-microglobulin (β2m) is a 99-residue β-sandwich protein with a typical seven-stranded immunoglobulin fold, whose two facing β-sheets are held together by a disulphide bridge buried in the protein core (Figure 4.1). β2m is the light chain of the class-I major histocompatibility complex (MHC-I) (Wearsch and Cresswell 2008). The *in vivo* turnover of the complex results in the release of β2m to the blood and its degradation in the kidneys. Dialysis Related Amyloidosis (DRA) is a pathological state afflicting patients suffering from kidney failure treated by long-term hemodialysis, related to inefficient β2m catabolism (Floege and Ehlerding 1996; Menaa et al. 2008). Under such pathological conditions the concentration of circulating β2m increases by a factor of 10–50 (Floege and Ehlerding 1996; Drueke 2000), leading to β2m amyloid fibrils that deposit principally in the skeletal joints, forming deposits of hundreds of grams (Gejyo et al. 1985) that hamper normal joint

functionality. *In vivo*, β2m normally circulates at ~0.1 μM concentration (Floege and Ehlerding 1996). However, β2m is soluble and stable *in vitro* at mM concentrations (Okon et al. 1992; Morgan et al. 2001) and can be reversibly unfolded (Chiti et al. 2001b; Kameda et al. 2005). Several conditions that enable *in vitro* β2m amyloid formation have been found. These include partial denaturation by acids (Kad et al. 2003), proteolytic treatment, and addition of various chemicals such as sodium dodecyl phosphate (Yamamoto et al. 2004a) and trifluoroethanol (TFE) (Yamamoto et al. 2004b; Platt and Radford 2009). The formation of amyloid fibrils is highly dependent on pH, since no aggregation takes place above pH 5 in the absence of chemical agents, while long and straight fibrils spontaneously growth after incubation at acidic pH, with an optimum at pH 2.5 (McParland et al. 2000). At neutral pH, aggregation has been obtained by the addition of preformed fibrils acting as seeds (Chiti et al. 2001a). Further addition of moderate concentrations of TFE stimulates β2m fibrillation at neutral pH (Yamamoto et al. 2004a). TFE, among several protein denaturants and organic solvents examined, was the most effective in accelerating amyloid fibril extension at a neutral pH, and its effect is maximized at a concentration of 20% (v/v) (Yamamoto et al. 2005). According to Transmission Electron Microscopy results, the fibrils obtained at neutral pH and in presence of TFE are indistinguishable from those produced at acidic pH (Rennella et al. 2010).

Mutagenesis has shed some light on the structural determinants leading native β2m to associate into soluble oligomers and, eventually, into insoluble fibrils. Numerous β2m mutants have been prepared and their biochemical and structural features analyzed in relation to their amyloidogenic properties. Mutations of Pro32 showed that the slow phase observed in β2m folding mainly reflects the residue's *trans/cis* isomerisation (Kameda et al. 2005; Sakata et al. 2008). Interestingly, also a lag phase in amyloid fibril formation

disappears in the P32A mutant, suggesting that the *cis/trans* isomerisation of Pro32 can also limit the kinetics of amyloid aggregation (Eakin et al. 2006). β2m Cu^{2+}-dependent fibrillogenesis has been observed and related to His31. This residue chelates the copper ion, which promotes Pro32 isomerization (Eakin et al. 2002; Rosano et al. 2004) and triggers protein dimerization (Mendoza et al. 2010).

A trimmed β2m variant, Δ6-β2m devoid of the first six N-terminal residues, is found as a minor component (about 20%) of β2m amyloid plaques *in vivo*. Such a truncated form has a higher tendency to form amyloid aggregates, being not completely folded under physiological conditions (Esposito et al. 2000).
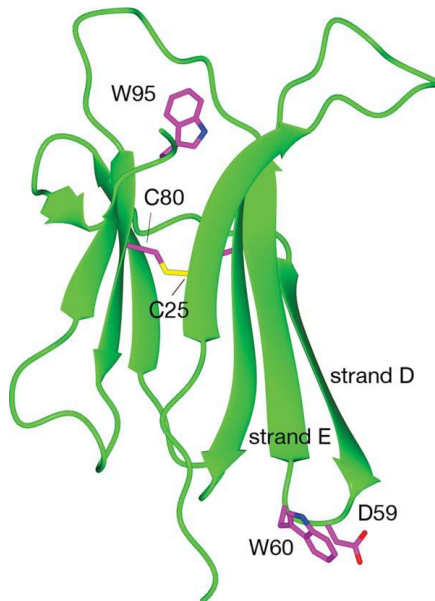

Figure 4.1 Cartoon representation of β2m tertiary structure. The residues mutated in this study (Trp60 and Asp59), the residues forming the disulphide bridge (Cys80 and Cys25), and the residue mainly responsible for intrinsic protein fluorescence (Trp95) are shown in sticks.

Recently, the β2m P5G mutant has been shown to behave similarly to Δ6-β2m (Eichner and Radford 2009). Furthermore, the crystal structure of the hexameric H13F mutant has shed light on the intermolecular interactions supporting β2m oligomeric state. In particular, residues Phe56 and Trp60 have been shown to mediate hydrophobic interactions at the subunit interfaces (Calabrese et al. 2008).

Ashcroft and coworkers (Borysik et al. 2004) pioneered the use of mass spectrometry methods for biochemical and biophysical characterization of β2m. In particular, analysis of the charge-state distributions (CSDs) obtained by electrospray ionization allowed identification of distinct conformers, which were further resolved and characterized by ion-mobility experiments (Smith et al. 2009). These studies depicted the distribution and the evolution of the molecular species as a function of time and pH, in w.t. β2m and two mutants (V9A, F30A) (Borysik et al. 2004; Smith et al. 2007). Protein oligomerization has also been investigated by these methods, showing distinct stoichiometry patterns under conditions that lead to different fibrillation kinetics and different fibril morphologies (Smith et al. 2006a).

Recently, the surface loop linking the D and E strands, and in particular the residue Trp60, have gathered increasing attention. Trp60 is a surface residue that is buried in the MHC-I complex, but becomes totally exposed to solvent in monomeric β2m. Its role in intermolecular interactions likely justifies its strong evolutionary conservation through vertebrates. The β2m W60G mutant shows marked differences from the wild type protein (Esposito et al. 2008), showing marked differences from the wild type protein. This mutant is more stable to chemical unfolding, does not aggregate into amyloid fibrils under mild conditions, has a more regular DE-loop conformation and shows a reduced tendency to form oligomers (Esposito et al. 2008). It has been suggested that such increased

stability might be the result of a more acceptable backbone conformation within the DE-loop (Esposito et al. 2008). Such observations prompted the design of other DE-loop mutants. The D59P mutant confirmed that mutations affecting the backbone conformational strain of the DE-loop may have long-range effects. Accordingly, the D59P mutation increases the loop rigidity, resulting in lower stability and higher tendency to aggregation (Ricagno et al. 2008). The W60V mutant showed the same stability of w.t. β2m to chemical denaturants, but reduced amyloidogenic propensity under mild conditions (Ricagno et al. 2009).

Altogether, the available evidence supports the hypothesis that β2m fibrillation is promoted by the accumulation of partially folded intermediates (Radford et al. 2005). NMR experiments have described residual structural features of the protein at acidic pH (Platt et al. 2005), or at neutral pH and in the presence of TFE (Rennella et al. 2010). However, the structural features of these intermediates are not known in detail.

In this work, wild type β2m conformational properties under two fibrillation conditions – pH 2.5 or pH 7.4 in the presence of 20% TFE – are compared by complementary biophysical techniques in order to obtain a deeper description of the β2m intermediates that characterize these two conditions. In addition, the effect of several mutations on β2m are investigated in terms of the stability and the oligomerization properties at neutral pH, as well as the structural features of the partially folded intermediate at pH 2.5.

Human ataxin-3

A group of nine neurodegenerative disorders, like Huntington or Kennedy diseases, is characterized by the presence of consecutive glutamines in the sequences of the proteins involved (Zoghbi and Orr

2000). The disease appears when the tract of poly-glutamines (poly-Q) exceeds a certain threshold, and its severity increases with the poly-Q length (Gusella and MacDonald 2000). One of this pathology, the spinocerebellar ataxia type-3 (SCA3) – also known as Machado-Joseph disease – is caused by the amyloid aggregation of the protein ataxin-3 (AT3).

Ataxin-3 is a 41kDa protein (Kawaguchi et al. 1994). Its structure consists in a globular, prevalently α-helical domain in the N-terminal region – called Josephin – that involves the first 182 amino acids. The Josephin domain is followed by a flexible and disordered region (residues 183-291) that contains several ubiquitin interacting motifs (UIMs). Finally, a series of 12-40 consecutive glutamines (the poly-Q tract) follows the unstructured region in the non-pathogenic form of the protein. The expansion of this glutamines to 55-84 comports the appearance of the disease (Masino et al. 2003). Ataxin-3 is emerging as a multifunctional protein (Mazzucchelli et al. 2009): it has a role as a transcriptional repressor (Li et al. 2002), it regulates aggresome-mediated protein degradation (Bilen and Pittman 2005) and it is involved in the ubiquitin (Ub) proteasome pathway as a ubiquitin-specific cysteine protease, binding the polyubiquitylated proteins via its UIMs (Nicastro et al. 2005). The latter function is the most characterized from a molecular point of view: UIMs were identified inside the flexible region of the protein, and the enzymatic site was localized in the Josephin domain, together with other two distinct ubiquitin-binding sites (Nicastro et al. 2009). The catalytic activity of the Josephin justifies its highly conserved sequence across a large number of living organisms, differently from the C-terminal of the protein (with the exception of the UIMs). Due to its importance, the structure and the stability of the Josephin has been extensively studied by different techniques. The results highlight a domain (24 kDa) composed by the first 182 amino acids, globular and monomeric, prevalently α-helical and fairly resistant to thermal

and chemical denaturation (Chow et al. 2004b; Masino et al. 2004; Nicastro et al. 2006) (Figure 4.2). Despite the functional and structural characterization of the N-terminal part of the protein, the role of the disordered region and that of the poly-Q tract remains unknown. In other poly-Q proteins this region was suggested to work as polar zippers, joining specific transcription factors bound to separate DNA segments (Perutz et al. 1994). In AT3, however, the activity of poly-Q remains obscure, and it is also possible that this tract has no specific function. In the murine variant, the not-expanded poly-Q tail is only six-amino acids long (Gusella and MacDonald 2000).

The aggregation mechanism of AT3 – and of the other proteins belonging to the poly-Q family – is extremely complex and object of controversial opinions. A commonly accepted point is the primary role played by the poly-Q tract. Indeed, proteins involved in poly-Q diseases share no sequence homology, and present the glutamine tract in different regions of the chain (Ross et al. 2003). The expansion of the poly-Q tract accelerates nucleation and elongation phases during aggregation, and confers to mature fibrils an increased stability that make them resistant to sodium dodecyl sulphate (Ellisdon et al. 2006). The peptide QBP1, selectively binding a sequence of consecutive glutamines with a pathogenic length, inhibits aggregation both *in vitro* and *in vivo* (Nagai et al. 2000). The insertion of poly-Q tracts of different lengths into a folded and stable protein, myoglobin, induced, for the longest variants, the formation of fibrils with small changes in the protein three-dimensional structure (Tanaka et al. 2001). Synthetic poly-Q peptides have been shown to assemble, directly from the monomeric state (Chen et al. 2002), into fibrils composed of β-sheets, stabilized by hydrogen bonds between the glutamine side chains (Perutz et al. 1994).

Even if the poly-Q tract represents the main modulator of the aggregation process, the precise mechanism adopted is not completely understood, and different models have been proposed.

Breuer and coworkers suggested that a proteolytic cleavage of the protein around the residue 250 is necessary to start the aggregation of AT3 (Haacke et al. 2006). Since fragments containing the poly-Q are the predominant species detected in neuronal inclusions in human brain, a "toxic fragment hypothesis" was proposed, whereby C-terminal fragments containing the extended poly-Q tract undergo a conformational transition and form the nucleus of the fibril. These fragments can also induce structural modification – particularly concerning the poly-Q tract – in the full-length protein, driving its sequestration into the nucleus through the poly-Q.
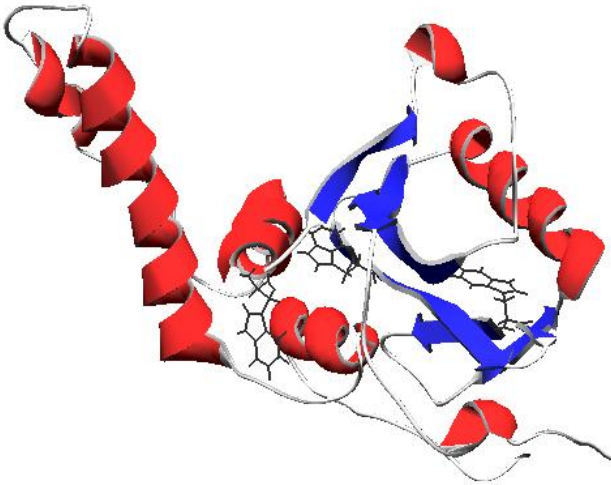

Figure 4.2 Cartoon representation of human ataxin-3 Josephin domain (PDB ID code 1YZB). The α-helical regions are colored in red, the β-strands in blue. The residues mainly responsible of intrinsic protein fluorescence (Trp87, Trp120 and Trp130) are shown in sticks.

A result supporting this hypothesis is given by the fact that the inhibition of calcium-dependent calpain, the putative protease responsible of the cleavage, also suppresses protein ability to aggregate (Haacke et al. 2007).

In a completely different scenario, the triggering event of fibril formation is the destabilization of protein structure by the expansion of the poly-Q. This hypothesis is closer to the general "misfolding mechanism", with accumulation of a conformational intermediate that is prone to aggregate (Chow et al. 2004c). Alternatively, the process has been suggested to follow a "domain swapping" mechanism, where the interaction of two structural elements is weakened in favor of an equivalent inter-molecular interaction, leading to oligomeric assemblies (Liu et al. 2001). In AT3 the poly-Q acts as a linker of the domains, and its expansion promotes both the swapping – as a result of conformational destabilization – and the formation of glutamine polar zippers. The β-sheets composing these zippers should orient in a parallel fashion, in such a way that the strands are stabilized by a continuous network of side-chain–side-chain hydrogen bonds (Bevivino and Loll 2001).

The C-terminal part of the protein has been shown not to affect the structure of the Josephin domain, based on NMR comparison of the full-length protein and the Josephin domain alone (Masino et al. 2004). However, apparently contrasting evidence has been reported on the effect of the poly-Q tail on conformational stability of the Josephine domain. No difference has been detected in the acid-induced unfolding of three ataxin-3 variants containing 15, 28 and 55 glutamines (Chow et al. 2004a). Furthermore, a similar behavior of the isolated Josephine domain and the full-length protein has been observed by CD and fluorescence during thermal denaturation and chemical denaturation by guanidine hydrochloride (Masino et al. 2004). On the other hand, a slight but significant destabilizing effect of the C-terminus is observed by fluorescence during chemical

denaturation by urea (Chow et al. 2004b). So, no conclusive evidence supports the hypothesis of a destabilizing effect of the C-terminal protein moiety on the folded N-terminal domain. Furthermore, no experiments have been reported dissecting the role, within the C-terminal region, of the disordered tract and the poly-Q tail.

Based on ThT fluorescence measurements on AT3 variants with differently expanded poly-Q tracts, a two-stage mechanism for AT3 fibrillation has been proposed. The first step leads to formation of SDS-soluble aggregates. This step is not dependent on poly-Q expansion, although expansion of the poly-Q above the threshold increases nucleation and elongation rates. The second step leads to formation of SDS-insoluble fibrils and depends on poly-Q expansion (Ellisdon et al. 2006; Ellisdon et al. 2007). In the "induced misfit" model, the interaction with a putative partner under physiological conditions prevents the Josephin domain to switch to the aggregation-prone state. In this view, expansion of the poly-Q could hinder the interaction with the partner, favoring in this way the beginning of fibril assembly (Masino et al. 2004).

The results described above underscore the importance of other regions, distinct from the poly-Q, for the fibrillation process (Saunders and Bottomley 2009). Indeed, different fragments lacking the glutamines are able to form aggregates. Both isolated the Josephin domain and the full-length protein with poly-Q replaced by the tripeptide QHQ have been shown to possess some fibrillation propensity (Masino et al. 2004; Ellisdon et al. 2007), but a direct comparison of the influence on aggregation of the components distinct from the poly-Q has not been performed yet. In particular, it is well known that disordered regions are particularly important in the fibrillation of many proteins (Baxa et al. 2006) and that they can affect the overall stability of a multi-domain protein (Saunders and Bottomley 2009). Thus, the disordered tract in AT3 (residues 183-

291) could be a critical region for the conformational and amyloidogenic properties of the protein.

In this work, two different truncated forms of AT3, the Josephin domain alone (1-182) and the Josephin domain followed by the unstructured region preceding the poly-Q tract (1-291), are compared in their aggregation propensity, in the conformational states populated under different conditions, and in the properties of assembling into oligomers. The aim of the measurements is to shed light on the role of the disordered region preceding the poly-Q tail in the conformational properties and the fibrillation mechanism of the protein.

Bovine β-lactoglobulin

Bovine β-lactoglobulin (BLG) is an 18.3 kDa protein belonging to the lipocalin family (Kontopidis et al. 2004). It is the major whey protein of cow's milk and it is usually found in two prevalent variants, A and B, both with two disulfide bonds (Cys66–Cys160 and Cys106–Cys119) and a free thiol (Cys121), but differing in two of the 162 amino acids (Asp-64-Gly and Val-118-Ala, respectively). BLG structure is characterized by a long α-helix and a three-turn helix at the C-terminus, and by a β-barrel composed of eight anti-parallel β-strands (A-H) that creates a hydrophobic pocket – the calyx – whose entrance is covered by the flexible E-F loop (Qin et al. 1998). The calyx accessibility is controlled by the so-called Tanford transition, a structural change of the protein that determines the open or closed protein conformation. This transition is regulated by pH, via protonation of a glutamate (E89) with an unusual pKa of 7.3 (Eberini et al. 2004). When E89 is in its protonated state, it becomes buried, causing E-F loop closure. Deprotonation, instead, causes its opening and exposure of the charged side chain to the solvent.

Despite the fact that the physiological role of BLG is not certainly established, the protein is thought to be involved in the transport of fatty acids bound inside the calyx. Indeed, the protein tightly binds small hydrophobic compounds like palmitic acid and retinol inside the β-barrel (Wu et al. 1999; Martins et al. 2008).

At neutral pH, BLG is prevalently dimeric. The oligomeric state is favored by several hydrogen bonds distributed across the A-B loops and a small intermolecular β-sheet formed by β-strands I (Brownlow et al. 1997) (Figure 4.3). However, the dimer interface area is small – ~ 1000 Å^2 – and complex formation is energetically dependent on the sequestering of water molecules between the monomers (Bello et al. 2008). Consequently, the interaction between the subunits is weak – the association constant is only 5×10^4 M^{-1} at pH 6.5 (Sakai et al. 2000) – and the dimer disassemblies easily in several circumstances, like acidic and alkaline pH, high temperature, low protein concentration and in the presence of organic solvents (Invernizzi et al. 2006), as well as a consequence of particular point mutations (Sakurai and Goto 2002). On the other hand, the presence of salt at acidic pH shifts the equilibrium towards the dimer, because the electrostatic repulsion between the monomers is weakened (Sakurai et al. 2001). Despite the frail quaternary structure, BLG monomeric fold is stable over a wide range of temperature and pH values (Invernizzi et al. 2006), suggesting that dimer separation is not caused by a destabilization of protein tertiary structure.

Even if BLG is not known to be involved in any amyloidogenic disease, a series of studies *in vitro* demonstrated the ability of the protein to assembly into amyloid fibrils under several experimental conditions, confirming the hypothesis that amyloid aggregation is probably an intrinsic feature of the polypeptide chain. A widely employed method to obtain BLG fibrils is by incubation at high temperatures, usually in the range 60-90°C. When the experiments are performed at pH values away from the isoelectric point of the

protein (pI ~ 5.1 for BLG), elongated amyloid aggregates form. On the other hand, spherical particles of micrometrical dimensions are produced incubating the sample near the isoelectric point (Bromley et al. 2005). In the latter case, however, the building blocks of the particles consist as short fibrillar components packed into the spherical structure (Krebs et al. 2009). Regardless of the pH, the aggregation kinetics are accelerated by increasing temperature (Domike and Donald 2007), ionic strength or protein concentration (Schokker et al. 2000). The early stages of the assembly have been monitored by several techniques, leading to the characterization of the conformational and oligomeric intermediates involved in the process.
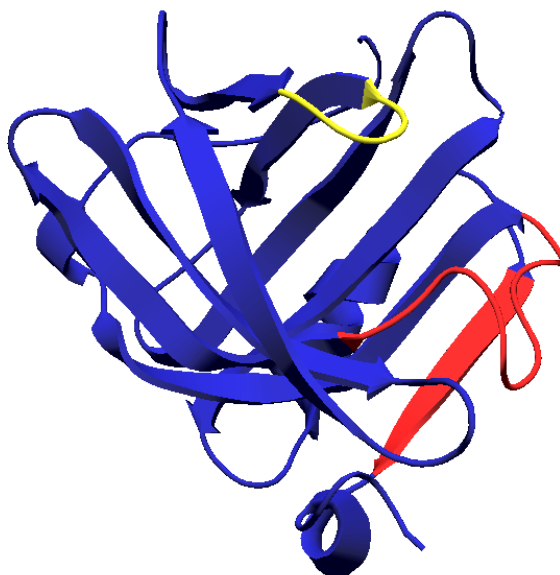
Figure 4.3 Cartoon representation of monomeric bovine BLG (PDB ID code 3BLG). The E-F loop responsible of the Tanford transition is shown in yellow. The A-B loop and the β-strand I involved in the dimer interface are shown in red.

The "activated" species has been identified in a partially unfolded monomer (Schokker et al. 1999) that shows the typical features of a molten globule (Carrotta et al. 2001). During the aggregation, non-native, covalently linked dimers and trimers were detected (Schokker et al. 1999; Carrotta et al. 2001).

BLG transition to the fibrillar form has been also induced under high hydrostatic pressure (Zasypkin et al. 1996; Dumay et al. 1998) or in the presence of chemical denaturants (Katsuta et al. 1997; Renard et al. 1999). In particular, samples incubated at 37°C in 5M urea – a concentration close to the midpoint of BLG monomer unfolding transition – have been shown to assemble into amyloid aggregates by the classical nucleation-elongation mechanism influenced by seeding (Hamada and Dobson 2002). The lag phase of the process has been studied in detail, leading to the identification of three subsequent steps: the protein is initially destabilized with a partial unfolding, then oligomers assemble and finally they convert into a more stable nucleus with the complete disruption of protein fold (Giurleo et al. 2008). In order to obtain a more localized information on the regions involved in the aggregation mechanism, fibrils formed by peptide corresponding to different BLG β-strands with an intrinsic propensity to aggregate have been tested in their capability to promote the aggregation of the whole protein (Hamada et al. 2009). Such a seeding effect has obtained only for the peptide corresponding to the βA strand, the only accessible strand in the intermediate conformation assumed by the protein under the employed experimental conditions. Indeed, the other tested regions– strands βG, βH and βI – are protected by the protein structure, especially thanks to disulphide bonds. Therefore, βA was suggested to have a key role in the nucleation reaction (Hamada et al. 2009).

BLG aggregation has also been induced by TFE addition. The first evidence that the incubation of BLG in TFE permits the assembly of the protein into β-sheet aggregates was obtained by Fourier-

transform infrared spectroscopy (Dong et al. 1998). The fibrillar nature of these aggregates was confirmed later by Atomic Force Microscopy, highlighting different morphologies depending on pH (Gosal et al. 2002). However, conformational and oligomeric intermediates have not been characterized under these circumstances. Another crucial element for BLG aggregation is represented by the free thiol in position 121, even if its role in the assembly process is controversial. The incomplete refolding *in vitro* of denatured BLG has been attributed to this free thiol group (Burova et al. 1998). The implication of Cys121 as an intermolecular linker for dimer formation has been attested both in heat- (Schokker et al. 1999) and urea-induced (Giurleo et al. 2008) fibrillations. A thiol-disulfide exchange has been shown to take place during fibril formation in urea, and the absence of the free thiol group in iodoacetamide-treated BLG preventes aggregation under this condition. Nevertheless, after solubilization of urea-induced fibrils, no covalent dimers have been detected by mass spectrometry (Hamada and Dobson 2002). Furthermore, the C121S mutation has been shown to promote protein aggregation during heterologous expression in *E. coli* (Invernizzi et al. 2008). On the other hand, intermolecular disulfide bonds have been shown not to be required for the formation of BLG amyloid fibrils upon prolonged treatment with potassium thiocyanate (Rasmussen et al. 2007). Thus, there is no clear evidence on the effective role of Cys121 in the BLG fibrillation process.

It is noteworthy how millions of years of evolution has worked to prevent aggregation for a protein with intrinsic amyloidogenic features. In the early stages of the folding process, two regions with an intrinsic propensity to aggregate, βG and βH, are rapidly driven to a native-like structure through the Cys106–Cys119 bond (Kuwata et al. 2001), and a non-native α-helix in the region corresponding to the native βA strand is formed, probably to avoid the exposition of a dangerous amyloidogenic tract (Chikenji and Kikuchi 2000). The

final structure of BLG reduces the intrinsic ability of the different regions to assembly, and the only dangerous section left, the strand βI, is protected by dimer formation (Hamada et al. 2009). Finally, Cys121, whose presence is justified by a possible role in the folding pathway (Invernizzi et al. 2008), is subsequently buried in the protein core and protected to dangerous expositions by the short C-terminal helix. In spite of this remarkable protection mechanism, a deflection from physiological conditions can induce BLG to undergo conformational changes that expose the protein to the aggregation risk.

In this work, the aggregation properties of BLG in the presence of TFE have been studied. In particular, the effect of TFE on the distinct structural levels of the protein have been analyzed by different techniques, and covalently bound intermediates of the fibrillation process have been identified. The role of Cys121 in the aggregation mechanism is discussed. Studies on BLG Tanford transition and native dimerization have also been performed. Even if these analyses are not directly connected to fibrillation, they permit to optimize analysis of this protein by nano-ESI-MS. The results obtained in this part of the work are exposed in Appendix A.

RESULTS AND DISCUSSION

5. HUMAN β2-MICROGLOBULIN

A probable cause of the different aggregation kinetics and morphologies observed for β2m under several different conditions is the accumulation of distinct intermediate states. Even if the characterization of these species is difficult, the employment of complementary biophysical techniques can shed some light on the intermediates that are formed during the process. The first part of this chapter focuses this topic. The second part focuses on the DE-loop, which has been highlighted as a crucial region for protein aggregation. The effect of several mutations in this loop is analyzed in terms of native oligomerization and conformational stability.

Protein conformational intermediates

Two fibrillation protocols *in vitro* are generally employed to induce aggregation in β2m: incubation at pH 2.5 or incubation at pH 7.4 in presence of 20% TFE. One goal of this work is the structural analysis of the protein under these two conditions and the characterization of the intermediate states involved in the fibrillation mechanism. The nano-ESI-MS spectra in Figure 5.1 show the effect of TFE and pH on β2m CSD. The raw spectra were deconvoluted by Gaussian fitting, in order to identify and quantify the individual components. The results are reported in the inset of each panel. Under non-denaturing conditions (pH 7.4, 0% TFE), the spectrum presents only a few peaks centred around the 7+ ion (Figure 5.1A). Narrow distributions corresponding to relatively low charge states are typical of folded, globular proteins. This result is consistent with previous nano-ESI-MS analyses under non-denaturing conditions (Smith et al.

2007; Santambrogio et al. 2010) and is compatible with β2m being folded in its native, immunoglobulin-like domain. A single species is detected by the fitting algorithm under these conditions. The addition of moderate amounts of TFE (20% final concentration) induces the appearance of a new CSD at lower mass-to-charge ratios, with the main charge state 9+ (Figure 5.1B). The deconvolution algorithm identifies this peak envelope as a new species corresponding to ~13% of the population. The slightly increased average charge state of this component suggests that it represents a partially folded form of the protein. This component progressively accumulates at increasing temperatures of the instrument interface (Figure 5.1C) or at increasing TFE concentrations (Figure 5.1D,E), becoming the prominent state (~52%) at 60% TFE (Figure 5.1E).
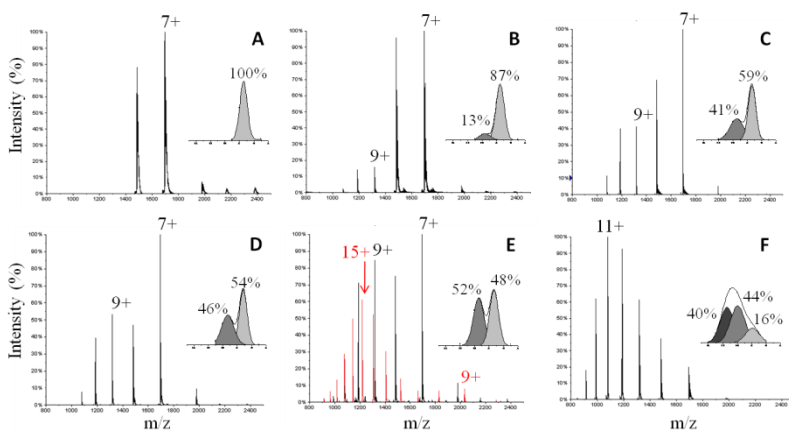
Figure 5.1 Nano-ESI-MS of 5 µM β2m (black) and BLG (red) in 10 mM ammonium acetate at variable pH, IHT and TFE concentrations. A) pH 7.4, IHT off; B) pH 7.4, 20% TFE, IHT off; C) pH 7.4, 20% TFE, IHT 150°C; D) pH 7.4, 40% TFE, IHT off; E) pH 7.4, 60% TFE, IHT off; F) pH 2.5, IHT off. Peak envelopes are labeled by the main (most intense) charge state. Each frame includes the Gaussian fitting results, where each component is labeled by its relative amount.

A residual ~48% of the protein in a native-like state at 60% TFE might seem surprising. In order to rule out a bias in the detection system, another protein of similar size, BLG, was used as an internal reference. As shown in Figure 5.1E, the CSD of BLG, from an equimolar mixture with β2m in 60% TFE, is typical of an almost completely denatured protein. This result indicates that the observed effects are quite protein-specific and that there is no systematic enhancement of low charge states under the employed experimental conditions. The presence of similar amounts of two distinct conformers of β2m with different hydrodynamic radius has been confirmed by gel filtration in 60% TFE (data not shown).

Two conformational intermediates of β2m in 20% TFE have been identified by Bellotti and coworkers (Rennella et al. 2010): a native-like intermediate (I) and a denatured form (D). The letter has been recognized as the amyloidogenic intermediate. The 9+ form resolved here can be identified as D, since it represents the maximally unfolded species induced by TFE. Higher TFE concentrations, indeed, lead to higher accumulation of the 9+ form but no further shift in the CSD. In agreement with Bellotti and coworkers, this species accumulates further also by increasing the temperature of the sample in 20% TFE. We conclude that the I intermediate is not resolved in the here reported nano-ESI-MS spectra, having very similar overall compactness to that of the native protein.

The conformational properties of β2m appear to be quite different at pH 2.5 in the absence of TFE (Figure 5.1F). The CSD is further shifted to lower *m/z* values, compared to pH 7.4, 20% TFE, with the most intense peak corresponding to the 11+ ion. Similar CSDs have been reported for β2m at pH 2.5 (Borysik et al. 2004). Three components have been identified by ion mobility measurements under these conditions: native-like, denatured and partially folded. Therefore, the broad CSD of β2m at pH 2.5 has been deconvoluted by Gaussian fitting using three components. The results are in good

agreement with those from ion mobility data, with a minor fraction of native-like compactness (~16.6%) centred on the 8+ ion, an intermediate (~43.6%) centred on the 10+ ion, and a denatured form (~39.8%), centred on the 11+/12+ ion (Smith et al. 2009). The latter has been previously recognized as the most "left-shifted" form of β2m under non-reducing conditions, since the addition of denaturing agents does not further shift the distribution (Santambrogio et al. 2010). Therefore, this conformer should represent the minimal β2m compactness, compatible with maintenance of the disulphide bridge (Santambrogio et al. 2010). An average charge state of about 10+, instead, has been previously recognized as characteristic of a partially folded form that accumulates in WT and mutant β2m at pH below 5 (Borysik et al. 2004). This species reveals intermediate compactness by ion mobility measurements (Smith et al. 2007). Residual structure of β2m at pH 2.5 has been documented also by NMR (Platt et al. 2005). Altogether, these results indicate that β2m under the two compared fibrillation conditions populates different molecular ensembles. However, each condition leads to formation of a metastable, partially folded form. These two intermediates are similar in terms of structural compactness.

The effect of TFE and pH was also analysed by intrinsic fluorescence (Figure 5.2A). Under non-denaturing conditions (pH 7.4, 0% TFE), β2m has a strong intrinsic fluorescence, which is almost completely due to the buried Trp95 (Esposito et al. 2000). Upon the addition of 20% TFE, the spectrum presents only minor changes, indicating that a large fraction of the molecular population has a still structured hydrophobic core that maintains Trp95 in a native-like environment. At higher TFE concentrations, the fluorescence emission peak progressively decreases and shifts towards higher wavelengths. At pH 2.5, the fluorescence emission spectrum is drastically different from that at 20% TFE (pH 7.4) reflecting more pronounced protein denaturation (Figure 5.2A, thick

line). Thus, fluorescence properties point to different average tertiary structure of β2m under the two different fibrillation conditions, in agreement with the different species distributions revealed by nano-ESI-MS. Far-UV CD was used to monitor protein secondary structure. Under non-denaturing conditions (pH 7.4, 0% TFE), the β2m spectrum has a broad minimum between 210 and 220 nm and a positive maximum at 200 nm (Figure 5.2B). These features are characteristic of proteins rich in β-strands. Minor spectral changes are observable in the presence of 20% TFE (Figure 5.2B). The isodichroic point around 200 nm is typical of a beta-to-alpha transition.

As the TFE concentration is raised from 20 to 40%, the spectrum acquires features of predominantly helical structures, with a minimum at 207 nm, a shoulder at 220 nm, and positive ellipticity values at 190 nm (Figure 5.2B). These results are consistent with the known stabilizing effect of TFE on protein secondary structures, especially α-helices (Shiraki et al. 1995) and with previously reported CD data (Yamamoto et al. 2004b; Rennella et al. 2010). The transitions monitored by CD and fluorescence are very similar, revealing cooperative unfolding with a midpoint between 20% and 30% TFE. The spectrum recorded at 20% TFE and 50 °C is shown as a dotted line in Figure 5.2B. At pH 2.5 in the absence of TFE (Figure 5.2C), the spectrum reveals significant accumulation of random coil relative to ordered secondary structure, with a shift of the minimum to 205 nm and a strong decrease in the signal at 190 nm. Therefore, independent biophysical probes consistently show that the conditions compared here, pH 2.5 and pH 7.4, 20% TFE , induce drastically different conformational states of β2m in solution both in terms of secondary and tertiary structure. All of the here reported CD measurements have also been performed in 10 mM ammonium acetate, in order to guarantee identical solvent conditions as in the

nano-ESI-MS experiments. Very similar results were obtained compared to phosphate buffer (data not shown).

It would be desirable to directly compare the two partially folded forms induced by the conditions considered here. Difficulties arise by the fact that they are represented only by a minor fraction of the molecular population. However, the average information obtained by CD measurements can be deconvoluted on the basis of the species distributions derived by nano-ESI-MS, in order to obtain the calculated spectrum of the intermediate.
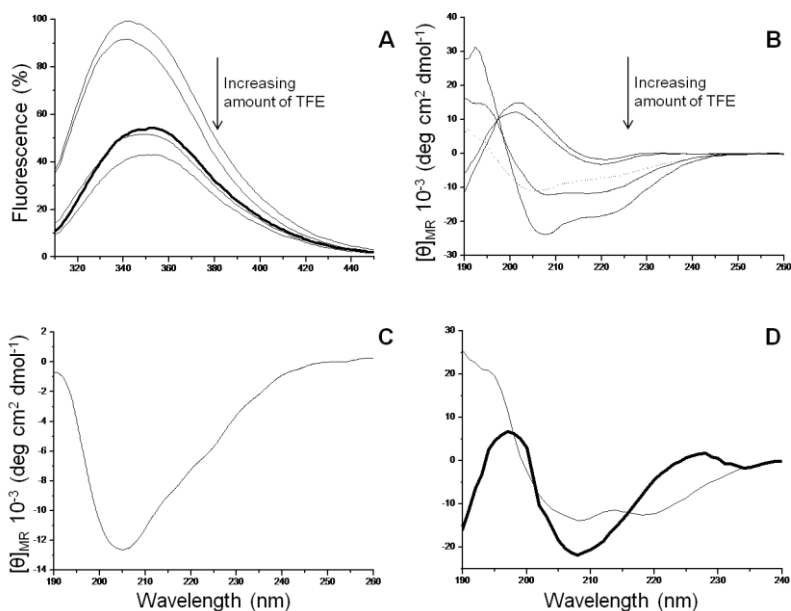
Figure 5.2 Spectroscopic measurements on 37 μM β2m in 10 mM ammonium acetate (fluorescence) and 14 μM β2m in 10 mM phosphate buffer (CD). A) Tryptophan fluorescence at pH 7.4, 25 °C, in the presence of 0%, 20%, 30% and 40% TFE (thin lines), or pH 2.5, 25 °C (bold line). B) Far-UV CD at pH 7.4, 25 °C, in the presence of 0%, 20%, 30% and 40% TFE (thin lines), or 20% TFE, 55 °C (dotted line). C) Far-UV CD at pH 2.5, 25 °C. D) Calculated Far-UV CD spectra for the intermediate in 20% TFE (thin line) and that at pH 2.5 (bold line).

Although protein signal yield in nano-ESI-MS tend to increase for proteins in the denatured state, the area obtained by Gaussian fitting of the CSDs can be taken as an approximate estimate of the relative amounts of distinct conformers, particularly if not fully unfolded, as in this case. It should also be noted that the relative amount of the TFE-induced intermediate measured by ESI-MS (~13%) is not overestimated relative to the 20% inferred by CD titrations. Therefore, a deconvolution procedure has been implemented as described in Appendix B. In the case of 20% TFE (pH 7.4), it was assumed that the compact conformation detected by nano-ESI-MS (Figure 5.1B, main charge state 7+, relative amount 87%) represents the native state, whose CD spectrum is known. The intermediate (Figure 5.1B, main charge state 9+, relative amount 13%) is the unknown species. Therefore, the CD signal of the sample at 20% TFE at each wavelength can be seen as a linear combination of the known value of the native protein and the unknown value of the intermediate, weighted by the relative abundance obtained by Gaussian fitting of the nano-ESI-MS data. This procedure yields, by difference, the spectrum of the unknown intermediate. Figure 5.2D (thin line) illustrates the results of this procedure performed on the basis of the spectra at pH 7.4, in the presence of 0% or 20% TFE. The resulting difference spectrum is indicative of a predominantly helical structure, with a double negative minimum at ~207 and ~220 nm, and positive signal at 190 nm. For the sample at pH 2.5, the coexistence of three distinct species must be taken into account, where the native and the denatured forms should be considered as the known species, in order to calculate the spectrum of the intermediate. To this aim, we simplistically describe the denatured form by the reference spectrum of the random coil. This assumption will lead to a minor overestimation of the beta-structure content of the intermediate, since the denatured form is not 100% random coil

(Platt et al. 2005). The spectrum of the intermediate calculated on this basis is quite different from that of the intermediate at pH 7.5, 20% TFE, revealing a lower contribution of alpha *versus* beta secondary structure, as indicated by the single negative band around 210 nm and the strongly decreased signals at 190 and 220 nm. These differences suggest that the fraction of ordered secondary structure in these intermediates is mainly beta at low pH and mainly helical at pH 7.4, 20% TFE.

In conclusion, the two partially folded forms identified in this work, either in the presence of TFE or at low pH, do not represent the same β2m conformer. Although they have similar overall compactness, their ordered secondary structure is quite different. The low-pH intermediate has a predominant beta component, while the TFE-induced intermediate has a more helical character. Although it is possible that the two intermediates have partially structured regions in common, the differences in protein conformation likely account for the different aggregation propensity of β2m under the two compared conditions.

Effect of mutations in the DE-loop

This part of the work deals with a more localized study of the role of the DE-loop, known as a critical region for β2m stability and aggregation. The effects of several point mutations in this region on protein stability and dynamics are analyzed in the attempt to relate these features to the effects of the mutations on β2m aggregation propensity. Protein thermal stability has been studied by near- and far-UV CD for w.t. β2m and the mutants W60G, W60V and D59P. Thermal stability increases remarkably in the W60G mutant, remains unaffected in W60V and decreases slightly in D59P (Santambrogio et al. 2010). Such overall effects on β2m properties may be related to

the fact that residues in the DE-loop (either Trp60 or Ser57) populate unfavorable regions of the Ramachandran plot in all the known three-dimensional structures of β2m, either monomeric or in complex with the MHC-I heavy chain (Ricagno et al. 2009). Thus, their mutation may release or accentuate the conformational strain of the protein backbone in this region of the structure. Glycine is known to be unique among amino acids in conferring backbone conformational freedom, thanks to the wider range of allowed φ and ψ angles. A closer-to-ideality Ramachandran plot for the W60G mutant, compared to the w.t. protein, is well reflected by the higher T_m (~7°C higher than w.t.) observed for the mutant, consistent with data by chemical denaturation (Esposito et al. 2008). On the other hand, a loss in backbone flexibility due to the D59P mutation may explain the aggregation propensity and the small drop in T_m observed for the D59P mutant (Ricagno et al. 2008). The similar thermal stability displayed by w.t. and W60V mutant is in agreement with similar conformational restraints exerted on the protein backbone by the Trp or Val side chain at this site (Ricagno et al. 2009).

The four β2m variants have been analyzed by means of nano-ESI-MS under non-denaturing conditions, to detect intermolecular interactions as a function of protein concentration. Soluble oligomers formed by the β2m variants have been analyzed under mild desolvation conditions, to favor detection of non-covalent complexes (Figure 5.3). The CSDs of the w.t. and mutant proteins are narrow, with main charge state 6+/7+, consistent with a compact protein conformation. The spectra of w.t. β2m reveal oligomer-specific peaks corresponding to dimers, trimers, and tetramers of a natively folded protein, which can be already detected starting at 15 μM protein concentration. Such aggregation pattern is here referred to as "native aggregation". The reported concentrations are well below the generally observed threshold (100 μM) for unspecific protein aggregation under electrospray conditions (Invernizzi and Grandori

2007). Furthermore, the concentration dependence of the spectra rules out that the observed oligomeric species result from dissociation of higher-order aggregates during the analysis. Therefore, these data can be interpreted in terms of protein oligomerization already present in the original solvated samples. As shown in Figure 5.3, all three β2m mutants display a reduced propensity to form oligomers. The W60G and W60V mutants are slightly more affected than the D59P mutant. Thus, the DE-loop, and the exposed side chain of Trp60 in particular, seem to mediate protein–protein interaction in β2m oligomerization under non-denaturing conditions. The effect of the D59P mutation can be interpreted according to previously published evidence on electrostatic interactions. A salt bridge between Asp59 and Lys19 has been observed at the subunit interface of hexameric β2m and has been implicated in dimer association (Srikanth et al. 2009; Mendoza et al. 2010). The effect of Cu^{2+} on protein oligomerization has also been interpreted by repositioning of Asp59 into a more favorable orientation for ion pairing with Lys19 (Srikanth et al. 2009; Mendoza et al. 2010). Our observation that the D59P mutant has a reduced propensity to native aggregation is consistent with a contribution of Asp59 in w.t. protein association even in the absence of Cu^{2+}.

Protein oligomerization is also affected by mutations at site 60. Both the W60G and W60V mutants display a distinctly lower tendency to form native oligomers compared to w.t. β2m. In the MHC-I complex, β2m Trp60 is hosted inside a hydrophobic pocket of the heavy chain, contributing favorably to complex stability. When β2m is in the monomeric, unbound state, Trp60 becomes completely exposed to the solvent, thus in an energetically unfavorable state for a bulky aromatic residue (Esposito et al. 2008). The exposed Trp60 side chain in isolated β2m may promote intermolecular association via aromatic interactions. The W60V mutant, the variant with the lowest tendency to native aggregation, cannot support aromatic interactions

at this site; it may also face scarcely favorable intermolecular packing posed by the Cβ-branched aliphatic side chain. Although the observed aggregation propensity likely relates to the intrinsic amyloidogenic properties of β2m, additional investigations will be needed to understand the role that such species may play in the fibrillation process.

The reduction kinetics by dithiothreitol (DTT) at pH 2.5 of each mutant compared to the w.t. protein has been analyzed by nano-ESI-MS. To ensure identical sample handling, each β2m mutant has been analyzed in a mixture with equimolar amounts of the w.t. protein, over a time course after DTT addition (Figure 5.4).
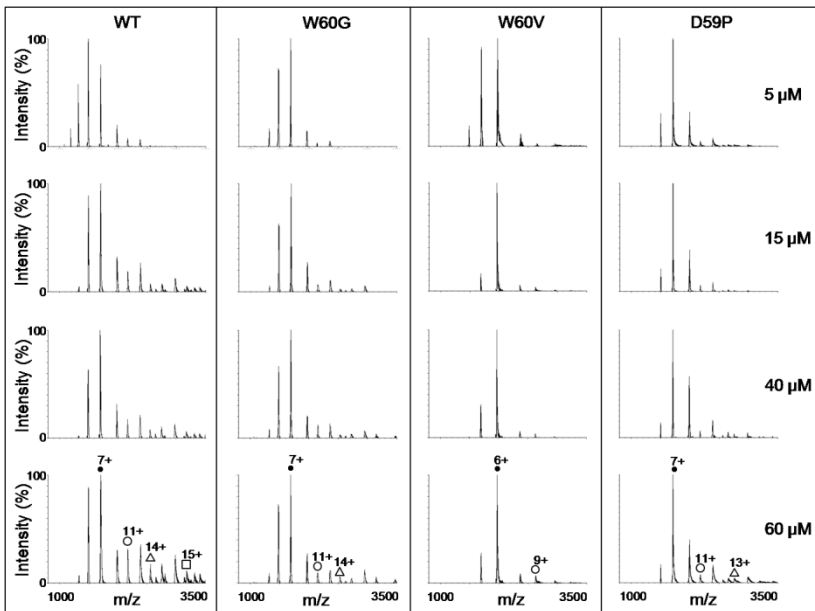
Figure 5.3 Protein oligomerization. Nano-ESI-MS of w.t., W60G, W60V and D59P β2m, at variable protein concentration (in milli-Q water, pH 6.5). The most intense peak of the monomer (●), dimer (○), trimer (△) and tetramer (□) is labeled by the corresponding symbol and by the charge state.

Peaks can be assigned unambiguously to each protein thanks to their different masses. In agreement with previous reports, reduction of the disulfide bridge in the folded w.t. protein (at pH 7.4) does not occur at an appreciable rate (data not shown), while the fibrillogenic intermediate that accumulates at pH 2.5 is prone to reduction (Hong et al. 2002). The spectra before the addition of DTT indicate that the CSDs of the mutants at pH 2.5 are superimposable to those of the w.t. protein, centered on the 11+/12+ ion. After the addition of 20 mM DTT to the acid-denatured protein at room temperature, the disulphide becomes slowly reduced, as indicated by the appearance of new peaks revealing a shift of 2 Da in protein masses. The reduced form also maps in a different region of the spectrum, with a CSD centered on the 15+ ion, indicating a more extended conformation than the oxidized, denatured protein. While the W60V and D59P mutants behave similarly to the w.t. (Figure 5.4B and 5.4C), W60G displays a significantly delayed reduction (Figure 5.4A). After 5 hours from DTT addition, the CSD of the w.t. protein is clearly bimodal, indicating accumulation of the reduced, fully unfolded protein (~35%). At the same time point, the W60G mutant is still almost completely oxidized. After 24 hours, only ~20% of the mutant protein is reduced. This result indicates that, in spite of the similar CSDs, there are fine differences in the structure and/or dynamics of the acid-induced intermediate of W60G that make its disulfide bridge less accessible to the solvent compared to the other variants. Thus, mutation of the Trp60 residue to Gly, not only affects stability and oligomerization in the native state, but also alters the structural properties of the amyloidogenic intermediate at low pH. The distinct behavior of W60G during reduction kinetics could stem from either a slightly more structured conformation, or reduced flexibility, or both. Thus, the long-range effects of the W60G mutation appear to affect not only the mutant native state, but also its partially unfolded intermediate. The possibility to describe fine

structural differences among partially folded forms under fibrillation conditions is highly promising for future studies on the distinct fibrillation properties of β2m variants.

The results reported herein on β2m and its DE-loop mutants stress the crucial role played by a protein loop, possibly by just one residue in this region, in determining protein stability, dynamics, and oligomerization. The amyloidogenic propensity of β2m mutants may therefore result from a subtle combination of several biophysical and molecular properties.
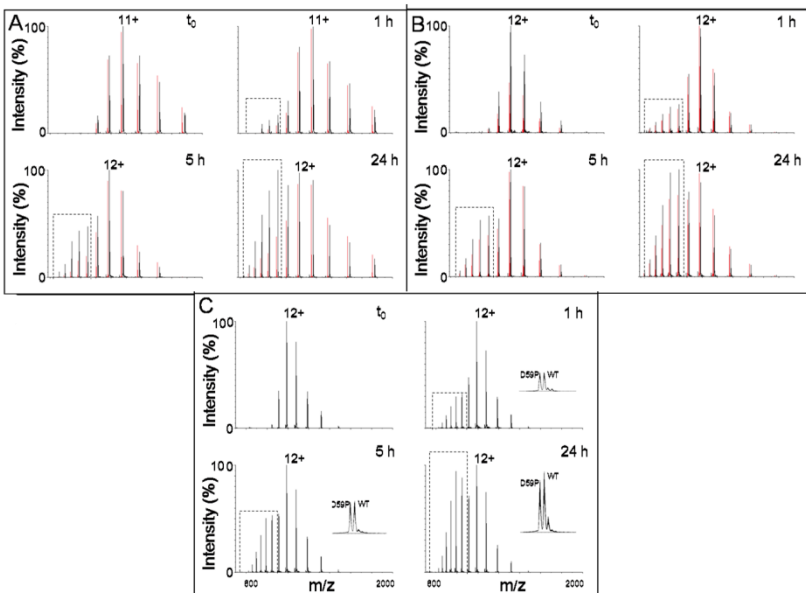
Figure 5.4 Reduction kinetics. Nano-ESI-MS of equimolar (2.5 µM) mixtures of w.t. (black) and mutant (red) proteins at variable time of incubation in 10 mM ammonium acetate pH 2.5, 20 mM DTT, at room temperature. (A) w.t. and W60G; (B) w.t. and W60V; (C) w.t. and D59P. Frames include peaks that reveal an average-mass shift of 2 Da. Peak envelopes are labeled by the main charge state. Due to the similar molecular weight of w.t. and D59P, the presence of the two components is shown in zoomed images of the 15+ peaks in the insets of panel C.

We will undertake further studies on the effects of these mutations on β2m fibrillation as a function of pH and TFE concentration, in the attempt to clarify which step(s) of amyloidogenesis are affected by the DE-loop mutations.

6. HUMAN ATAXIN-3

The goal of this part of the work is to investigate the specific role of the disordered region of AT3 (residues 183-291) in determining the conformational and aggregation properties of the protein. This region is immediately after the Josephin domain (1-182) and just before the poly-Q tract. To this purpose the protein fragments 1-182 (AT3/182) and 1-291 (AT3/291) have been produced and analyzed by nano-ESI-MS, CD, intrinsic and ThT fluorescence. Conformational features and stability to chemical denaturation of the two fragments are compared. Then, the amyloidogenic properties and the oligomerization propensities induced by heat are studied for the two species.

Conformational features of the fragments

Nano-ESI-MS is employed to investigate the conformational properties of the truncated forms AT3/182 and AT3/291, which are very soluble and readily detectable by this method. The comparison of the CSDs of the two truncated forms was performed on equimolar mixtures of the two fragments, in order to guarantee analysis under identical experimental conditions. As shown in Figure 6.1, the Josephin domain (AT3/182) under non-denaturing conditions is characterized by a narrow peak envelope at relatively high *m/z* values (main charge state 10+), as typically observed for folded proteins. This result is in agreement with NMR data indicating that the isolated Josephin domain has a well-defined globular structure (Masino et al. 2004; Nicastro et al. 2005). On the contrary, the longer fragment (AT3/291) displays a bimodal CSD with a more compact

form centered on the 13+ ion and a less compact form centered on the 26+ ion. This result indicates that, while AT3/182 exists as a stable, compact structure, AT3/291 interconverts between two distinct conformational states characterized by different compactness. The largely predominant component at 26+ in the AT3/291 spectrum most likely represents the protein containing the folded Josephin domain and the disordered segment 183-291. This interpretation is supported by the fact that such a form can be further unfolded to a fully denatured state with transitions almost identical to those of the isolated Josephin domain (see below).
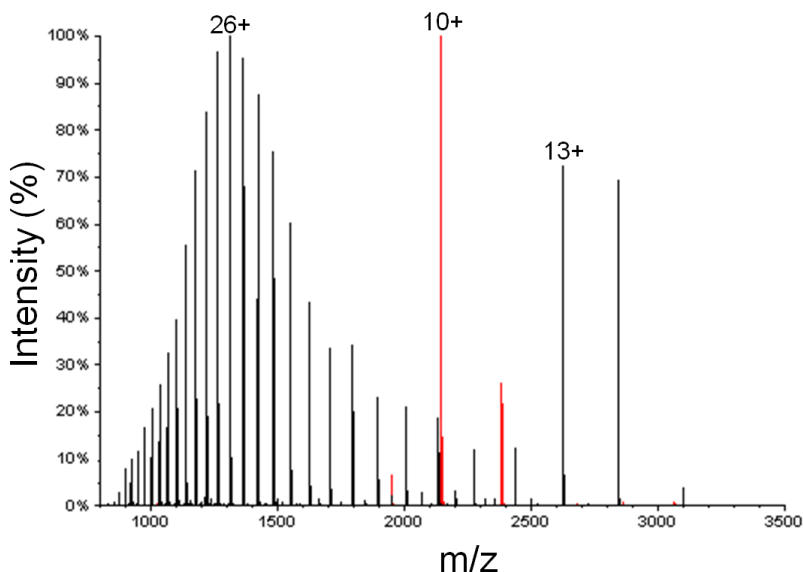
Figure 6.1 Nano-ESI-MS spectrum of a mixture of 10 µM AT3/182 (red) and 10 µM AT3/291 (black) in 5 mM ammonium acetate pH 6.5. Numeric labels indicate the main charge state of each peak envelope. The spectrum was recorded at room temperature.

Thus, the existence of a second component at lower charge states (13+) is unexpected. Such a form is also of surprising compactness for a chimera of an ordered domain and a disordered region of similar lengths. Indeed, its average charge state (12.6, calculated from the peaks of the 11+, 12+, 13+ and 14+ ions) approximates the expected limit value (Heck and Van Den Heuvel 2004) for folded, globular proteins of that size (14.4). This result indicates that the disordered segment can fold into a considerably compact, metastable conformation. Such a component represents approximately ~12% of the total signal intensity, which is a lower-bound estimate of the corresponding fraction in the molecular population, considering that unfolded protein ions may have higher ESI signal response than folded protein ions (Kuprowski and Konermann 2007). The AT3 C-terminal moiety (including the disordered region and the poly-Q tract) has been shown by far-UV CD to contain a considerable degree of helical secondary structure (Masino et al. 2004). However, no significant tertiary structure content could be detected by NMR (Masino et al. 2004). A compact state of AT3 C-terminal moiety has been hypothesized for the protein bound to intracellular partners (Masino et al. 2004). Here it is shown that the disordered region of the protein is capable to acquire transient tertiary structure, at least when linked to the Josephin domain. How the average secondary structure revealed by CD (Masino et al. 2004) is distributed in the two conformers detected in this work is an issue that cannot be addressed with the present data. However, the average secondary structure of the disordered tract (residues 183-291) was estimated by differential CD measurements. Figure 6.2 shows comparison of the far-UV CD spectra of AT3/182 and AT3/291, indicating, as expected, a higher relative amount of random-coil AT3/291. Nevertheless, the difference spectrum, which indicates the contribution of the 183-291 tract, corresponds to a disordered structure with some helical component, as indicated by the minima at

205 and 222 nm. Deconvolution results indicate a helical content around 17%. These results indicate that the 183-291 region is not completely disordered and has an intrinsic propensity to form helical structures. This propensity can be due to the α-helical conformation characteristic of the UIMs motifs present in this region (Sgourakis et al. 2010).

Chemical unfolding of the fragments has been investigated by near-UV CD and Trp fluorescence, as alternative methods to probe tertiary structure. Both these techniques sense the environment of Trp residues. Since the 183-291 segment contains no Trp, while three of them map into the Josephin domain, these methods provide specific information on the conformational state of the Josephin domain within either the AT3/182 or the AT3/291 protein context.
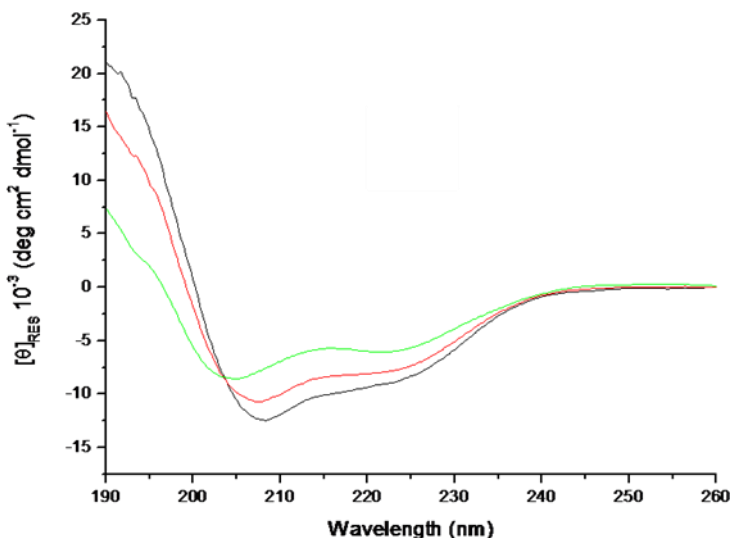
Figure 6.2 Measured far-UV CD spectra of AT3/182 (black line), AT3/291 (red line) and the calculated contribution of the 183-291 region (green line).

As can be seen in Figure 6.3, the unfolding transitions revealed by these techniques for the two compared protein fragments in the presence of increasing acetonitrile concentrations are very similar. These results indicate that the conformational stability of the Josephin domain is not significantly affected by the presence of the disordered region, in agreement with what previously reported based on far-UV CD (Chow et al. 2004b; Masino et al. 2004). Figure 6.3 also shows that the near-UV CD spectrum at the end of the transition is not flat but, rather, has lost the initial positive peak and acquired a new negative one.
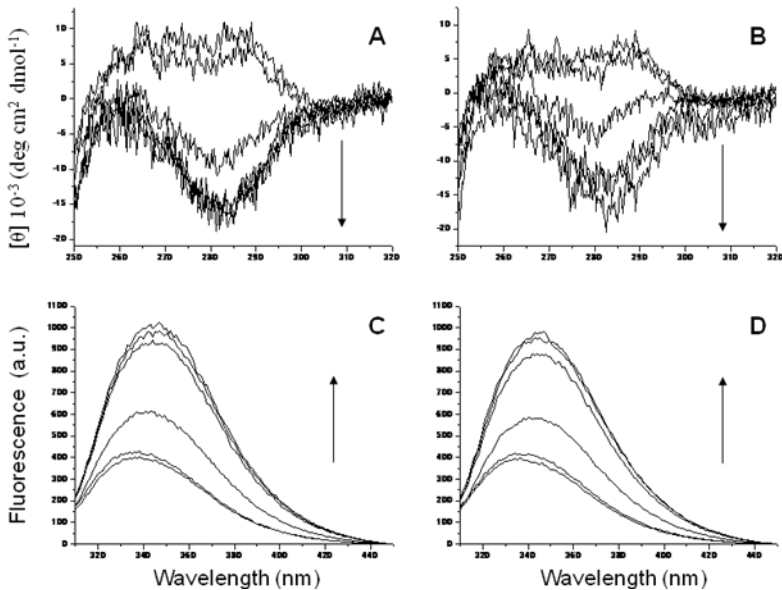
Figure 6.3 A, B) Circular dichroism in the near UV. C, D) Tryptophan fluorescence (excitation at 295 nm). Samples contain 20 μM AT3/182 (A, C) or AT3/291 (B, D) in 5 mM ammonium acetate, pH 6.5, and variable concentrations of acetonitrile (10%, 20%, 30%, 40% and 50%). The arrows indicate the direction of spectral changes at increasing concentrations of acetonitrile. Spectra were recorded at room temperature.

This feature indicates that the denatured state induced by acetonitrile is characterized by some degree of ordered tertiary structure in the N-terminal region, and that this structure is similar in the two protein variants.

The results obtained by the local structural probe of Trp side chains can be combined with information about global structural compactness obtained by nano-ESI-MS. Figure 6.4 shows nano-ESI-MS spectra of each AT3 variant at increasing acetonitrile concentrations in the range 10-75%. The organic solvent induces unfolding of both proteins, resulting in bimodal CSDs and progressive accumulation of highly charged components (16+ for AT3/182 and 29+ for AT3/291). The loss of the AT3/291 form centered on the 26+ ion during the transition indicates that this form contains structure and supports its identification as the protein with folded Josephin domain and disordered region 183-291. The most highly charged components, 16+ for AT3/182 and 29+ for AT3/291, progressively accumulates at increasing acetonitrile concentrations without further shift. Thus, they can be identified as the fully unfolded forms. Therefore, the unfolding of the Josephin domain is reflected by the transition from main charge state 9+ to 16+ in AT3/182 and from 26+ to 29+ in AT3/291. The unfolding of the compact conformation of the disordered region, instead, is reflected by the transition from main charge state 13+ to 26+ in AT3/291. This transition occurs earlier during the titration, compared to the unfolding of the Josephin domain, consistent with a more labile structure. Figure 6.4 also shows that, at high concentrations of acetonitrile, an intermediate species accumulates in both protein samples. Such species is centered on the 12+ ion for AT3/182 and on the 23+ ion for AT3/291 and becomes resolved from both the compact and the fully unfolded form of the respective protein. These species are characterized by intermediate compactness and represent, therefore, partially folded forms of each protein. In order to identify

and quantify the intermediate species confidently, the spectra were converted to z as abscissa axis and deconvoluted by Gaussian fit (data not shown). One intermediate is detected in AT3/182, while two distinct intermediates are identified in the AT3/291 spectra. The AT3/291 intermediates also accumulate to a larger extent. At 75% acetonitrile, the relative amount of the AT3/182 intermediate is ~9%, while the AT3/291 intermediates represent ~27% and ~14% of the molecular population. Accumulation of a partially folded conformation at high acetonitrile concentrations is consistent with the near-UV CD data reported above. That chemical unfolding of AT3 proceeds *via* an intermediate had first been postulated based on fluorescence profiles in the presence of GuHCl (Masino et al. 2004). Here, the intermediate is directly visualized, showing its accumulation under conditions of reduced solvophobic effect.
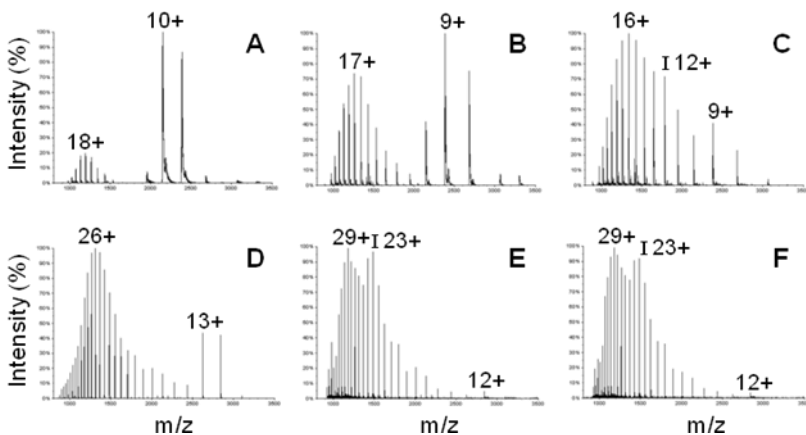
Figure 6.4 Nano-ESI-MS spectra of 10 µM AT3/182 (A-C) and 10 µM AT3/291 (D-F). Samples are in 5 mM ammonium acetate pH 6.5 and variable concentrations of acetonitrile: 10% (A, D), 50% (B, E) and 75% (C, F). Numeric labels indicate the main charge state of each peak envelope. The intermediate form of each protein is labeled as I. Spectra were recorded at room temperature.

Dimerization and amyloid aggregation of the fragments

AT3 aggregation can be elicited by heat. Thus, nano-ESI-MS spectra at variable temperatures were recorded, in order to explore heat-induced changes in protein conformation and/or oligomerization. Figure 6.5 shows data acquired after equilibration of the instrument interface at 150-175°C. Samples containing 10 µM protein in 10 mM ammonium acetate pH 6.5 show bimodal CSDs, revealing accumulation of the same partially folded forms observed at high concentrations of acetonitrile, with main charge state 16+ for AT3/182 and 23+ for AT3/291. In both the spectra, dimer-specific signals can be detected, showing, themselves, a bimodal distribution. In other words, dimers with different compactness are present, indicating that also partially folded forms of the proteins can associate into oligomeric structures. These partially folded dimers could represent intermediates of the aggregation process that leads to protofibrils.
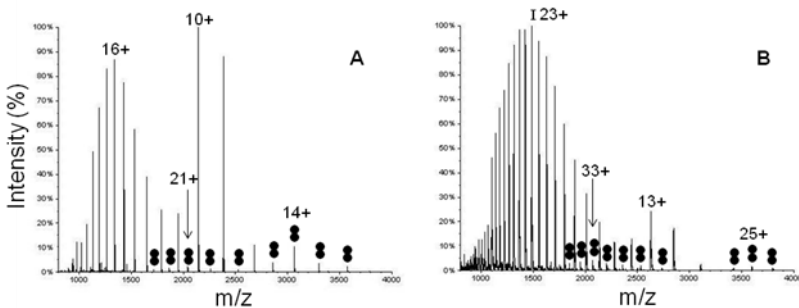
Figure 6.5 Nano-ESI-MS spectra of 10 µM AT3/182 (A) and 10 µM AT3/291 (B) in 5 mM ammonium acetate pH 6.5. Spectra were recorded at interface temperature 175 °C (A) and 150 °C (B). The labels show the main charge state of each peak envelope. The filled circles indicate the peaks unambiguously associated to dimers.

It should be pointed out that protein concentration is about one order of magnitude below the critical value that induces formation of unspecific dimers during protein electrospray. More studies will be needed to elucidate the aggregation pathway in which these species are involved.

The fibrillation kinetics of the two pure proteins has been monitored by ThT fluorescence. As shown in Figure 6.6, both proteins induce an increase in ThT fluorescence during the first 65 hours of incubation. However, AT3/291 displays enhanced kinetics relative to AT3/182, with a virtually absent lag phase and already approaching a plateau within 65 hours. AT3/182, instead, presents a much less fluorescence intensity respect to AT3/291, has a lag phase of about 15 hours and is still in exponential growth after 65 hours.
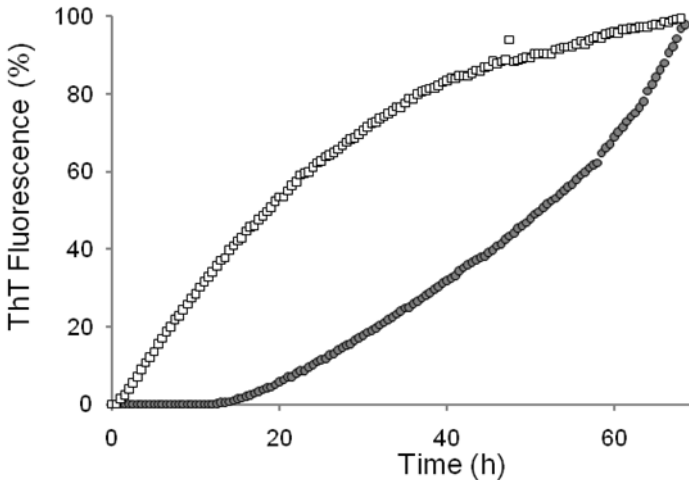
Figure 6.6 Fluorescence of 20 μM ThT in the presence of 25 μM AT3/182 (grey circles) or 25 μM AT3/291 (white squares) in PBS (20 mM KH$_2$PO$_4$, 150 mM NaCl), measured every 30 minutes for 65 hours. Each curve is normalized to its highest intensity.

The aggregation process has also been monitored by FT-IR in the Amide-I region. In agreement with the above ThT assays, the band of inter-molecular β-sheets and protein aggregates accumulates much faster in AT3/291 than in AT3/182 (data not shown). Altogether, the results presented in this section indicate that both protein variants form oligomeric species and fibrils under the employed experimental conditions, but that the longer fragment has a much higher amyloidogenic potential. Thus, the disordered region promotes aggregation even in the absence of the poly-Q tail and, therefore, likely contributes determining AT3 amyloidogenic properties *in vivo*. That the isolated Josephin domain can fibrillate *in vitro* had been already shown by Pastore and coworkers (Masino et al. 2004). Here it is shown that the addition of the disordered region 183-291 dramatically enhances aggregation. These findings suggest a direct involvement of residues belonging to the 183-291 segment in protein-protein interactions during the fibrillation process. The fact that the species with higher amyloidogenic potential (AT3/291) displays stronger propensity to accumulate partially folded forms under denaturing conditions (Figure 6.4) suggests that such forms might play a role promoting protein aggregation.

7. BOVINE β-LACTOGLOBULIN

The effect of TFE on BLG structure is analyzed by nano-ESI-MS and CD, in order to study the conformational properties of the protein in the presence of this cosolvent. Then, the aggregation process of the protein in TFE is monitored in order to identify covalent oligomeric intermediates. Finally, the role of Cys121 in the transition to the fibril form is analyzed by the employment of alkylated proteins.

BLG amyloid fibrils in TFE

TFE has been shown to promote amyloid aggregation of BLG (Dong et al. 1998; Gosal et al. 2002), but this condition has not been extensively studied as in the case of heat- or urea-induced fibrillation. In order to get a more detailed picture, the effect of TFE on the structure of the protein was analyzed by nano-ESI-MS and far-UV CD, monitoring the changes in quaternary, tertiary and secondary structures at increasing percentages of the cosolvent (Figure 7.1). A relatively small amount of TFE is enough to completely disassemble BLG native dimers. Indeed, at 7.5% TFE (Figure 7.1B) nano-ESI-MS spectra show that the protein is almost entirely monomeric, with no traces of the peaks that are characteristic of the dimeric population in the absence of the cosolvent (see Figure 9.3B in Appendix A). Furthermore, a new distribution centered on the 14+/15+ ion progressively increases in response to TFE addition, indicating unfolding of the tertiary structure of the protein. At 40% TFE this component dominates the spectrum, and the native CSD has almost completely disappeared

(Figure 7.1C). TFE also induces a transition from a prevalently β-sheet structure – characteristic of native BLG – to a prevalently α-helical conformation (Figure 7.1A). This change in secondary structure takes place at higher TFE percentages with respect to the transitions monitored by nano-ESI-MS, having a midpoint at ~15% TFE. Thus, the distinct structural levels have individual responses to TFE: quaternary structure is easily disassembled, the tertiary fold is a little more stable, and secondary structure is further more stable (Figure 7.1D). However, above 40% TFE protein conformation is completely reorganized respect to the native state.
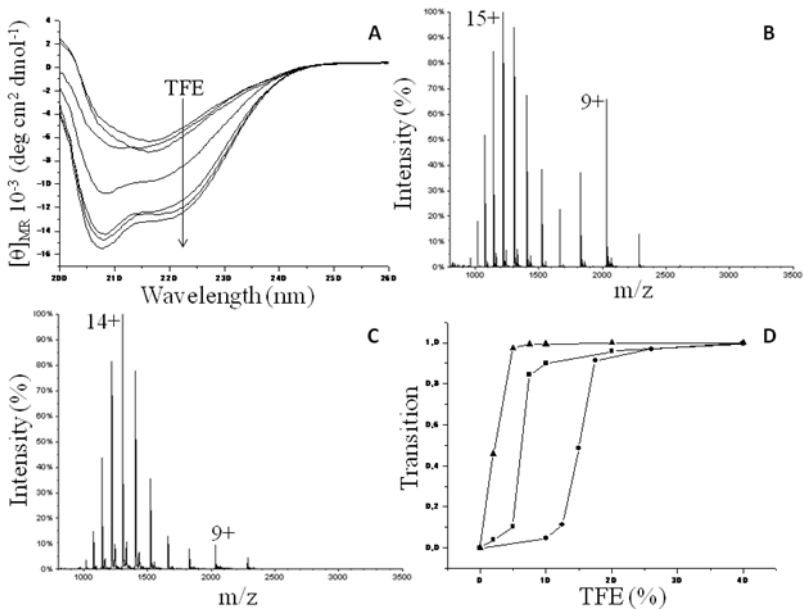
Figure 7.1 Response to TFE of 10 μM BLG in 10 mM ammonium acetate. (A) far-UV CD spectra in 0-40% TFE. (B) nano-ESI-MS at 7.5% TFE. (C) nanoESI-MS at 40% TFE. (D) change in BLG distinct structures as a function of TFE: disassembly of the dimer monitored by Bayesan reconstructions (▲), unfolding of monomer monitored by Gaussian fits (■), change of secondary structure monitored by mean residue [θ] at 222 nm (●).

Next, the formation of intermolecular disulfide bridges during BLG aggregation in TFE was monitored. The protein was incubated at room temperature in 50% TFE for 2 months and several aliquots during the incubation were analyzed, after clarification, by nano-ESI-MS and far-UV CD.

One week after the beginning of incubation, a dimeric population starts to be detected in nano-ESI-MS spectra (Figure 7.2A) and progressively accumulates during the incubation (Figure 7.2B). The same samples were also subjected to far-UV CD analysis (Figure 7.2C). The characteristic α-helical profile does not change during incubation, suggesting that the secondary structure of the soluble material is maintained constant. However, signal intensity progressively approaches to zero, confirming that the protein accumulates in non-soluble aggregates removed by the clarification step or not sensitive to absorption in this range of wavelengths. The presence of non-soluble aggregates at the end of the incubation period is also confirmed by Fourier-transform infrared spectroscopy: the characteristic absorption band at ~1620 cm^{-1}, typical of intermolecular β-sheets, is detected in the pellet of the centrifugation (data not shown).

The dimer population survives in the nano-ESI-MS spectra even upon dilution and further addition of denaturants (data not shown). Furthermore, the spectra reported in Figure 7.2 were collected under quite harsh instrumental conditions (IHT 125°C, DP 150V). This evidence already suggests that the observed dimers are covalently linked. To test this hypothesis an aliquot of the sample was incubated in 1 mM DTT at room temperature for one hour. The addition of DTT dissociates the complex, as assessed by the loss of its diagnostic odd peaks in the spectrum (Figure 7.3). These results suggest that the detected dimers form by intermolecular disulphide bridging.

In order to shed light on the role of covalent dimers in the aggregation mechanism, the free thiol of the protein has been

blocked by an alkylation reaction, obtaining a modified BLG (alkBLG). The alkylation has been performed by bromoethanol in acetonitrile, employing a standard protocol (Hale et al. 2004). Reaction completion has been confirmed by nano-ESI-MS: alkBLG presents a mass shift of 44 Da relative to BLG (18277 Da), consistent with the addition of one ethanol group to the free thiol (Figure 7.4B). The modification does not comport any relevant change in secondary or tertiary structures for the protein in 50% TFE, as confirmed by a comparison of BLG and alkBLG analyzed by CD and intrinsic fluorescence (Figure 7.4C,D).
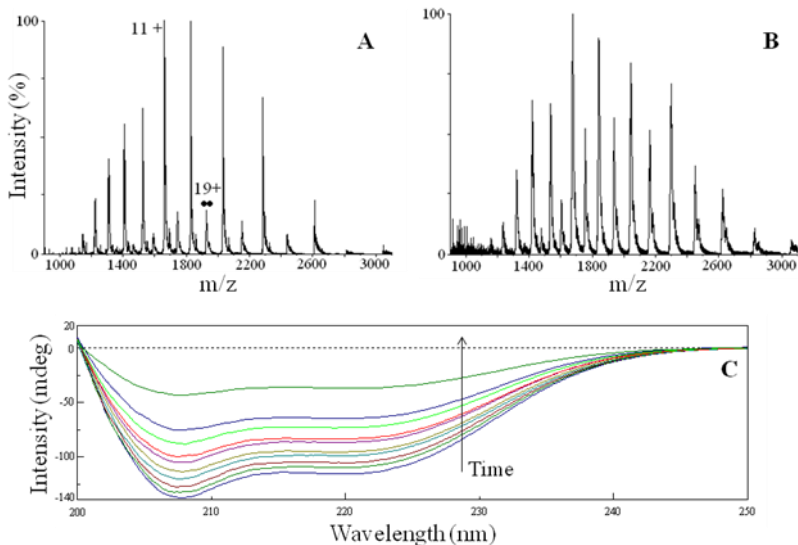
Figure 7.2 Incubation of 35 μM BLG in 10 mM ammonium acetate, 50% TFE. (A) Nano-ESI-MS spectrum after one week. The most intense peak of the monomer (11+) and of the dimer (●●, 19+) are labeled. (B) NanoESI-MS spectrum after four weeks. (C) far-UV CD spectra after 0-8 weeks.

BLG and alkBLG have been incubated in 50% TFE in the presence of ThT, to test whether there are differences in the aggregation kinetics of the two species. Figure 7.5 shows the normalized ThT fluorescence for BLG and alkBLG as a function of incubation time in 50% TFE. After a small lag-phase of ~3 hours, fluorescence intensity starts to rise. Despite this, BLG and alkBLG present distinct fluorescence responses. In the analyzed time interval (48 hours) the BLG fluorescence profile reveals three main growth phases: a small increase after ~3 hours, a second, big increment at ~12 hours and a further growth after ~32 hours.
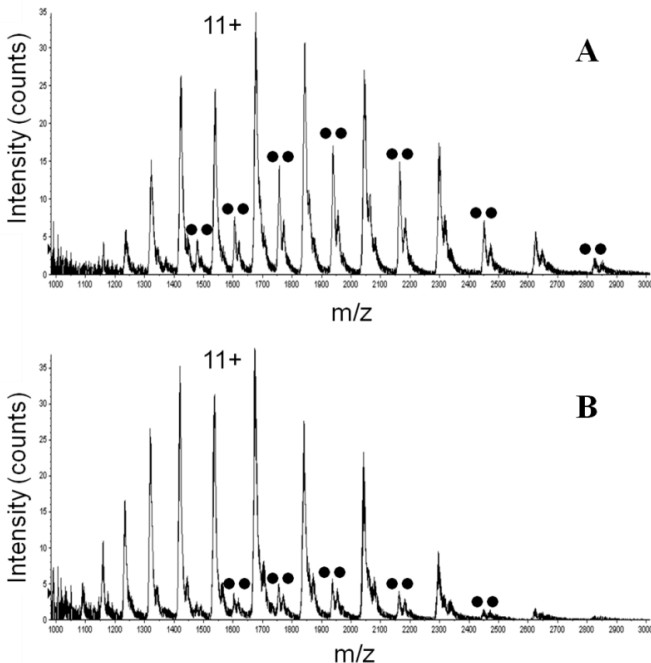
Figure 7.3 Nano-ESI-MS spectra of 35 μM BLG in 10 mM ammonium acetate, 50% TFE (A) after 28 days (B) after 28 days and 1 hour in 1 mM DTT at room temperature. The most intense peak of the monomer (11+) and the diagnostic peaks of the dimer (●●) are labeled.

On the other hand, alkBLG has a smaller increment after 3 hours and does not present any change in fluorescence intensity after 30 hours. This result suggests that Cys121 plays a role in the formation of species prone to ThT binding in the presence of TFE. Therefore, cystein oxidation could favor formation of conformational and oligomeric intermediates of the fibrillation process.

The study of BLG aggregation in our laboratories is still at a preliminary stage. Further experiments are in progress to better characterize the species that accumulate during the incubation in TFE and discriminate on- and off-patway intermediates of the fibrillation process.
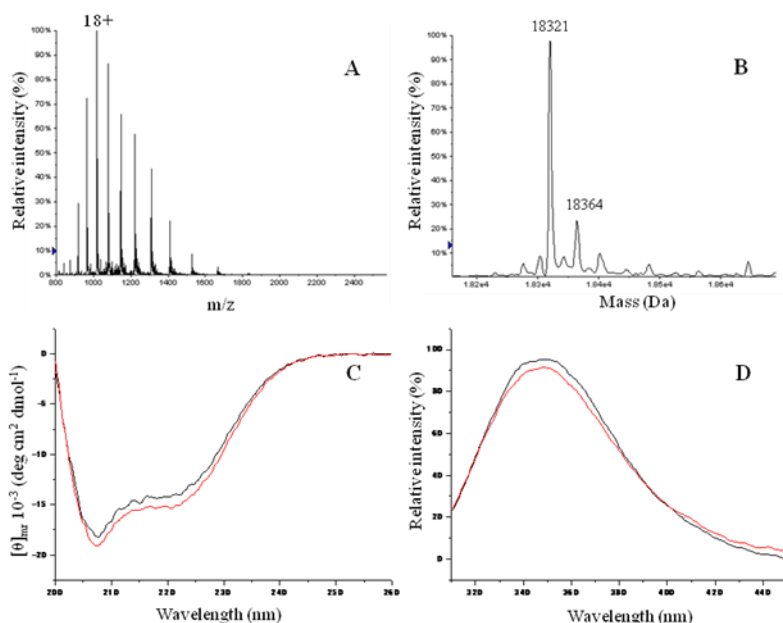
Figure 7.4 Nano-ESI-MS spectrum of alkBLG after alkylation reaction (A). Mass deconvolution of panel A confirms alkylation in one position (18321 Da) and, partially, also in second position (18364 Da) (B). Far-UV CD (C) and intrinsic fluorescence (D) spectra of BLG (red) and alkBLG (black) in 50% TFE.
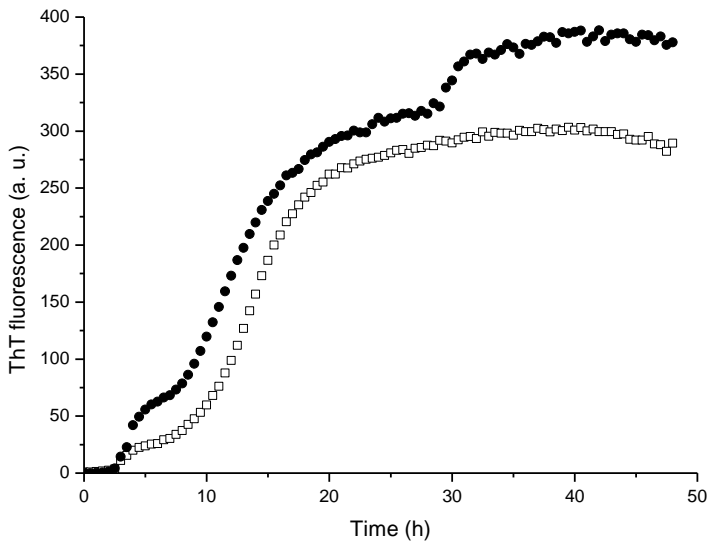
Figure 7.5 Fluorescence of 20 μM ThT in the presence of 400 μM BLG (black circles) or 400 μM alkBLG (white squares) in milli-Q water, measured every 30 minutes for 48 hours. The signal of the ThT without proteins was subtracted.

8. CONCLUSIONS AND PERSPECTIVES

A full elucidation of amyloid fibrillation is possible only by the characterization of the intermediate species that become populated during the process. However, this goal is particularly challenging for two main reasons: i) the intermediate species are often populated only transiently and in a low amount, and are therefore difficult to be detected and characterized ii) a complex pattern of competing pathways is present under fibrillation conditions, and it is therefore difficult to discriminate in this network among on- and off-pathway species. In spite of these problems, the results obtained in this work provide some novel information to the characterization of the molecular ensembles of the three considered proteins under fibrillation conditions.

The conformational intermediates of β2m under two different aggregation conditions have been characterized, thanks to the complementary information of nano-ESI-MS and CD. These two species show similar compactness but distinct secondary structures. New experiments are in progress to relate these features with the differences in fibrillation kinetics and to detect the oligomeric intermediates at pH 2.5 or at neutral pH in the presence of TFE. Our results on the mutants in the DE-loop confirm previous evidence, highlighting the importance of this region for protein stability. New mutants have been designed and produced and are currently under investigation in order to obtain further details on the roles played by the different protein regions in the fibrillation mechanism. In order to characterize β2m and its mutant variants better, ion mobility mass spectrometry is going to be employed in addition to CSD analysis.

The features of the disordered region 183-291 in ataxin-3 have been analyzed for the first time, since previous studies were prevalently

focused on the poly-Q tract or on the Josephin domain. The disordered region has been shown to promote amyloid aggregation and to affect protein conformation under both denaturing and non-denaturing conditions, indicating an active role in the fibrillation process played by the domains flanking the poly-Q tract. Our research in this field is developing in two directions: first of all, experiments similar to those presented here are going to be performed on murine variants AT3/182 and AT3/291. This choice is motivated by the fact that human and murine species are almost identical in the Josephin domain (only two conservative mutations) but are extremely different in the 183-291 region (only 20% identity). Therefore, a comparison of human and murine variants could lead to a better characterization of the disordered tract. On the other hand, we also want to study the intrinsic properties of the isolated fragment 183-291 in terms of amyloidogenity and conformation.

BLG has been extensively used as a reference protein in our laboratory, and has also represented a useful model system for the implementation of the different methods for this PhD thesis. BLG fibrillation is a very young topic in our research program: however, the preliminary studies reported here indicate feasibility of this kind of analysis for BLG aggregation induced by TFE and indicates targets for further studies.

In conclusion, the employment of complementary biophysical techniques has helped to shed some light on the aggregation mechanism of different amyloid proteins. In particular, nano-ESI-MS has emerged as an extremely versatile method that permits combined analysis of oligomeric state and conformation. Coupled with conventional spectroscopic techniques can give valuable information on amyloid protein aggregation.

9. APPENDIX A

BLG-ANS interaction

Measurements on protein-ANS interaction are performed with a double intent: to detect an important, pH-dependent structural change – the Tanford transition – and to test the possibility to monitor conformation-dependent ANS interactions by nano-ESI-MS (Santambrogio and Grandori 2008). The effect of solution pH on the complex between BLG and ANS is investigated at equimolar (2 μM) protein and ligand concentrations (Figure 9.1). ANS (MW: 299 Da) appears to bind to a much lesser extent to BLG at pH 5.9 than at pH 7.9 (panels A and B). This effect is evident both as relative abundance and stoichiometry of the detectable adducts. This result suggests that, against the background of nonspecific binding, conformation-specific complexes contribute to the spectrum at pH 7.9. Indeed, the BLG conformational changes by the Tanford transition can explain stronger binding at pH 7.9, while adduct formation by electrostatic interactions, if anything, should be reduced increasing the pH. Therefore, the additional binding observed at pH 7.9 relative to pH 5.9 most likely reflects the protein conformational change. Exposure of hydrophobic binding sites, thus, seems to contribute to ANS binding, although also surface binding sites might become better available to complex formation as a consequence of lid opening. Lysozyme is considered as a negative control. This protein, which does not display major conformational transitions between pH 6 and 8, gives almost identical ANS adducts at those pH values (data not shown). Thus, it seems that both binding modes of ANS to BLG, on the protein surface and inside the calyx, can be detected by nano-ESI-MS and both seem involved in the complexes

observed with the protein in its open conformation. Figure 9.1C shows data for the BLG-ANS complexes at pH 2.5. The nano-ESI-MS spectra under these conditions display more adducts than at pH 5.9, but still less than at pH 7.9. The BLG structure is known to be quite resistant to acid denaturation (Fogolari et al. 1998). Accordingly, the CSD reveals a still folded protein, although a shift of the main charge state from 9+ to 10+ is suggestive of a slightly perturbed structure relative to neutral pH.
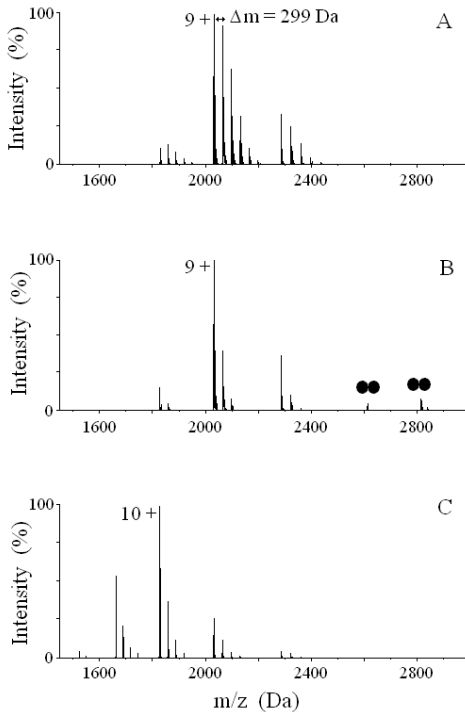

Figure 9.1 Nano-ESI-MS spectra of 2 μM BLG in the presence of 2 μM ANS, in 10 mM ammonium acetate pH 7.9 (A), pH 5.9 (B), and pH 2.5 (C). The main peak of the protein without bound ANS is labeled by the corresponding charge state. Peaks corresponding to the dimeric protein are labeled by double circles.

However, the protein should be in its closed conformation at this pH, making the calyx not accessible. Thus, more extensive binding of ANS to BLG at pH 2.5 than at pH 5.9 could be explained by a less compact structure or by the higher positive charge of the protein, or both. The data reported here are in agreement with a previous report by Molinari and coworkers, which also shows stronger binding of ANS to BLG at pH 2.5 than at pH 6 by nano-ESI-MS (Hamdan et al. 1996). In agreement with previous fluorescence studies and with the nano-ESI-MS results reported here, the ANS fluorescence in the presence of BLG at pH 7.9 is much stronger than at pH 5.9 (Figure 9.2). It has been shown previously that ANS fluorescence in the presence of BLG at pH 6.3 and 8.2 is not reduced, but rather slightly enhanced, at increasing ionic strength, indicating that electrostatic interactions do not play an important role in ANS binding to BLG in solution and in this pH range (D'Alfonso et al. 1999). Thus, fluorescence enhancement going from pH ~6 to pH ~8 seems to be due to stronger binding mediated by hydrophobic interactions and can be interpreted in terms of protein conformational changes during the Tanford transition. On the other hand, even stronger fluorescence than at pH 7.9 is observed at pH 2.5 (Figure 9.2). This result is in agreement with previous fluorescence studies but in contrast with the nano-ESI-MS results reported here, which show weaker binding at pH 2.5 than at pH 7.9. As shown previously, ANS fluorescence in the presence of BLG at pH 2.5 is strongly reduced at increasing ionic strength and this effect is largely due to changes in the binding affinity rather than to quantum-yield variations (D'Alfonso et al. 1999). Thus, fluorescence enhancement at acidic pH seems to be due to stronger binding mediated by electrostatic interactions and can be assumed to take place on the protein surface, both because the protein is in its closed conformation and because charged residues are mostly exposed to the solvent.

The results obtained with BLG suggest that ANS binding within a pocket is better detectable by nano-ESI-MS than ANS binding to the protein surface, regardless of the nature of the interaction that has led to formation of the complex in solution. Apo-Myoglobin (apoMb) was chosen as another example of a protein structure with a preformed hydrophobic pocket (Stryer 1965). This protein also offers the possibility to compare the folded structure to the well-characterized molten globule induced by moderate acidification (Hughson et al. 1990).
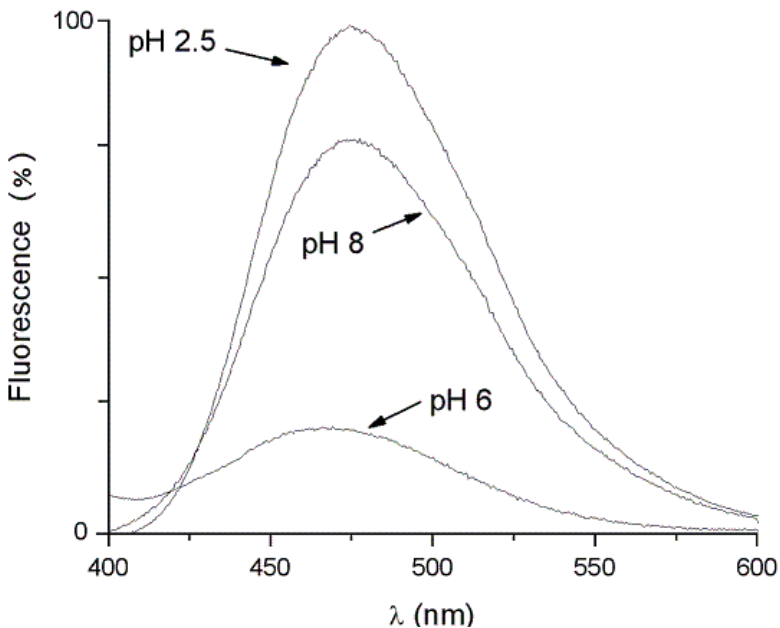
Figure 9.2 Emission fluorescence spectra of 2 µM BLG in the presence of 2 µM ANS, in 10 mM ammonium acetate at variable pH. The wavelength for excitation is 374 nm.

Our results support the hypothesis that ANS binding in the heme pocket at pH 6.7 is well detectable under nano-ESI-MS conditions. However, binding to apoMb is reduced at pH 5, a condition where a molten globule state and a perturbed, native-like conformation coexist. On the other hand, fluorescence spectra of the same apoMb samples confirm that, in solution, stronger ANS binding takes place at pH 5 than at pH 6.7, in agreement with the affinity enhancement generally observed with proteins in the molten-globule state (data not shown).

The results reported here suggest that ANS complexes with disordered protein structures are not stable enough under nanoESI conditions for the application of this method to the study of folding intermediates, as typically done by fluorescence spectroscopy in solution. Thus, it seems that a well-structured cavity like the BLG calyx or the apoMb heme pocket is important for efficient detection of protein-ANS complexes by nano-ESI-MS.

BLG native dimerization

Exploiting the ability of mass spectrometry to directly detect the distinct species present in a heterogeneous sample, the relative amount of monomer and native dimer is monitored as a function of total protein concentration (Figure 9.3). The experiments were performed at pH 6.7 in 10 mM ammonium acetate, a good condition for the detection of BLG dimer by nano-ESI-MS (Invernizzi et al. 2006). At the lowest concentration tested (1 μM, panel A) a single CSD centered on the 9+ ion dominates the spectrum: the deconvolution of its peaks returns a mass of 18277 ± 0.5 Da, that matches with the molecular weight of monomeric BLG. The value of 9+ charges for this mass is in agreement with the theoretical average charge state for a globular protein of this size (Heck and Van Den

Heuvel 2004). The homogeneity of the compact fold is also confirmed by the sharpness of the CSD. Only three peaks characterize the CSD under these conditions. Thus, BLG monomer is able to populate a folded state even if not bounded in the dimeric form. A second, small distribution is detected at higher m/z values, and corresponds to the mass of the dimer. The features of this CSD, too, suggest a compact conformation, leading to the conclusion that native-like dimers have been detected. Each subunit in the dimer, indeed, acquires an average of 6.5+ charges.
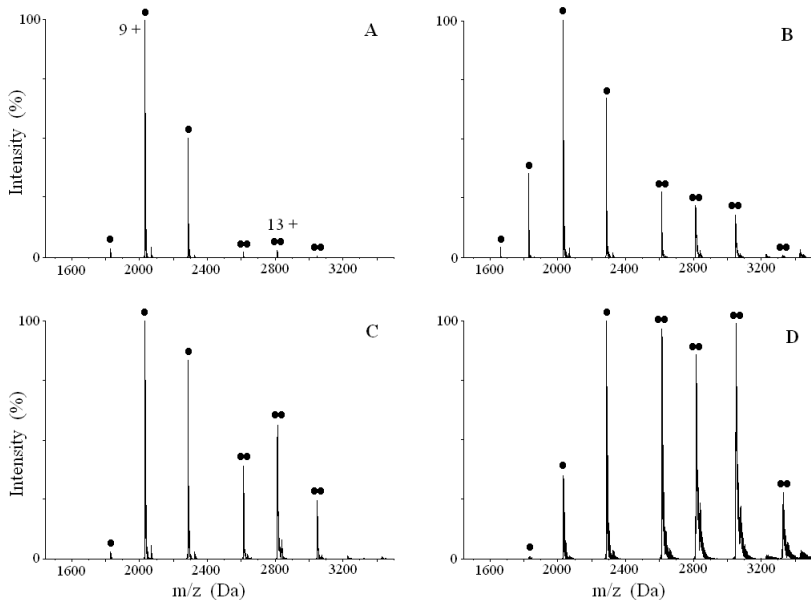
Figure 9.3 Nano-ESI-MS spectra of BLG in 10 mM ammonium acetate at room temperature as a function of protein concentration. (A) 1 µM, (B) 10 µM, (C) 20 µM, (D) 40 µM. Peaks corresponding to prevalent monomer (●) and prevalent dimer (●●) are labeled. In panel (A) the main peaks corresponding to monomer and dimer are labeled by their charge state.

As the protein concentration in the sample is increased, the fraction of dimers progressively rises, without substantial changes in the CSDs (Figure 9.3, panels B,C and D). The concentration dependence of the relative amount of the species suggests that the latter reflects the pre-existing equilibrium in the liquid sample.

It is possible to shift the monomer-dimer equilibrium by changing the interface temperature (IHT) of the mass spectrometer. This parameter is usually employed to improve the instrument sensitivity, since high temperature permits a better desolvatation of the droplets. However, in this study the heat provided by the interface has been used to affect the stability of both BLG quaternary and tertiary structures. For instance, the increase of temperature at 20 μM protein concentration comports a lower intensity in the CSD associated to the dimeric population (Figure 9.4), that almost completely disappeares at the highest tested IHT (150°C). This behavior is particularly highlighted by a diagnostic peak of the dimer, the 13+ charge state, since it is the only peak of the complex that does not overlap with the CSD of the monomer. In the tested range of temperatures, there is no shift in the peak envelope of the dimer, which remains centered on the 13+ charge state. These results suggest that the main effect caused by the rise of temperature is a thermal dissociation of the complex, without significant changes in the conformation of the survived dimers. The CSD of the monomer is constant over a wide interval of IHT values, becoming bimodal only at very high temperatures. Indeed, at 150°C a new component at lower m/z values appears, suggesting that thermal unfolding occurs. This confirms that the tertiary structure of the monomer is not only independent of dimerization, but also very stable to thermal denaturation.

The separate detection of monomer and dimer by nano-ESI-MS at different concentrations and temperatures permits to estimate the fractions of the two species in distinct conditions by Bayesan

reconstructions, and thus to perform a thermodynamic analysis on the monomer-dimer equilibrium.

The dissociation reaction dimer $\leftrightarrow$ 2 monomers can be analyzed as a function of protein concentration and temperature. In the hypothesis that the measured fractions of the two species reflect their relative concentrations, it follows that monomer and total concentrations are related by a parabolic equation (demonstration is in Appendix B):
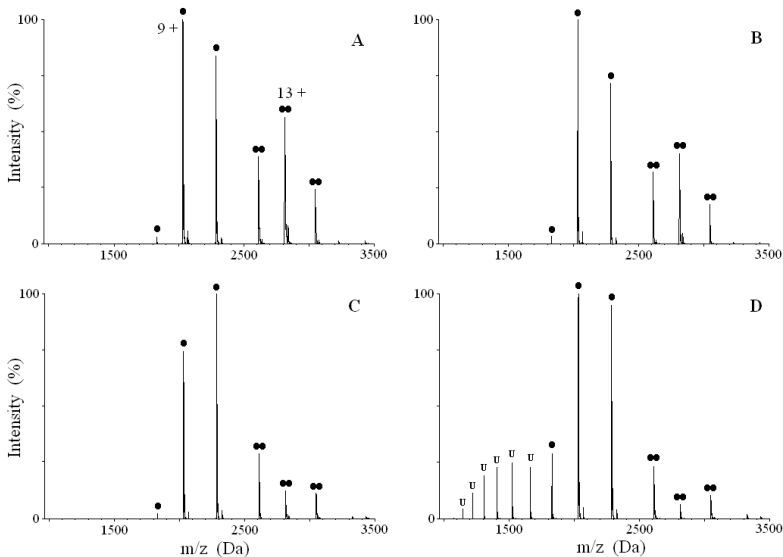
$$[T] = \frac{2}{K}[M]^2 + [M]$$

Figure 9.4 Nano-ESI-MS spectra of 20 μM BLG in 10 mM ammonium acetate as a function of IHT. (A) turned off, (B) 50°C, (C) 100°C, (D) 150°C. Peaks corresponding to folded monomer (●), unfolded monomer (U) and dimer (●●) are labeled. In panel (A) the main peaks corresponding to monomer and dimer are labeled by their charge state.

where K is the apparent dissociation constant. Thus, calculating the monomer concentration from nano-ESI-MS spectra, it is possible to estimate the apparent dissociation constant at each temperature by parabolic fits (Figure 9.5). Then, a Van't Hoff analysis of the results (Figure 9.5, insert) allows to calculate the standard enthalpy and entropy variations of dimer dissociation by the following linear equation (Kelly and Reithel 1971):

$$\ln K = -\frac{\Delta H^0}{R} \cdot \frac{1}{T} + \frac{\Delta S^0}{R}$$

where R is the universal gas constant.

The apparent dissociation constant at room temperature obtained by this analysis is $K = (10.8 \pm 0.6)\,\mu M$. This relatively high value is consistent with those calculated at pH 6.5 by different methods (Sakai et al. 2000; Sakurai and Goto 2002), and confirms the weakness of the binding between dimer subunits. The results for standard enthalpy and entropy variations are $\Delta H^0 = (35 \pm 4)\,kJ/mol$ and $\Delta S^0 = (19 \pm 12)\,J/mol \cdot K$, corresponding to $\Delta G^0 \approx 7\,kcal/mol$ at 20°C, also in agreement with previous works (Sakurai and Goto 2002; Bello et al. 2008). It is quite surprising that the numerical results obtained by nano-ESI-MS match so well the data of conventional techniques. Several arguments, indeed, would argue against the application of nano-ESI-MS for quantitative thermodynamic analysis. Primarily, the strength of the interactions changes while passing from the liquid to the gas-phase. Moreover, the equilibrium condition is lost. Finally, in the here described experiments, there is no control on the real temperature inside the droplets. It is possible to clarify this discrepancy adducing two main arguments: first of all, in a nano-ESI-MS measurement the time elapse between desolvation and

detection is only few hundreds of μs. Thus, below a certain threshold in the kinetic dissociation constant the detection of the complex is possible, in spite of the lost of the thermodynamic equilibrium. Secondly, the increased strength of electrostatic interactions and hydrogen bonds in the gas-phase can contribute to maintenance of native-like structure (Grandori 2003b; Marchese et al. 2010). These properties suggest that, even if a completely different environment characterizes the ions analysed by nano-ESI-MS, this technique could return to us a sort of snapshot of the pre-existing scenario in the liquid phase (Invernizzi et al. 2007).
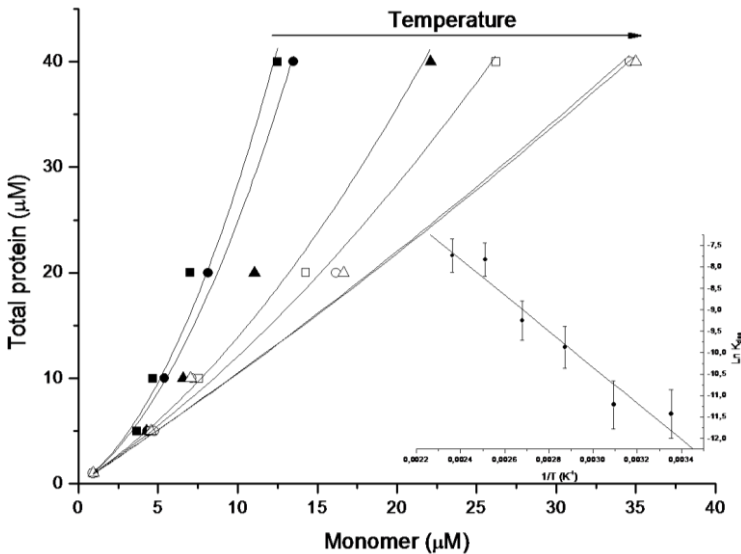
Figure 7.5 Parabolic fits to obtain the apparent dissociation constants at different IHTs. Turned off (25°C) (■), 50°C (●), 75°C (▲), 100°C (□), 125°C (○), 150°C (Δ). Insert: fit by Van't Hoff equation to obtain the thermodynamic parameters of dimer dissociation.

To validate the experiments carried out in this section, a comparison with a canonical method – like isothermal titration calorimetry (ITC) – under the same conditions will be necessary. Indeed, ITC measures the heat exchanged by the system after the injection of an aliquot of concentrated protein into a calorimetric cell, as a consequence of complex dissociation. This method has been already employed for the study of BLG dimer dissociation (Bello et al. 2008), and can be an useful techniques to confirm the results obtained by nano-ESI-MS.

10. APPENDIX B

Gaussian fit

In order to quantify the relative amount of the distinct components of nano-ESI-MS spectra, deconvolution of raw data have been performed by Gaussian multi-component reconstruction. Mass spectra have been transformed from $x=m/z$ to $x=z$ as abscissa axis and data points have been fitted by a series of Gaussian functions:

$$y = y_0 + \frac{A}{w\sqrt{\pi/2}} e^{-\frac{2(x-x_0)^2}{w^2}}$$

were y_0 is the base line off-set, A is the area under the curve, w is the width of the curve at half height and x_0 is the abscissa of the maximum. The number of component and their position was let vary, in order to obtain a final interpolation with $R^2 > 0.99$. The analysis was performed by the software Origin7 (Originlab, Northampton, MA, USA).

Estimation of far-UV CD spectra for β2m intermediate species

For the calculation of the far-UV CD spectrum of the intermediate at 20% TFE, the mean residue ellipticity at each wavelength was obtained by the equation:

$$[\vartheta]_X = \frac{[\vartheta]_{N+X} - f_N [\vartheta]_N}{f_X}$$

where $[\vartheta]_X$, $[\vartheta]_N$ and $[\vartheta]_{N+X}$ are the values for the intermediate (X), the native protein at neutral pH and 0% TFE (N) and their mixture at neutral pH and 20% TFE (N+X). The relative fractions f_N and f_X were estimated by Gaussian fitting of the nanoESI-MS spectra as described above. For the intermediate at pH 2.5, instead, the following equation was employed:

$$[\vartheta]_X = \frac{[\vartheta]_{N+X} - f_D [\vartheta]_D - f_N [\vartheta]_N}{f_X}$$

where the reference ellipticity values of the random coil $[\vartheta]_D$ were assumed, simplistically, to describe the protein in the denatured state. The relative fractions f_N, f_D, and f_X were estimated by Gaussian fitting of the nanoESI-MS spectra as described above.

Analysis of BLG native dimerization

The reaction under examination is the dissociation of the dimer into two monomers:

$$D \leftrightarrow 2M$$
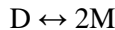
Total protein concentration $[T]$ is given by the sum of monomer concentration $[M]$ and twice dimer concentration $[D]$ (since each dimer is formed by two molecules). Assuming that species concentrations are proportional to the relative peak areas in the

deconvoluted mass spectrum, it is possible to find the monomer concentration in each sample:

$$\begin{cases} [T] = [M] + 2[D] \\ \dfrac{A_M}{A_D} = \dfrac{[M]}{[D]} \end{cases} \quad \rightarrow \quad [M] = \dfrac{A_M[T]}{A_M + 2A_D}$$

From the definition of the apparent dissociation constant it follows that total and monomer concentrations are related by a parabolic function:

$$\begin{cases} [T] = [M] + 2[D] \\ K = \dfrac{[M]^2}{[D]} \end{cases} \quad \rightarrow \quad [T] = \dfrac{2}{K}[M]^2 + [M]$$

Fitting the data with the parabola $y = Cx^2 + x$ it is so possible to estimate the apparent dissociation constant. The previous procedure permits to find K values for all the temperatures tested. Using Van't Hoff equation and the relationship between the apparent dissociation constant and the main thermodynamic functions we obtain:

$$\begin{cases} \left.\dfrac{\partial \ln K}{\partial T}\right|_P = \dfrac{\Delta H^0}{RT^2} \\ \Delta G^0 = -RT \ln K = \Delta H^0 - T\Delta S^0 \end{cases} \quad \rightarrow \quad \ln K = -\dfrac{\Delta H^0}{R} \cdot \dfrac{1}{T} + \dfrac{\Delta S^0}{R}$$

where ΔG^0, ΔH^0 and ΔS^0 are the Gibbs energy, enthalpy and entropy standard variations, T is the temperature (in Kelvin degrees) and R is the universal gas constant. From the linear interpolation of the data in a Van't Hoff plot it is possible to find the values of the thermodynamic functions.

REFERENCES

Adams, P.D., Chen, Y., Ma, K., Zagorski, M.G., Sonnichsen, F.D., McLaughin, M.L., and Barkley, M.D. 2002. Intramolecular quenching of tryptophan fluorescence by the peptide bond in cyclic hexapeptides. *J. Am. Chem. Soc.* **124:** 9278-9286.

Akashi, S. 2006. Investigation of molecular interaction within biological macromolecular complexes by mass spectrometry. *Med. Res. Rev.* **26:** 339-368.

Altekar, W. 1977. Fluorescence of proteins in aqueous neutral salt solutions. I. Influence of anions. *Biopolymers* **16:** 341-368.

Šamalikova, M., Carey, J., and Grandori, R. 2005. Assembly of the hexameric Escherichia coli arginine repressor investigated by nano-ESI-TOF mass spectrometry. *Rapid Commun. Mass Spectrom.* **19:** 2549-2552.

Andersen, E.S., Dong, M., Nielsen, N.M., Jahn, K., Subramani, R., Mamdouh, W., Golas, M.M., Sander, B., Stark, H., Oliveira, C.L., et al. 2009. Self-assembly of a nanoscale DNA box with a controllable lid. *Nature* **459:** 73-76.

Ayed, A., Krutchinsky, A.N., Ens, W., Standing, K.G., and Duckworth, H.W. 1998. Quantitative evaluation of protein-protein and ligand-protein equilibria of a large allosteric enzyme by electrospray ionization time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* **12:** 339-344.

Ban, T., Hamada, D., Hasegawa, K., Naiki, H., and Goto, Y. 2003. Direct observation of amyloid fibril growth monitored by thioflavin T fluorescence. *J. Biol. Chem.* **278:** 16462-16465.

Baxa, U., Cassese, T., Kajava, A.V., and Steven, A.C. 2006. Structure, function, and amyloidogenesis of fungal prions: filament polymorphism and prion variants. *Adv. Protein Chem.* **73:** 125-180.

Baxa, U., Speransky, V., Steven, A.C., and Wickner, R.B. 2002. Mechanism of inactivation on prion conversion of the

Saccharomyces cerevisiae Ure2 protein. *Proc. Natl. Acad. Sci. USA* **99:** 5253-5260.

Bello, M., Perez-Hernandez, G., Fernandez-Velasco, D.A., Arreguin-Espinosa, R., and Garcia-Hernandez, E. 2008. Energetics of protein homodimerization: effects of water sequestering on the formation of beta-lactoglobulin dimer. *Proteins* **70:** 1475-1487.

Benesch, J.L., Ruotolo, B.T., Simmons, D.A., and Robinson, C.V. 2007. Protein complexes in the gas phase: technology for structural genomics and proteomics. *Chem. Rev.* **107:** 3544-3567.

Bevivino, A.E., and Loll, P.J. 2001. An expanded glutamine repeat destabilizes native ataxin-3 structure and mediates formation of parallel beta -fibrils. *Proc. Natl. Acad. Sci. USA* **98:** 11955-11960.

Bhak, G., Choe, Y.J., and Paik, S.R. 2009. Mechanism of amyloidogenesis: nucleation-dependent fibrillation versus double-concentrated fibrillation. *BMB Rep.* **42:** 541-551.

Bilen, J., and Pittman, R.N. 2005. The polyglutamine neurodegenerative protein ataxin 3 regulates aggresome formation. *Proc. Natl. Acad. Sci. USA* **102:** 4330-4335.

Borysik, A.J., Radford, S.E., and Ashcroft, A.E. 2004. Co-populated conformational ensembles of beta2-microglobulin uncovered quantitatively by electrospray ionization mass spectrometry. *J. Biol. Chem.* **279:** 27069-27077.

Bousset, L., Thomson, N.H., Radford, S.E., and Melki, R. 2002. The yeast prion Ure2p retains its native alpha-helical conformation upon assembly into protein fibrils in vitro. *EMBO J.* **21:** 2903-2911.

Bromley, E.H., Krebs, M.R., and Donald, A.M. 2005. Aggregation across the length-scales in beta-lactoglobulin. *Faraday Discuss.* **128:** 13-27.

Brownlow, S., Morais Cabral, J.H., Cooper, R., Flower, D.R., Yewdall, S.J., Polikarpov, I., North, A.C., and Sawyer, L. 1997. Bovine

beta-lactoglobulin at 1.8 Å resolution-still an enigmatic lipocalin. *Structure* **5:** 481-495.

Burova, T.V., Choiset, Y., Tran, V., and Haertle, T. 1998. Role of free Cys121 in stabilization of bovine beta-lactoglobulin B. *Protein Eng.* **11:** 1065-1073.

Calabrese, M.F., Eakin, C.M., Wang, J.M., and Miranker, A.D. 2008. A regulatable switch mediates self-association in an immunoglobulin fold. *Nat. Struct. Mol. Biol.* **15:** 965-971.

Careri, M., Elviri, L., Zagnoni, I., Cavazzini, D., and Rossi, G.L. 2004. Complexes between recombinant intracellular carriers of vitamin A and their specific ligands. *Eur. J. Mass Spectrom.* **10:** 429-436.

Carrotta, R., Bauer, R., Waninge, R., and Rischel, C. 2001. Conformational characterization of oligomeric intermediates and aggregates in beta-lactoglobulin heat aggregation. *Protein Sci.* **10:** 1312-1318.

Caughey, B., and Lansbury, P.T. 2003. Protofibrils, pores, fibrils, and neurodegeneration: separating the responsible protein aggregates from the innocent bystanders. *Annu. Rev. Neurosci.* **26:** 267-298.

Chen, S., Ferrone, F.A., and Wetzel, R. 2002. Huntington's disease age-of-onset linked to polyglutamine aggregation nucleation. *Proc. Natl. Acad. Sci. USA* **99:** 11884-11889.

Chen, Y., and Barkley, M.D. 1998. Toward understanding tryptophan fluorescence in proteins. *Biochemistry* **37:** 9976-9982.

Chikenji, G., and Kikuchi, M. 2000. What is the role of non-native intermediates of beta-lactoglobulin in protein folding? *Proc. Natl. Acad. Sci. USA* **97:** 14273-14277.

Chiti, F., Calamai, M., Taddei, N., Stefani, M., Ramponi, G., and Dobson, C.M. 2002a. Studies of the aggregation of mutant proteins in vitro provide insights into the genetics of amyloid diseases. *Proc Natl Acad Sci U S A* **99:** 16419-16426.

Chiti, F., De Lorenzi, E., Grossi, S., Mangione, P., Giorgetti, S., Caccialanza, G., Dobson, C.M., Merlini, G., Ramponi, G., and

Bellotti, V. 2001a. A partially structured species of beta 2-microglobulin is significantly populated under physiological conditions and involved in fibrillogenesis. *J. Biol. Chem.* **276:** 46714-46721.

Chiti, F., and Dobson, C.M. 2006. Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* **75:** 333-366.

Chiti, F., and Dobson, C.M. 2009. Amyloid formation by globular proteins under native conditions. *Nat. Chem. Biol.* **5:** 15-22.

Chiti, F., Mangione, P., Andreola, A., Giorgetti, S., Stefani, M., Dobson, C.M., Bellotti, V., and Taddei, N. 2001b. Detection of two partially structured species in the folding process of the amyloidogenic protein β2-microglobulin. *J. Mol. Biol.* **307:** 379-391.

Chiti, F., Stefani, M., Taddei, N., Ramponi, G., and Dobson, C.M. 2003. Rationalization of the effects of mutations on peptide and protein aggregation rates. *Nature* **424:** 805-808.

Chiti, F., Taddei, N., Baroni, F., Cappani, C., Stefani, M., Ramponi, G., and Dobson, C.M. 2002b. Kinetic partitioning of protein folding and aggregation. *Nat. Struct. Biol.* **9:** 137-143.

Chow, M.K., Ellisdon, A.M., Cabrita, L.D., and Bottomley, S.P. 2004a. Polyglutamine expansion in ataxin-3 does not affect protein stability: implications for misfolding and disease. *J. Biol. Chem.* **279:** 47643-47651.

Chow, M.K., Mackay, J.P., Whisstock, J.C., Scanlon, M.J., and Bottomley, S.P. 2004b. Structural and functional analysis of the Josephin domain of the polyglutamine protein ataxin-3. *Biochem. Biophys. Res. Commun.* **322:** 387-394.

Chow, M.K., Paulson, H.L., and Bottomley, S.P. 2004c. Destabilization of a non-pathological variant of ataxin-3 results in fibrillogenesis via a partially folded intermediate: a model for misfolding in polyglutamine disease. *J. Mol. Biol.* **335:** 333-341.

Chowdhury, S.K., Katta, V., and Chait, B.T. 1990. Probing conformational changes in proteins by mass spectrometry. *J. Am. Chem. Soc.* **112:** 9012-9013.

Clark, S.M., and Konermann, L. 2004. Determination of ligand-protein dissociation constants by electrospray mass spectrometry-based diffusion measurements. *Anal. Chem.* **76:** 7077-7083.

Corrigan, A.M., Muller, C., and Krebs, M.R. 2006. The formation of nematic liquid crystal phases by hen lysozyme amyloid fibrils. *J. Am. Chem. Soc.* **128:** 14740-14741.

Cui, D., and Gao, H. 2003. Advance and prospect of bionanomaterials. *Biotechnol. Prog.* **19:** 683-692.

D'Alfonso, L., Collini, M., and Baldini, G. 1999. Evidence of heterogeneous 1-anilinonaphthalene-8-sulfonate binding to beta-lactoglobulin from fluorescence spectroscopy. *Biochim. Biophys. Acta.* **1432:** 194-202.

Das, K., Sarkar, N., Nath, D., and Bhattacharyya, K. 1992. Non-radiative pathways of aniline-naphthalene sulphonates: twisted intramolecular charge transfer versus intersystem crossing. *Spectrochim, Acta* **48A:** 1701-1705.

de la Mora, J.F., Van Berkel, G.J., Enke, C.G., Cole, R.B., Martinez-Sanchez, M., and Fenn, J.B. 2000. Electrochemical processes in electrospray ionization mass spectrometry. *J. Mass Spectrom.* **35:** 939-952.

Declerck, N., Dutartre, H., Receveur, V., Dubois, V., Royer, C., Aymerich, S., and van Tilbeurgh, H. 2001. Dimer stabilization upon activation of the transcriptional antiterminator LicT. *J. Mol. Biol.* **314:** 671-681.

Di Stasio, E., Bizzarri, P., Misiti, F., Pavoni, E., and Brancaccio, A. 2004. A fast and accurate procedure to collect and analyze unfolding fluorescence signal: the case of dystroglycan domains. *Biophys. Chem.* **107:** 197-211.

Dobo, A., and Kaltashov, I.A. 2001. Detection of multiple protein conformational ensembles in solution via deconvolution of

charge-state distributions in ESI MS. *Anal. Chem.* **73:** 4763-4773.

Dole, M., Mack, L.L., Hines, R.L., Mobley, R.C., Ferguson, L.D., and Alice, M.B. 1968. Molecular beams of macroions. *J. Chem. Phys.* **49:** 2240-2249.

Domike, K.R., and Donald, A.M. 2007. Thermal dependence of thermally induced protein spherulite formation and growth: kinetics of beta-lactoglobulin and insulin. *Biomacromolecules* **8:** 3930-3937.

Donald, L.J., Stokell, D.J., Holliday, N.J., Ens, W., Standing, K.G., and Duckworth, H.W. 2005. Multiple equilibria of the Escherichia coli chaperonin GroES revealed by mass spectrometry. *Protein Sci.* **14:** 1375-1379.

Dong, A., Matsuura, J., Manning, M.C., and Carpenter, J.F. 1998. Intermolecular beta-sheet results from trifluoroethanol-induced nonnative alpha-helical structure in beta-sheet predominant proteins: infrared and circular dichroism spectroscopic study. *Arch. Biochem. Biophys.* **355:** 275-281.

Drueke, T.B. 2000. Beta2-microglobulin and amyloidosis. *Nephrol. Dial. Transplant.* **15 (Suppl1):** 17-24.

Dumay, E.M., Kalichevsky, M.T., and Cheftel, J.C. 1998. Characteristics of pressure-induced gels of beta-lactoglobulin at various times after pressure release. *Food Sci. Technol.* **31:** 10-19.

Dumoulin, M., Last, A.M., Desmyter, A., Decanniere, K., Canet, D., Larsson, G., Spencer, A., Archer, D.B., Sasse, J., Muyldermans, S., et al. 2003. A camelid antibody fragment inhibits the formation of amyloid fibrils by human lysozyme. *Nature* **424:** 783-788.

Eakin, C.M., Berman, A.J., and Miranker, A.D. 2006. A native to amyloidogenic transition regulated by a backbone trigger. *Nat. Struct. Biol.* **13:** 202-208.

Eakin, C.M., Knight, J.D., Morgan, C.J., Gelfand, M.A., and Miranker, A.D. 2002. Formation of copper specific binding site in non-

native state of beta-2-microglobulin. *Biochemistry* **41:** 10646-10656.

Eakin, C.M., and Miranker, A.D. 2005. From chance to frequent encounters: origins of beta2-microglobulin fibrillogenesis. *Biochim. Biophys. Acta* **1753:** 92-99.

Eberini, I., Baptista, A.M., Gianazza, E., Fraternali, F., and Beringhelli, T. 2004. Reorganization in apo- and holo-beta-lactoglobulin upon protonation of Glu89: molecular dynamics and pKa calculations. *Proteins* **54:** 744-758.

Eichner, T., and Radford, S.E. 2009. A generic mechanism of beta2-microglobulin amyloid assembly at neutral pH involving a specific proline switch. *J. Mol. Biol.* **386:** 1312-1326.

Ellisdon, A.M., Pearce, M.C., and Bottomley, S.P. 2007. Mechanisms of ataxin-3 misfolding and fibril formation: kinetic analysis of a disease-associated polyglutamine protein. *J. Mol. Biol.* **368:** 595-605.

Ellisdon, A.M., Thomas, B., and Bottomley, S.P. 2006. The two-stage pathway of ataxin-3 fibrillogenesis involves a polyglutamine-independent step. *J. Biol. Chem.* **281:** 16888-16896.

Esposito, G., Michelutti, R., Verdone, G., Viglino, P., Hernandez, H., Robinson, C.V., Amoresano, A., Dal Piaz, F., Monti, M., Pucci, P., et al. 2000. Removal of the N-terminal hexapeptide from human beta2-microglobulin facilitates protein aggregation and fibrils formation. *protein Sci.* **9:** 831-845.

Esposito, G., Ricagno, S., Corazza, A., Rennella, E., Gumral, D., Mimmi, M.C., Betto, E., Pucillo, C.E., Fogolari, F., Viglino, P., et al. 2008. The controlling roles of Trp60 and Trp95 in beta2-microglobulin function, folding and amyloid aggregation properties. *J. Mol. Biol.* **378:** 887-897.

Ferguson, N., Becker, J., Tidow, H., Tremmel, S., Sharpe, T.D., Krause, G., Flinders, J., Petrovich, M., Berriman, J., Oschkinat, H., et al. 2006. General structural motifs of amyloid protofilaments. *Proc. Natl. Acad. Sci. USA* **103:** 16248-16253.

Fink, A.L., Calciano, L.J., Goto, Y., Kurotsu, T., and Palleros, D.R. 1994. Classification of acid denaturation of proteins: intermediates and unfolded states. *Biochemistry* **33:** 12504-12511.

Fligge, T.A., Bruns, K., and Przybylski, M. 1998. Analytical development of electrospray and nanoelectrospray mass spectrometry in combination with liquid chromatography for the characterization of proteins. *J. Chromatogr. B Biomed. Sci. Appl.* **706:** 91-100.

Floege, J., and Ehlerding, G. 1996. Beta-2-microglobulin-associated amyloidosis. *Nephron.* **72:** 9-26.

Fogolari, F., Ragona, L., Zetta, L., Romagnoli, S., De Kruif, K.G., and Molinari, H. 1998. Monomeric bovine beta-lactoglobulin adopts a beta-barrel fold at pH 2. *FEBS Lett.* **436:** 149-154.

Gazit, E. 2007a. Self-assembled peptide nanostructures: the design of molecular building blocks and their technological utilization. *Chem. Soc. Rev.* **36:** 1263-1269.

Gazit, E. 2007b. Use of biomolecular templates for the fabrication of metal nanowires. *FEBS J.* **274:** 317-322.

Gejyo, F., Yamada, T., Odani, S., Nakagawa, Y., Arakawa, M., Kunitomo, T., Kataoka, H., Suzuki, M., Hirasawa, Y., Shirahama, T., et al. 1985. A new form of amyloid protein associated with chronic hemodialysis was identified as beta 2-microglobulin. *Biochem. Biophys. Res. Commun.* **129:** 701-706.

Giurleo, J.T., He, X., and Talaga, D.S. 2008. Beta-lactoglobulin assembles into amyloid through sequential aggregated intermediates. *J. Mol. Biol.* **381:** 1332-1348.

Gordiyenko, Y., and Robinson, C.V. 2008. The emerging role of MS in structure elucidation of protein-nucleic acid complexes. *Biochem. Soc. Trans.* **36:** 723-731.

Gosal, W.S., Clark, A.H., Pudney, P.D.A., and Ross-Murphy, S.B. 2002. Novel amyloid fibrillar networks derived from a

globular protein: beta-lactoglobulin. *Langmuir* **18:** 7174-7181.

Grandori, R. 2003a. Electrospray-ionization mass spectrometry for protein conformational studies. *Curr. Org. Chem.* **7:** 1589-1603.

Grandori, R. 2003b. Origin of the conformation dependence of protein charge-state distributions in electrospray-ionization mass spectrometry. *J. Mass Spectrom.* **38:** 11-15.

Grandori, R., Santambrogio, C., Brocca, S., Invernizzi, G., and Lotti, M. 2009. Electrospray-ionization mass spectrometry as a tool for fast screening of protein structural properties. *Biotechnol. J.* **4:** 73-87.

Greenfield, N.J. 1999. Applications of circular dichroism in protein and peptide analysis. *Trends Analyt. Chem.* **18:** 236-244.

Greenfield, N.J. 2004. Circular dichroism analysis for protein-protein interactions. *Methods Mol. Biol.* **261:** 55-78.

Griffin, J.H., Rosenbusch, J.P., Weber, K.K., and Blout, E.R. 1972. Conformational changes in aspartate trancarbamylase. I. Studies of ligand binding and of subunit interactions by circular dichroism spectroscopy. *J. Biol. Chem.* **247:** 6482-6490.

Groenning, M., Olsen, L., van der Weet, M., Flink, J.M., Frokjaer, S., and Jorgensen, F.S. 2007. Study of the binding of Thioflavin T to beta-sheet-rich and non-beta-sheet-rich cavities. *Struct. Biol.* **158:** 358-369.

Gusella, J.F., and MacDonald, M.E. 2000. Molecular genetics: unmasking polyglutamine triggers in neurodegenerative disease. *Nat. Rev. Neurosci.* **1**.

Haacke, A., Broadley, S.A., Boteva, R., Tzvetkov, N., Hartl, F.U., and Breuer, P. 2006. Proteolytic cleavage of polyglutamine-expanded ataxin-3 is critical for aggregation and sequestration of non-expanded ataxin-3. *Hum. Mol. Genet.* **15:** 555-568.

Haacke, A., Hartl, F.U., and Breuer, P. 2007. Calpain inhibition is sufficient to suppress aggregation of polyglutamine-expanded ataxin-3. *J. Biol. Chem.* **282:** 18851-18856.

Hale, J.E., Butler, J.P., Gelfanova, V., You, J.S., and Knierman, M.D. 2004. A simplifed procedure for the reduction and alkylation of cysteine residues in proteins prior to proteolytic digestion and mass spectral analysis. *Anal. Biochem.* **333:** 174-181.

Hamada, D., and Dobson, C.M. 2002. A kinetic study of beta-lactoglobulin amyloid fibril formation promoted by urea. *Protein Sci.* **11:** 2417-2426.

Hamada, D., Tanaka, T., Tartaglia, G.G., Pawar, A., Vendruscolo, M., Kawamura, M., Tamura, A., Tanaka, N., and Dobson, C.M. 2009. Competition between folding, native-state dimerisation and amyloid aggregation in beta-lactoglobulin. *J. Mol. Biol.* **386:** 878-890.

Hamada, D., Yanagihara, I., and Tsumoto, K. 2004. Engineering amyloidogenicity towards the development of nanofibrillar materials. *Trends Biotechnol.* **22:** 93-97.

Hamdan, M., Curcuruto, O., Molinari, H., Zetta, L., and Ragona, L. 1996. Electrospray mass spectrometry: Complexation between 1-anilinonaphtalene-8-sulphonate and proteins. *J. Mass Spectrom.* **31:** 1261-1264.

Hannemann, F., Bera, A.K., Fischer, B., Lisurek, M., Teuchner, K., and Bernhardt, R. 2002. Unfolding and conformational studies on bovine adrenodoxin probed by engineered intrinsic tryptophan fluorescence. *Biochemistry* **41:** 11008-11016.

Hawe, A., Sutter, M., and Jiskoot, W. 2008. Extrinsic fluorescent dyes as tools for protein characterization. *Pharm. Res.* **25:** 1487-1499.

Heck, A.J., and Van Den Heuvel, R.H. 2004. Investigation of intact protein complexes by mass spectrometry. *Mass Spectrom. Rev.* **23:** 368-389.

Hong, D.P., Gozu, M., Hasegawa, K., Naiki, H., and Goto, Y. 2002. Conformation of beta2-microglobulin amyloid fibrils analyzed by reduction of the disulphide bond. *J. Biol. Chem.* **277:** 21554-21560.

Hughson, F.M., Wright, P.E., and Baldwin, R.L. 1990. Structural characterization of a partly folded apomyoglobin intermediate. *Science* **249:** 1544-1548.

Ilag, L.L., Videler, H., McKay, A.R., Sobott, F., Fucini, P., Nierhaus, K.H., and Robinson, C.V. 2005. Heptameric (L12)6/L10 rather than canonical pentameric complexes are found by tandem MS of intact ribosomes from thermophilic bacteria. *Proc. Natl. Acad. Sci. USA* **102:** 8192-8197.

Invernizzi, G., Šamalikova, M., Brocca, S., Lotti, M., Molinari, H., and Grandori, R. 2006. Comparison of bovine and porcine beta-lactoglobulin: a mass spectrometric analysis. *J. Mass Spectrom.* **41:** 717-727.

Invernizzi, G., Annoni, E., Natalello, A., Doglia, S.M., and Lotti, M. 2008. In vivo aggregation of bovine beta-lactoglobulin is affected by Cys at position 121. *Protein Expr. Purif.* **62:** 111-115.

Invernizzi, G., and Grandori, R. 2007. Detection of the equilibrium folding intermediate of beta-lactoglobulin in the presence of trifluoroethanol by mass spectrometry. *Rapid Commun. Mass Spectrom.* **21:** 1049-1052.

Invernizzi, G., Natalello, A., Šamalikova, M., and Grandori, R. 2007. Protein-protein and protein-ligand interactions studied by electrospray-ionization mass spectrometry. *Protein Pept. Lett.* **14:** 894-902.

Iribarne, J.V., and Thomson, B.A. 1976. On the evaporation of small ions from charged droplets. *J. Chem. Phys.* **64:** 2287-2294.

Kad, N.M., Myers, S.L., Smith, D.P., Smith, D.A., Radford, S.E., and Thompson, N.H. 2003. Hierarchical assembly of beta2-microglobulin amyloid in vitro revealed by atomic force microscopy. *J. Mol. Biol.* **330:** 785-797.

Kaltashov, I.A., and Abzalimov, R.R. 2008. Do Ionic Charges in ESI MS Provide Useful Information on Macromolecular Structure? *J. Am. Soc. Mass Spectrom.* **19:** 1239-1246.

Kaltashov, I.A., and Eyles, S.J. 2002. Studies of biomolecular conformations and conformational dynamics by mass spectrometry. *Mass Spectrom. Rev.* **21:** 37-71.

Kaltashov, I.A., and Mohimen, A. 2005. Estimates of protein surface areas in solution by electrospray ionization mass spectrometry. *Anal. Chem.* **77:** 5370-5379.

Kaltashov, I.A., Zhang, M., Eyles, S.J., and Abzalimov, R.R. 2006. Investigation of structure, dynamics and function of metalloproteins with electrospray ionization mass spectrometry. *Anal. Bioanal. Chem.* **386:** 472-481.

Kameda, A., Hoshino, M., Higurashi, T., Takahashi, S., Naiki, H., and Goto, Y. 2005. Nuclear magnetic resonance characterization of the refolding intermediate of beta2-microglobulin trapped by non-native prolyl peptide bond. *J. Mol. Biol.* **348:** 383-397.

Katsuta, K., Hatakeyama, M., and Hiraki, J. 1997. Isothermal gelation of proteins. 1. Urea-induced gelation of whey proteins and their gelling mechanism. *Food Hydrocolloids* **11:** 367-372.

Katta, V., and Chait, B.T. 1991. Observation of the heme-globin complex in native myoglobin by electrospray-ionization mass spectrometry. *J. Am. Chem. Soc.* **113:** 8534-8535.

Kawaguchi, Y., Okamoto, T., Taniwaki, M., Aizawa, M., Inoue, M., and Katayama, S. 1994. CAG expansions in a novel gene for Machado-Joseph disease at chromosome 14q32.1. *Nature Genet.* **8:** 221-228.

Kebarle, P., and Verkerk, U.H. 2009. Electrospray: from ions in solution to ions in the gas phase, what we know now. *Mass Spectrom. Rev.* **28:** 898-917.

Kelly, M.J., and Reithel, F.J. 1971. A thermodynamic analysis of the monomer-dimer association of -lactoglobulin A at the isoelectric point. *Biochemistry* **10:** 2639-2644.

Kelly, S.M., Jess, T.J., and Prince, N.C. 2005. How to study proteins by circular dichroism. *Biochim Biophys Acta* **1751:** 119-139.

Kelly, S.M., and Price, N.C. 2000. The use of circular dichroism in the investigation of protein structure and function. *Curr. Protein Pept. Sci.* **1:** 349-384.

Kenney, J.M., Knight, D., Wise, M.J., and Vollrath, F. 2002. Amyloidogenic nature of spider silk. *Eur. J. Biochem.* **269:** 4159-4164.

Khurana, R., Coleman, C., Ionescu-Zanetti, C., Carter, S.A., Krishna, V., Grover, R.K., Roy, R., and Singh, S. 2005. Mechanism of thioflavin T binding to amyloid fibrils. *J. Struct. Biol.* **151:** 229-238.

Kodali, R., and Wetzel, R. 2007. Polymorphism in the intermediates and products of amyloid assembly. *Curr. Opin. Struct. Biol.* **17:** 48-57.

Konermann, L., and Douglas, D.J. 1997. Acid-induced unfolding of cytochrome c at different methanol concentrations: electrospray ionization mass spectrometry specifically monitors changes in the tertiary structure. *Biochemistry* **36:** 12296-12302.

Konermann, L., and Douglas, D.J. 1998. Equilibrium unfolding of proteins monitored by electrospray ionization mass spectrometry: distinguishing two-state from multi-state transitions. *Rapid Commun. Mass Spectrom.* **12:** 435-442.

Kontopidis, G., Holt, C., and Sawyer, L. 2004. Beta-lactoglobulin: binding properties, structure, and function. *J. Dairy Sci.* **87:** 785-796.

Krebs, M.R., Bromley, E.H., and Donald, A.M. 2005. The binding of thioflavin-T to amyloid fibrils: localisation and implications. *J. Struct. Biol.* **149:** 30-37.

Krebs, M.R., Devlin, G.L., and Donald, A.M. 2009. Amyloid fibril-like structure underlies the aggregate structure across the pH range for beta-lactoglobulin. *Biophys. J.* **96:** 5013-5019.

Krebs, M.R., Morozova-Roche, L.A., Daniel, K., Robinson, C.V., and Dobson, C.M. 2004. Observation of sequence specificity in the seeding of protein amyloid fibrils. *Protein Sci.* **13:** 1933-1938.

Krell, T., Horsburgh, M.J., Cooper, A., Kelly, S.M., and Coggins, J.R. 1996. Localization of the active site of type II dehydroquinases. Identification of a common arginine-containing motif in the two classes of dehydroquinases. *J. Biol. Chem.* **271:** 24492-24497.

Krittanai, C., and Johnson, W.C.J. 1997. Correcting the circular dichroism spectra of peptides for contributions of absorbing side chains. *Anal. Biochem.* **253:** 57-64.

Kumar, S., and Udgaonkar, J.B. 2009. Conformational conversion may precede or follow aggregate elongation on alternative pathways of amyloid protofibril formation. *J. Mol. Biol.* **385:** 1266-1276.

Kuprowski, M.C., and Konermann, L. 2007. Signal response of coexisting protein conformers in electrospray mass spectrometry. *Anal. Chem.* **79:** 2499-2506.

Kuwata, K., Shastry, R., Cheng, H., Hoshino, M., Batt, C.A., Goto, Y., and Roder, H. 2001. Structural and kinetic characterization of early folding events in beta-lactoglobulin. *Nat. Struct. Biol.* **8:** 151-155.

Lakowicz, J.R. 1987. *Principles of fluorescence spectroscopy*. Plenum Press, New York and London.

Li, F., Macfarlan, T., Pittman, R.N., and Chakravarti, D. 2002. Ataxin-3 is a histone-binding protein with two independent transcriptional corepressor activities. *J. Biol. Chem.* **277:** 45004-45012.

Liu, Y., Gotte, G., Libonati, M., and Eisenberg, D. 2001. A domain-swapped RNase A dimer with implications for amyloid formation. *Nat. Struct. Biol.* **8:** 211-214.

Makin, O.S., and Serpell, L.C. 2005. Structures for amyloid fibrils. *FEBS J.* **272:** 5950-5961.

Mann, M., and Wilm, M. 1995. Electrospray mass spectrometry for protein characterization. *Trends Biochem. Sci.* **20:** 219-224.

Marchese, R., Grandori, R., Carloni, P., and Raugei, S. 2010. On the zwitterionic nature of gas-phase peptides and protein ions. *PLoS Comput. Biol.* **6:** e1000775.

Martin, S.R., and Schilstra, M.J. 2008. Circular dichroism and its application to the study of biomolecules. *Methods Cell. Biol.* **84:** 263-293.

Martins, P.A., Gomes, F., Vaz, W.L., and Moreno, M.J. 2008. Binding of phospholipids to beta-Lactoglobulin and their transfer to lipid bilayers. *Biochim. Biophys. Acta.* **1778:** 1308-1315.

Masino, L., Musi, V., Menon, R.P., Fusi, P., Kelly, G., Frenkiel, T.A., Trottier, Y., and Pastore, A. 2003. Domain architecture of the polyglutamine protein ataxin-3: a globular domain followed by a flexible tail. *FEBS Lett.* **549:** 21-25.

Masino, L., Nicastro, G., Menon, R.P., Dal Piaz, F., Calder, L., and Pastore, A. 2004. Characterization of the structure and the amyloidogenic properties of the Josephin domain of the polyglutamine-containing protein ataxin-3. *J. Mol. Biol.* **344:** 1021-1035.

Mazzucchelli, S., De Palma, A., Riva, M., D'Urzo, A., Pozzi, C., Pastori, V., Comelli, F., Fusi, P., Vanoni, M., Tortora, P., et al. 2009. Proteomic and biochemical analyses unveil tight interaction of ataxin-3 with tubulin. *Int. J. Biochem. Cell Biol.* **41:** 2485-2492.

McParland, V.J., Kad, N.M., Kalverda, A.P., Brown, A., Kinwin-Jones, P., Hunter, M.G., Sunde, M., and Radford, S.E. 2000. Partially unfolded states of beta(2)-microglobulin and amyloid formation in vitro. *Biochemistry* **39:** 8735-8746.

Menaa, C., Esser, E., and Sprague, S.M. 2008. Beta2-microglobulin stimulates osteoclast formation. *Kidney Int.* **73:** 1275-1281.

Mendes, M.A., de Souza, B.M., dos Santos, L.D., Santos, K.S., and Palma, M.S. 2005. Analyzing glycerol-mediated protein oligomerization by electrospray ionization mass

spectrometry. *Rapid Commun. Mass Spectrom.* **19:** 2636-2642.

Mendoza, V.L., Antwi, K., Baron-Rodriguez, M.A., Blanco, C., and Vachet, R.W. 2010. Structure of the preamyloid dimer of beta-2 microglobulin from covalent labeling and mass spectrometry. *Biochemistry* **49:** 1522-1532.

Mirza, U.A., and Chait, B.T. 1994. Effects of anions on the positive ion electrospray ionization mass spectra of peptides and proteins. *Anal. Chem.* **66:** 2898-2904.

Monsellier, E., Ramazzotti, M., Taddei, N., and Chiti, F. 2008. Aggregation propensity of the human proteome. *PLoS Comput. Biol.* **4:** e1000199.

Morgan, C.J., Gelfand, M., Atreya, C., and Miranker, A.D. 2001. Kidney dialysis-associated amyloidosis: a molecular role for copper in fiber formation. *J. Mol. Biol.* **309:** 339-345.

Nagai, Y., Tucker, T., Ren, H., Kenan, D.J., Henderson, D.S., Keene, J.D., Strittmatter, W.J., and Burke, J.R. 2000. Inhibition of polyglutamine protein aggregation and cell death by novel peptides identified by phage display screening. *J. Biol. Chem.* **275:** 10437-10442.

Natalello, A., Doglia, S.M., Carey, J., and Grandori, R. 2007. Role of flavin mononucleotide in the thermostability and oligomerization of Escherichia coli stress-defense protein WrbA. *Biochemistry* **46:** 543-553.

Nerelius, C., Fitzen, M., and Johansson, J. 2010. Amino acid sequence determinants and molecular chaperones in amyloid fibril formation. *Biochem. Biophys. Res. Commun.* **396:** 2-6.

Nicastro, G., Habeck, M., Masino, L., Svergun, D.I., and Pastore, A. 2006. Structure validation of the Josephin domain of ataxin-3: Conclusive evidence for an open conformation. *J. Biomol. NMR* **36:** 267-277.

Nicastro, G., Masino, L., Esposito, V., Menon, R.P., De Simone, A., Fraternali, F., and Pastore, A. 2009. Josephin domain of

ataxin-3 contains two distinct ubiquitin-binding sites. *Biopolymers* **91:** 1203-1214.

Nicastro, G., Menon, R.P., Masino, L., Knowles, P.P., McDonald, N.Q., and Pastore, A. 2005. The solution structure of the Josephin domain of ataxin-3: structural determinants for molecular recognition. *Proc. Natl. Acad. Sci. USA* **102:** 10493-10498.

Okon, M., Bray, P., and Vucelic, D. 1992. 1H NMR assignments and secondary structure of human beta 2-microglobulin in solution. *Biochemistry* **31:** 8906-8915.

Osz, K., Nagy, Z., Pappalardo, G., Di Natale, G., Sanna, D., Micera, G., Rizzarelli, E., and Sóvágó, I. 2007. Copper(II) interaction with prion peptide fragments encompassing histidine residues within and outside the octarepeat domain: speciation, stability constants and binding details. *Chemistry* **13:** 7129-7143.

Pedersen, J.S., Dikov, D., Flink, J.L., Hjuler, H.A., Christiansen, G., and Otzen, D.E. 2006. The changing face of glucagon fibrillation: structural polymorphism and conformational imprinting. *J. Mol. Biol.* **355:** 501-523.

Perutz, M.F., Johnson, T., Suzuki, M., and Finch, J.T. 1994. Glutamine repeats as polar zippers: their possible role in inherited neurodegenerative diseases. *Proc. Natl. Acad. Sci. USA* **91:** 5355-5358.

Peschke, M., Verkerk, U.H., and Kebarle, P. 2004. Features of the ESI mechanism that affect the observation of multiply charged noncovalent protein complexes and the determination of the association constant by the titration method. *J. Am. Soc. Mass Spectrom.* **15:** 1424-1434.

Pilkington, S.M., Roberts, S.J., Meade, S.J., and Gerrard, J.A. 2010. Amyloid fibrils as a nanoscaffold for enzyme immobilization. *Biotechnol. Prog.* **26:** 93-100.

Plakoutsi, G., Bemporad, F., Calamai, M., Taddei, N., Dobson, C.M., and Chiti, F. 2005. Evidence for a mechanism of amyloid

formation involving molecular reorganisation within native-like precursor aggregates. *J. Mol. Biol.* **351:** 910-922.

Platt, G.W., McParland, V.J., Kalverda, A.P., Homans, S.W., and Radford, S.E. 2005. Dynamics in the unfolded state of beta2-microglobulin studied by NMR. *J. Mol. Biol.* **346:** 279-294.

Platt, G.W., and Radford, S.E. 2009. Glimpses of the molecular mechanisms of beta2-microglobulin fibril formation in vitro: aggregation on a complex energy landscape. *FEBS Lett.* **583:** 2623-2629.

Price, N.C., and Stevens, E. 1983. The denaturation of rabbit muscle phosphorylase b by guanidinium chloride. *Biochem J.* **213:** 595-602.

Ptitsyn, O.B. 1995. Molten globule and protein folding. *Adv. Protein Chem.* **47:** 83-229.

Qin, B.Y., Bewley, M.C., Creamer, L.K., Baker, H.M., Baker, E.N., and Jameson, G.B. 1998. Structural basis of the Tanford transition of bovine beta-lactoglobulin. *Biochemistry* **37:** 14014-14023.

Radford, S.E., Gosal, W.S., and Platt, G.W. 2005. Towards an understanding of the structural molecular mechanism of beta(2)-microglobulin amyloid formation in vitro. *Biochim Biophys Acta* **1753:** 51-63.

Rahimi, F., Shanmugam, A., and Bitan, G. 2008. Structure-function relationships of pre-fibrillar protein assemblies in Alzheimer'sdisease and related disorders. *Curr. Alzheimer Res.* **5:** 319-341.

Rasmussen, P., Barbiroli, A., Bonomi, F., Faoro, F., Ferranti, P., Iriti, M., Picariello, G., and Iametti, S. 2007. Formation of structured polymers upon controlled denaturation of beta-lactoglobulin with different chaotropes. *Biopolymers* **86:** 57-72.

Renard, D., Lefebvre, J., Robert, P., Llamas, G., Dufour, E., and Dufour, E. 1999. Structural investigation of beta-

lactoglobulin gelation in ethanol/water solutions. *Int. J. Biol. Macromol.* **26:** 35-44.

Rennella, E., Corazza, A., Giorgetti, S., Fogolari, F., Viglino, P., Porcari, R., Verga, L., Stoppini, M., Bellotti, V., and Esposito, G. 2010. Folding and fibrillogenesis: clues from beta2-microglobulin. *J. Mol. Biol.* **401:** 286-297.

Ricagno, S., Colombo, M., de Rosa, M., Sangiovanni, E., Giorgetti, S., Raimondi, S., Bellotti, V., and Bolognesi, M. 2008. DE loop mutations affect beta-2 microglobulin stability and amyloid aggregation. *Biochemical and biophysical research communications* **377:** 146-150.

Ricagno, S., Raimondi, S., Giorgetti, S., Bellotti, V., and Bolognesi, M. 2009. Human beta-2 microglobulin W60V mutant structure: Implications for stability and amyloid aggregation. *Biochemical and biophysical research communications* **380:** 543-547.

Rosano, C., Zuccotti, S., Mangione, P., Giorgetti, S., Bellotti, V., Pettirossi, F., Corazza, A., Viglino, P., Esposito, G., and Bolognesi, M. 2004. beta2-microglobulin H31Y variant 3D structure highlights the protein natural propensity towards intermolecular aggregation. *J. Mol. Biol.* **335:** 1051-1064.

Ross, C.A., Poirier, M.A., Wanker, E.E., and Amzel, M. 2003. Polyglutamine fibrillogenesis: the pathway unfolds. *Proc. Natl. Acad. Sci. USA* **100:** 1-3.

Roumestand, C., Boyer, M., Guignard, L., Barthe, P., and Royer, C.A. 2001. Characterization of the folding and unfolding reactions of a small beta-barrel protein of novel topology, the MTCP1 oncogene product P13. *J. Mol. Biol.* **312:** 247-259.

Royer, C.A. 2006. Probing protein folding and conformational transitions with fluorescence. *Chem. Rev.* **106:** 1769-1784.

Royer, C.A., Hinck, A.P., Loh, S.N., Prehoda, K.E., Peng, X., Jonas, J., and Markley, J.L. 1993a. Effects of amino acid substitutions on the pressure denaturation of staphylococcal nuclease as

monitored by fluorescence and nuclear magnetic resonance spectroscopy. *Biochemistry* **32:** 5222-5232.

Royer, C.A., Mann, C.J., and Matthews, C.R. 1993b. Resolution of the fluorescence equilibrium unfolding profile of trp aporepressor using single tryptophan mutants. *Protein Sci.* **2:** 1844-1852.

Saiki, M., Honda, S., Kawasaki, K., Zhou, D., Kaito, A., Konakahara, T., and Morii, H. 2005. Higher-order molecular packing in amyloid-like fibrils constructed with linear arrangements of hydrophobic and hydrogen-bonding side-chains. *J. Mol. Biol.* **348:** 983-998.

Sakai, K., Sakurai, K., Sakai, M., Hoshino, M., and Goto, Y. 2000. Conformation and stability of thiol-modified bovine beta-lactoglobulin. *Protein Sci.* **9:** 1719-1729.

Sakata, M., Chatani, E., Kameda, A., Sakurai, K., Naiki, H., and Goto, Y. 2008. Kinetic coupling of folding and prolyl isomerization of beta2-microglobulin studied by mutational analysis. *J. Mol. Biol.* **382:** 1242-1255.

Sakurai, K., and Goto, Y. 2002. Manipulating monomer-dimer equilibrium of bovine Beta -lactoglobulin by amino acid substitution. *J. Biol. Chem.* **277:** 25735-25740.

Sakurai, K., Oobatake, M., and Goto, Y. 2001. Salt-dependent monomer-dimer equilibrium of bovine beta-lactoglobulin at pH 3. *Protein Sci.* **10:** 2325-2335.

Sanglier, S., Atmanene, C., Chevreux, G., and Dorsselaer, A.V. 2008. Nondenaturing mass spectrometry to study noncovalent protein/protein and protein/ligand complexes: technical aspects and application to the determination of binding stoichiometries. *Methods Mol. Biol.* **484:** 217-243.

Santambrogio, C., and Grandori, R. 2008. Monitoring the Tanford transition in beta-lactoglobulin by 8-anilino-1-naphthalene sulfonate and mass spectrometry. *Rapid Commun. Mass Spectrom.* **22:** 4049-4054.

Santambrogio, C., Ricagno, S., Colombo, M., Barbiroli, A., Bonomi, F., Bellotti, V., Bolognesi, M., and Grandori, R. 2010. DE-loop mutations affect beta2 microglobulin stability, oligomerization, and the low-pH unfolded form. *Protein Sci.* **19:** 1386-1394.

Santos, E.C., and Spector, A.A. 1972. Effect of fatty acids on the binding of 1-anilino-8-naphthalenesulfonate to bovine serum albumin. *Biochemistry* **11:** 2299-2302.

Saunders, H.M., and Bottomley, S.P. 2009. Multi-domain misfolding: understanding the aggregation pathway of polyglutamine proteins. *PEDS* **22:** 447-451.

Scalley, M.L., Yi, Q., Gu, H., McCormack, A., Yates, J.I., and Baker, D. 1997. Kinetics of folding of the IgG binding domain of peptostreptococcal protein L. *Biochemistry* **36:** 3373-3382.

Scheibel, T. 2005. Protein fibers as performance proteins: new technologies and applications. *Curr. Opin. Biotechnol.* **16:** 427-433.

Scheibel, T., Parthasarathy, R., Sawicki, G., Lin, X., and Jaeger, H. 2003. Conducting nanowires built by controlled self-assembly of amyloid fibers and selective metal deposition. *PNAS* **100:** 4527-4532.

Schneider, B.B., Douglas, D.J., and Chen, D.D. 2001. Ion fragmentation in an electrospray ionization mass spectrometer interface with different gases. *Rapid Commun. Mass Spectrom.* **15:** 249-257.

Schokker, E.P., Singh, H., Pinder, D.N., and Creamer, L.K. 2000. Heat-induced aggregation of beta-lactoglobulin AB at pH 2.5 as influenced by ionic strength and protein concentration. *Int. Dairy J.* **10:** 233-240.

Schokker, E.P., Singh, H., Pinder, D.N., Norris, G.E., and Creamer, L.K. 1999. Characterization of intermediates formed during heat-induced aggregation of beta-lactoglobulin AB at neutral pH. *Int. Dairy J.* **9:** 791-800.

Seeman, N.C. 2010. Nanomaterials based on DNA. *Annu. Rev. Biochem.* **79**.

Sgourakis, N.G., Patel, N.M., Garcia, A.E., Makhatadze, G.I., and McCallum, S.A. 2010. Conformational dynamics and structural plasticity play critical roles in the ubiquitin recognition of a UIM domain. *J. Mol. Biol.* **396:** 1128-1144.

Shah, D.S., Thomas, M.B., Phillips, S., Cisneros, D.A., Le Brun, A.P., Holt, S.A., and Lakey, J.H. 2007. Self-assembling layers created by membrane proteins on gold. *Biochem. Soc. Trans.* **35:** 522-526.

Shiraki, K., Nishikawa, K., and Goto, Y. 1995. Trifluoroethanol-induced stabilization of the alpha-helical structure of beta-lactoglobulin: implication for non-hierarchical protein folding. *J. Mol. Biol.* **245:** 180-194.

Sideras, K., and Gertz, M.A. 2009. Amyloidosis. *Adv. Clin. Chem.* **47:** 1-44.

Smith, A.M., Jahn, T.R., Ashcroft, A.E., and Radford, S.E. 2006a. Direct observation of oligomeric species formed in the early stages of amyloid fibril formation using electrospray ionisation mass spectrometry. *J. Mol. Biol.* **364:** 9-19.

Smith, D.P., Giles, K., Bateman, R.H., Radford, S.E., and Ashcroft, A.E. 2007. Monitoring copopulated conformational states during protein folding events using electrospray ionization-ion mobility spectrometry-mass spectrometry. *J. Am. Soc. Mass Spectrom.* **18:** 2180-2190.

Smith, D.P., Knapman, T.W., Campuzano, I., Malham, R.W., Berryman, J.T., Radford, S.E., and Ashcroft, A.E. 2009. Deciphering drift time measurements from travelling wave ion mobility spectrometry-mass spectrometry studies. *Eur. J. Mass Spectrom.* **15:** 113-130.

Smith, J.F., Knowles, T.P., Dobson, C.M., Macphee, C.E., and Welland, M.E. 2006b. Characterization of the nanoscale properties of individual amyloid fibrils. *Proc. Natl. Acad. Sci. USA* **103:** 15806-15811.

Sreerama, N., and Woody, R.W. 2000. Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal. Biochem.* **287:** 252-260.

Srikanth, R., Mendoza, V.L., Bridgewater, J.D., Zhang, G., and Vachet, R.W. 2009. Copper binding to beta-2-microglobulin and its pre-amyloid oligomers. *Biochemistry* **48:** 9871-9881.

Stefani, M., and Dobson, C.M. 2003. Protein aggregation and aggregate toxicity: new insights into protein folding, misfolding diseases and biological evolution. *J. Mol. Med.* **81:** 678-699.

Strickland, E.H. 1974. Aromatic contributions to circular dichroism spectra of proteins. *Crit. Rev. Biochem.* **2:** 113-175.

Stryer, L. 1965. The interaction of a naphthalene dye with apomyoglobin and apohemoglobin. A fluorescent probe of non-polar binding sites. *J. Mol. Biol.* **13:** 482-495.

Sun, J., Kitova, E.N., Sun, N., and Klassen, J.S. 2007. Method for identifying nonspecific protein-protein interactions in nanoelectrospray ionization mass spectrometry. *Anal. Chem.* **79:** 8301-8311.

Szabo, A.G., Stepanik, T.M., Wayner, D.M., and Young, N.M. 1983. Conformational heterogeneity of the copper binding site in azurin. A time-resolved fluorescence study. *Biophys. J.* **41:** 233-244.

Tanaka, M., Morishima, I., Akagi, T., Hashikawa, T., and Nukina, N. 2001. Intra- and intermolecular beta-pleated sheet formation in glutamine-repeat inserted myoglobin as a model for polyglutamine diseases. *J. Biol. Chem.* **276:** 45470-45475.

Uversky, V.N. 2002. Natively Unfolded Proteins: A Point Where Biology Waits for Physics. *Protein Sci.* **11:** 739-756.

Uversky, V.N. 2009. Intrinsic disorder in proteins associated with neurodegenerative diseases. *Front. Biosci.* **14:** 5188-5238.

Uversky, V.N., and Fink, A.L. 2004. Conformational constraints for amyloid fibrillation: the importance of being unfolded. *Biochim. Biophys. Acta* **1698:** 131-153.

Verkerk, U.H., and Kebarle, P. 2005. Ion-ion and ion-molecule reactions at the surface of proteins produced by nanospray. Information on the number of acidic residues and control of the number of ionized acidic and basic residues. *J. Am. Chem. Soc.* **16:** 1325-1341.

Voropai, E.S., Samtsov, M.P., Kaplevskii, K.N., Maskevich, A.A., Stepuro, V.I., Povarova, O.I., Kuznetsova, I.M., Turoverov, K.K., Fink, A.L., and Uversky, V.N. 2003. Spectral properties of Thioflavin T and its complexes with amyloid fibrils. *J. Appl. Spectros.* **70:** 868-874.

Wallace, B.A. 2000. Synchrotron radiation circular dichroism spectroscopy as a tool for investigating protein structures. *J. Synchrotron Radiat.* **7:** 289-295.

Wang, W., Kitova, E.N., and Klassen, J.S. 2005. Nonspecific protein-carbohydrate complexes produced by nanoelectrospray ionization. Factors influencing their formation and stability. *Anal. Chem.* **77:** 3060-3071.

Wearsch, P.A., and Cresswell, P. 2008. The quality control of MHC class I peptide loading. *Curr. Opin. Cell Biol.* **20:** 624-631.

Woody, R.W., and Dunker, A.K. 1996. Aromatic and cystine side-chain circular dichroism in proteins. In *Circular dichroism and the conformational analysis of biomolecules*. (ed. G.D. Fasman), pp. 109-157. Plenum Press, New York.

Wu, S.Y., Perez, M.D., Puyol, P., and Sawyer, L. 1999. beta-lactoglobulin binds palmitate within its central cavity. *J. Biol. Chem.* **274:** 170-174.

Yamamoto, S., Hasegawa, K., Yamaguchi, I., Goto, Y., Gejyo, F., and Naiki, H. 2005. Kinetic analysis of the polymerization and depolymerization of beta(2)-microglobulin-related amyloid fibrils in vitro. *Biochim. Biophys. Acta* **1753:** 34-43.

Yamamoto, S., Hasegawa, K., Yamaguchi, I., Tsutsumi, S., Kardos, J., Goto, Y., Gejyo, F., and Naiki, H. 2004a. Low concentrations of sodium dodecyl sulfate induce the extension of beta 2-microglobulin-related amyloid fibrils at neutral pH. *Biochemistry* **43:** 11075-11082.

Yamamoto, S., Yamaguchi, I., Hasegawa, K., Tsutsumi, S., Goto, Y., Gejyo, F., and Naiki, H. 2004b. Glycosaminoglycans enhance the trifluoroethanol-induced extension of beta 2-microglobulin-related amyloid fibrils at a neutral pH. *J. Am. Soc. Nephrol.* **15:** 126-133.

Zasypkin, D.V., Dumay, E., and Cheftel, J.C. 1996. Pressure- and heat-induced gelation of mixed -lactoglobulin/xanthan solutions. *Food Hydrocolloids* **19:** 203-211.

Zoghbi, H.Y., and Orr, H.T. 2000. Glutammine repeats and neurodegeneration. *Annu. Rev. Neurosci.* **23:** 217-247.

ACKNOWLEDGMENTS