Universidad de Granada

Programa de Doctorado en Farmacia

Departamento de Fisicoquímica

TESIS DOCTORAL

STRUCTURAL CHANGES IN PRE-AMYLOIDOGENIC AGGREGATES OF THE SH3 DOMAIN OF α-SPECTRIN

Fabio Castello

GRANADA 2016

Editor: Universidad de Granada. Tesis Doctorales
Autor: Fabio Castello
ISBN: 978-84-9125-999-2
URI: http://hdl.handle.net/10481/44169

El doctorando, Fabio Castello, y los directores de la tesis, Ángel Orte Gutiérrez, Salvador Casares Atienza y María José Ruedas Rama

Garantizamos, al firmar esta tesis doctoral, que el trabajo ha sido realizado por el doctorando bajo la dirección de los directores de la tesis y hasta donde nuestro conocimiento alcanza, en la realización del trabajo, se han respetado los derechos de otros autores a ser citados, cuando se han utilizado sus resultados o publicaciones.

Granada, a 3 de Junio de 2016

Director/es de la Tesis

Fdo.: Ángel Orte Gutiérrez

Fdo.: Salvador Casares Atienza

Fdo.: María José Ruedas Rama

Doctorando

Fdo.: Fabio Castello

Contents

Abbreviations

A488: Atto 488

A647N: Atto647N

BIFL: burst integrated fluorescence lifetime

CD: circular dichroism

DLS: dynamic light scattering

DSC: differential scanning microscopy

DTT: dithiothreitol

***E*:** FRET efficiency

FCS: Fluorescence correlation spectroscopy

FLIM: Fluorescence lifetime imaging measurements

FRET: Forster resonance energy transfer

IRF: instrumental reconstruction function

MFD: Multi-parameter Fluorescence Detection

MW: molecular weight

N47A-SH3: N47A variant of the α-spectrin SH3-domain engineered with a six-residue tag (Gly-Ser-Gly-Ser-Gly-Cys)

N47A-SH3-A488: A488-labeled N47A-SH3

N47A-SH3-A647N: A647N-labeled N47A-SH3

PIE: Pulsed Interleaved excitation

SDS: Sodium dodecyl sulfate

SMF: single-molecule fluorescence Spectroscopy

TCCD: Two color coincidence spectroscopy

TEM: Transmission electron microscopy

τ_{A488}: donor lifetime

1. INTRODUCTION

1. INTRODUCTION

1.1 THE ABILITY OF PROTEINS TO FOLD

The survival and health of all organisms depend on the ability of their constituent proteins to fold into precise three-dimensional conformations in order to develop their biological functions (1-7). However, after over 50 years history of the protein folding problem, there are still some broad issues unresolved, as to explain the mechanism by which disordered polypeptide chains convert into functional *native* proteins on a finite, biologically relevant time, how the fold is encoded in the sequence as well as if it is possible to predict protein structures from their sequence (8-15).

Anfinsen's *in vitro* studies on ribonuclease, in the 1960s and 1970s, demonstrated that proteins are able to reach their correct structure without auxiliary factors (16). He postulated that the native structure of a protein corresponds to the most thermodynamically stable structure, and that this information resides within the amino acids sequence (16). Following Anfinsen's evidences and using computer simulations, Levinthal perceived that if folding occurred by a purely random search of all possible conformations, the time to find the correct structure might take several orders of magnitude longer than the age of the universe (*Levinthal's paradox*) (14, 17-19). So, he supposed that proteins must follow predetermined pathways to fold. From experimental and theoretical simulations of the folding process (20), emerged that more compact, low-energy conformational ensembles may possess fewer conformations (12, 20–23). This, together with the awareness of the inability to equilibrate to a common structure (14, 19, 21), drove a more

statistical *new view* of the folding. This assumes that there must be different pathways that different individual molecules of the same protein sequence can take, to reach the same native structure (21-25). This vision, known as the *theory of the energy landscape* or *funnel folding hypothesis*, involves the concept of the energy landscapes and permits to solve the Levinthal's paradox (5, 21, 23, 24, 26-28). The new view provide a statistical description of the energy landscape in multi-dimensional funnels, which describe the Gibbs energy of the polypeptide chain as a function of its conformational properties (Figure 1.1), and assume multiple pathways leading to the same native state. The thermodynamic stability corresponds to the deepness within the diagram whereas the kinetic stability is related to the heights of the barriers, which separate the Gibbs energy minima (5, 21, 28-30). The process is stochastic, involving an intricate interplay between global topology and local structure: moving to a lower energy level along the diagram, the conformational space gets reduced and consequently, the number of conformations to be explored by the protein is also decreased. Remarkably, as the landscape is encoded by the amino-acid sequence, evolution has enabled protein sequence to fold both efficiently and quickly, specifically selecting for protein sequences competent to function and not simply for being stable and/or able to fold (31). Under the energy landscape theory and according to the thermodynamic laws, the driving force of the process is the Gibbs energy, which depends on the slight but decisive balance of entropy and enthalpy ($\Delta G° = \Delta H - T\Delta S$) (13). The native state possess lower entropy compared to other intermediates, but the establishment of a large number of favorable specific interactions favors a large negative enthalpy, driving the system towards more ordered states (13, 32-35). However, the forces that direct protein folding are not describable in exact terms. They are delicately balanced, interlocking, and their determination results extremely challenging, due to the difficulty of the study of microscopic interactions within the heterogeneous interior of the proteins, but also for the time

scale of the events, ranging from microsecond to second (36, 37). Much information on the forces leading the process comes from over 80,000 proteins contained in protein databases (38-43), including hydrogen bonds, van der Waals interactions, backbone angle preference, electrostatic interactions, hydrophobic interactions, chain entropy, etc. (13, 44, 45). All these are widely applied in computer simulations to study protein equilibria as well as dynamics (46, 47). In addition, the advances in experimental methods, such as mutational studies, hydrogen exchange pulse labeling, nuclear magnetic resonance (NMR), single-molecule methods, etc., have made possible the characterization of these folding intermediates, providing valuable information on the defined or multiple pathways of the folding process (48-52).

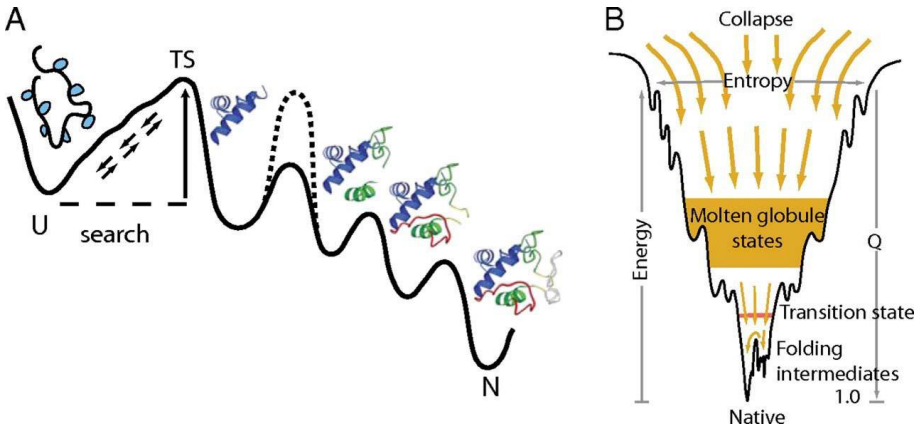


Figure 1.1 A: The classical view of the folding pathway: distinct intermediate states in distinct routes. (U: unfolded, N: native) **B:** The new view: multiple routes through a funneled landscape. In search of conformational space, proteins probe only a small number of all structural options during its transition from a random coil to a native structure, and they must fold energetically downhill and reduce in conformational extent as the native state is approached. Reprinted with permission from Proc Natl Acad Sci U S A, Vol. 111, 15873-15880 (2014).

1.2 FOLDING VERSUS AGGREGATION

Anfinsen's thermodynamic hypothesis (16) comes from in vitro studies and is suitable for diluted solutions and for individual molecules in isolation (8, 30, 53-55). Although the fundamental nature of the folding process is similar *in vitro* and *in vivo* (55), from the new quantitative methods to study proteins in living systems emerged that within the cells, protein folding occurs in a complex and highly populated environment where other macromolecules and metabolites are present (55-58). Therefore, the energy landscape for most proteins in the functional proteome is considerably more complex than in vitro (55, 59-62).

After their synthesis in the ribosome as nascent polypeptide chains, proteins are exposed to a series of potentially dangerous interactions with multiple elements in a heterogeneous environment before reaching their native states. Protein expression, folding, aggregation, regulation and degradation are balanced carefully to guarantee that the whole system functions properly. Specifically, in order to facilitate a correct folding, a complex of different molecules works in close collaboration to proteins, including molecular chaperones, proteolytic enzymes, folding catalysts or autophagic mechanisms (56-58, 63-66). Even with the existence of these protein quality-assurance systems, the folding process is far from perfect: some polypeptides or proteins can undergo an incorrect folding or *misfolding* (8, 67-69), a condition that is associated with several pathological states collectively called *protein misfolding* (*or protein conformational*) *diseases* (8, 10, 70-72). Some of these misfolded proteins or peptides can eventually aggregate into highly-organized fibrillar species, rich in cross-β sheets, which composed insoluble and pathogenic deposits of fibrillar aggregates found as intracellular inclusions or extracellular plaques (amyloids) in a variety of organs or tissues.

This process is associated with a loss of the original protein function and with the generation of toxic intermediates (Figure 1.3) (9-12, 73-78).

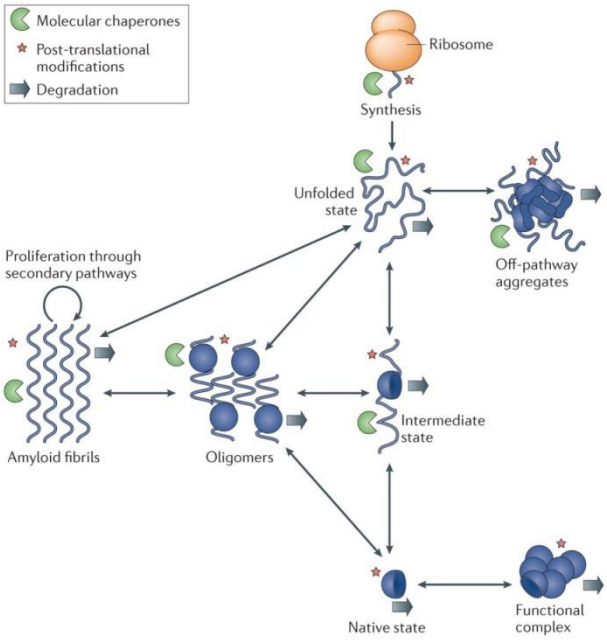


Figure 1.3 A schematic representation of some conformational states which can be adopted by polypeptide chains after their synthesis on the ribosome. Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Molecular Cell Biology, 15, 384–396 (2014).

Recently, *in vitro* studies on peptides and proteins, together with the advances in experimental methods able to detect rare species (79, 80), has provided new information about the complex nature of the multitude of conformational states available to a polypeptide chain. These conformations include those partially folded states normally accumulated during folding which can interconvert to each

other, determining the diversion of molecules towards alternative pathways (10, 81-84). Rather than the classic energy landscape of protein folding, a more complete view of this complex process can be displayed in combined energy landscapes for protein folding and aggregation (Figure 1.3) (55, 85). The species which populate the folding side (unfolded, intermediate and native-like states) are able to overcome the energy barrier to become aggregates via different pathways (85), so that the *aggregate state* can be considered as an alternative to the *folded* state (86).

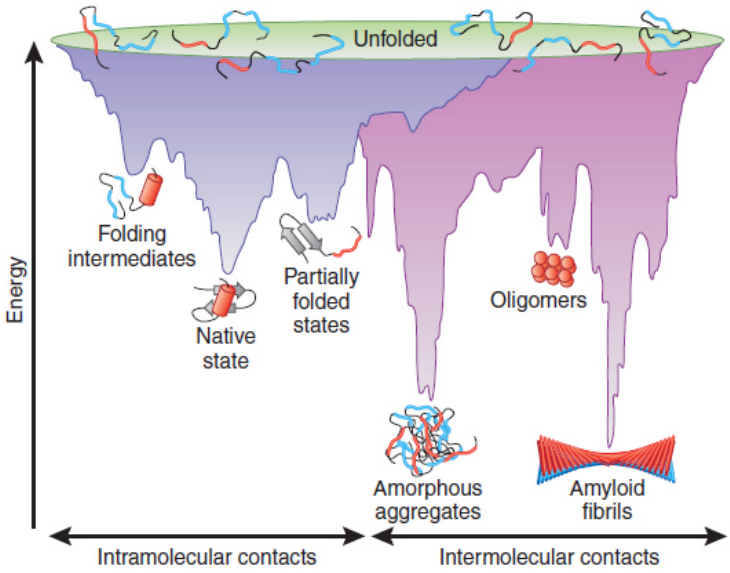


Figure 1.3 A combined energy landscape for protein folding (purple area) and aggregation (pink area) showing the wide range of conformations that a polypeptide chain can assume and the multitude of pathways available as it goes around the landscape. Reprinted by permission from Macmillan Publishers, Nature Structural and Molecular Biology, 16, 574-581 (2009).

1.3 THE IMPACT OF AMYLOID FIBRIL AGGREGATION

The largest group of misfolding diseases originate because misfolding occurs concomitantly with aggregation (70). To date, there are more than 40 known pathologies, some sporadic and/or hereditary and some transmissible, associated with the conversion of soluble and functional peptides and proteins into intractable aggregates, being amyloids the most significant examples (87-94). The term 'amyloidoses' is usually employed to indicate those diseases characterized by the presence of amyloid plaques, whose main constituent is a specific peptide or protein, different in the varying diseases (Figure 1.4).

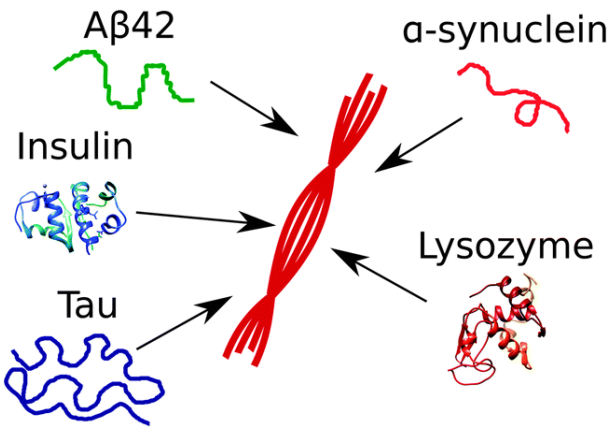


Figure 1.4 Examples of amyloid forming proteins associated to several human diseases. Reprinted from Phys. Chem. Chem. Phys., 2015, 17, 7606-7618 - Published by the PCCP Owner Societies.

Among these pathologies, there are some of the most devastating neurological as well as systemic diseases, such as Alzheimer's, Parkinson's and Huntington's

disease, type 2 diabetes, Transthyretin amyloidosis, Mad Cow disease and many others (Table 1.1) (8, 10, 95, 96).

Table 1.1 List of some misfolded proteins associated with human diseases.

Disease	Aggregating protein or peptide	Structure of protein or peptide
Alzheimer's disease	Aβ peptide	Intrinsically disordered
Parkinson's disease	α-synuclein	Intrinsically disordered
Huntington's disease	Huntingtin fragments	Mostly intrinsically disordered
Amyotrophic lateral sclerosis	Superoxide dismutase 1	β-sheet and Ig-like
Familial amyloid polyneuropathy	Transthyretin mutants	β-sheet
Spongiform encephalopathies	Prion protein or its fragments	Intrinsically disordered and α-helical
Type II Diabetes	Amylin	Intrinsically disordered
Lysozyme amyloidosis	Lysozyme mutants	α-helical and β-sheet

Regardless the nature of the corresponding precursor peptides or proteins related to the amyloidosis considered, the fibrillar deposits associated to diseases share common properties (7, 97). According to Transmission Electron Microscopy (TEM) or Atomic Force Microscopy (AFM) *in vitro* studies, the amyloid aggregates are organized in long and straight fibrils without ramification or associate laterally to form ribbons, (97-99). From circular dichroism (CD), X-ray

diffraction and infrared (IR) spectroscopy studies emerged that these aggregates are organized in each filament in a highly-ordered cross-β structure, forming β-sheets that extend throughout the entire length of the fibrils (100, 101). Nowadays we possess a great level of detail in the internal structure of the fibrils, thanks to the developments in methodologies, as magic angle spinning and solid-state NMR spectroscopy (102-105), growing nano- and micro-crystals of short peptides with amyloid characteristics (11, 106, 107) and amide hydrogen-deuterium studies combined with high-resolution NMR and mass spectrometry (108-111) (Figure 1.5).

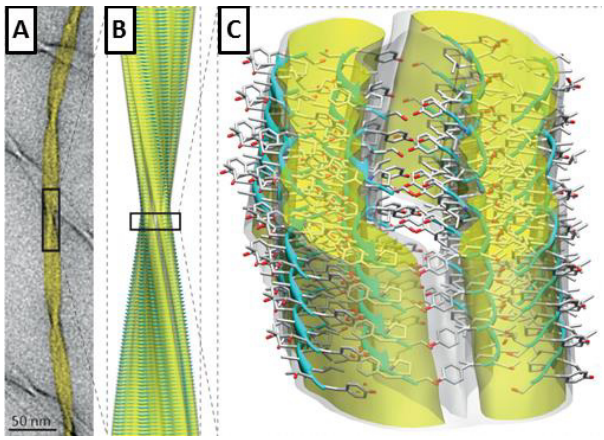


Figure 1.5 Structure of one of the several polymorphs of amyloid fibrils formed by an 11-residue fragment of transthyretin (112). **A:** Background image of the fibril taken using TEM (Scale bar, 50 nm). **B:** Atomic-level structure obtained combining cryo-electron microscopy reconstruction with solid-state NMR. **C:** Close-up view showing the hierarchical organization of an amyloid fibril. The three filaments forming the mature fibril are in turn formed by pairs of cross-β protofilaments, which are composed of pairs of β-sheets each. The fibril surfaces are shown as electron density maps, and the constituent β-sheets are shown in a ribbon representation; oxygen, carbon, and nitrogen atoms are shown in red, gray and blue, respectively. Reprinted with permission from Macmillan Publishers Ltd: Nature Reviews Molecular Cell Biology, 15, 384–396 (2014).

In amyloid disease patients, fibrillar aggregates of specific toxic peptides, proteins, or their fragments, can be found in a variety of organs or tissues (77, 78, 94, 113). They can form in organs including the brain, the liver and the spleen, or in skeletal tissue, depending on the pathology (or *amyloidosis*) they are involved in (114-116).

Many *in vivo* and *in vitro* studies on amyloidogenic proteins support the existence of a direct link between aggregation and the onset of these pathologies (the so-called *amyloid hypothesis*) (2, 117-119), whereas other studies report that not all amyloids are toxic but they can also develop biological functions (120-124). From more recent studies on many amyloidogenic proteins emerged that smaller aggregates and prefibrillar intermediates forming during the amyloid growth might be the main responsible agents of cell toxicity and dysfunction, whereas monomers and full-length fibrils cause limited or no toxicity (74, 76, 77, 125-130). Moreover, it has also been suggested the toxicity of the aggregation process *per se* (131). These recent acquisitions have emphasized the need to investigate the underlying molecular events of to obtain more information on the amyloid process.

1.4 STRUCTURE OF OLIGOMER AND AGGREGATES

Imaging techniques, such as TEM (99) and AFM (132), are commonly used to determine morphologies and sizes of the oligomers and the aggregates, on the nanometer length-scale. While TEM gives information about the sample surface in two-dimensions and involves staining of the sample to enhance the signal, AFM does not require any additional markers and provides three-dimensional nanoscale morphology of aggregates and fibrils. TEM images shows that the structure of the protofibrillar intermediates is similar to the mature fibrils. Protofibrils are thinner (diameters of usually less than 10 nm), shorter (length usually below 400 nm) and

more curvilinear (133, 134). They do not share the same shape or the symmetry of mature fibrils and they can be distinguished from oligomers by their elongated, linear structure (101, 129, 135) (Figure 1.5). On the contrary, oligomers do not present any defined shape as mature fibrils or protofibrils do, even if they are frequently referred to as spherical species (136-139). Compared to fibrils, they are usually considered to be small, however, it has been also documented that they can possess lateral extension reaching diameters of over 50 nm (137, 139-141).

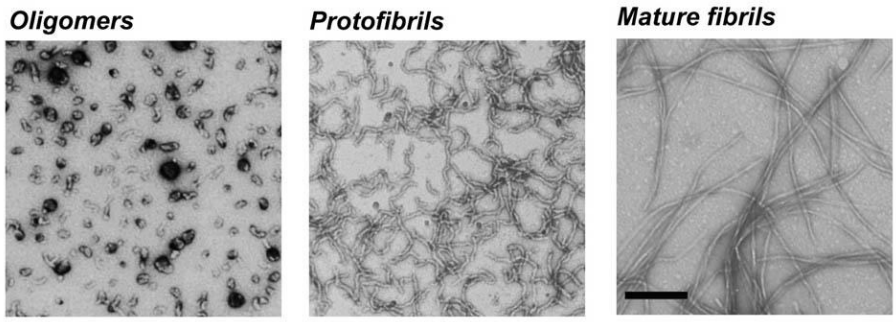


Figure 1.5 Structure of mature Aβ (1-40) amyloid fibrils, protofibrils and oligomers Reprinted from Journal of Molecular Biology**,** Volume 421, 427–440 (2012), with permission from Elsevier.

Other features permit to distinguish oligomers from fibrils, such as the content of cross-β structure typically present in amyloid fibrils (142), whereas oligomers can vary their secondary structure characteristics substantially: some of them have been described to exhibit high β-sheet content (76, 137, 141, 143, 144), while others show random-coil-like conformations (145). This demonstrates the high heterogeneity of oligomers (76, 146-148) and, at the same time, it remarks the inherent difficulty to reduce them to precise molecular identities (129).

1.5 EXPERIMENTAL APPROACHES TO STUDY PROTEIN AGGREGATION

The biophysical techniques typically used to study protein aggregations *in vitro* are briefly described as follows (128):

- far-UV (CD) spectroscopy (149-151), which is routinely used to determine the content of β-sheet and α-helical structure in protein monomers as well as in aggregates. As amyloid fibrils are mostly composed of cross-β sheet-rich structure, CD allows to monitor the conformational changes from any native conformation to a predominantly β-sheet amyloid;

- hydrogen/deuterium amide exchange coupled to NMR (152), which are used to probe which regions of the polypeptide chain are involved in persistent hydrogen-bonded structure or are excluded from solvent, providing the structural details of amyloid fibrils both at the residue-specific as well as at the core levels;

- X-ray diffraction (142, 153), which is used to confirm the presence of amyloid fibrils, providing information on their cross-β structures;

- electrospray ionization mass spectrometry (ESI-MS), which has revealed to be quite informative at various stages of fibril formation as well as to detect temporarily-formed oligomeric species (154, 155);

- fluorescence emission intensity of dyes such as Thioflavine T and Congo Red upon binding to amyloid fibrils, which permits to obtain information about intermolecular interactions and the formation of large proto-fibrils and fibrils;

- dynamic light scattering (DLS), which allows to estimate sizes and to study interactions between oligomers (156, 157);

- analytical ultracentrifugation, which allows separation of biomolecules according to their sedimentation coefficients, providing information about their mass and size (158, 159).

In addition to the reported methods, the analysis of chemical kinetics has become increasingly useful to complement the experimental characterization of the molecular mechanism of aggregation (160).

1.6 KINETICS OF AMYLOID FORMATION

From many *in vitro* and *in vivo* experiments on amyloid fibril formation with different peptides and proteins, it has become evident the difficulty to reduce the kinetics of the aggregation to a single model. Important mechanistic insights for elucidating the amyloid formation inherent mechanism have come from the combination of classical and modern kinetic theories, (161-170) which together provide the link between macroscopic measurements and the microscopic (molecular) steps in the process of aggregation. However, the manner by which the reaction proceeds is still uncertain. This is mainly due to the difficulty to detect and characterize the small, transient and structurally heterogeneous species which participate to the process. Moreover, with these transient species usually not being the dominant ones, they are even more difficult to be isolated. Many experimental techniques have been developed to characterize the kinetics of the aggregation processes and formulate and test the corresponding hypotheses. However, different experiments often report different results, especially when they focus on experimentally describing the early stages of the aggregation process (167, 170).

A clear definition of the agents involved in the aggregation reaction has been proposed by Arosio and colleagues (171):

- *monomers*: single protein or peptide chain;
- *aggregates*: any assembly containing more than one monomer (from dimers to higher order assemblies);
- *oligomers*: small aggregates of different structures, not resulting in fibers, which elongate more slowly with respect to fibrils. However, the distinction between oligomers and fibrils by the aggregate size is difficult, since the size limits are different for different proteins;
- *nuclei*: the smallest aggregate for which a growth by monomer additions would be faster than their dissociation into monomers;
- *amyloid fibrils*: linear aggregates formed by tens of thousands of monomers, with a repetitive cross-β structure.

Many improvements have been inferred thanks to the use of theoretical models, where the strategies used are often based on the comparison of integrated rate laws and experimental kinetic measurements (164, 166, 167, 172-177). The pioneer model for aggregation was developed by Oosawa and colleagues, studying the polymerization of actin. He suggested that the aggregation mechanism could be described by homogeneous (*primary*) nucleation of aggregates, followed by growth through monomer addition (161, 162, 178). The time course that describes this process commonly follows a sigmoidal profile, and consists of three distinct phases. The first one, the *lag phase*, where little or no fibrillar growth can be detected, precedes a *growth phase*, where a large mass percentage of the starting proteins or peptides is converted into fibrils. In the third, final phase, the amyloid accumulation reaches a *plateau* (166, 179-181). However, reducing the kinetics of the aggregation to a single model results difficult due to the multiple model

parameters (known as "force fields") that should be considered in computer simulations of protein aggregation (13, 182-184).

Many studies have revealed that monomers and fibrils are the prevailing species within the process: the *lag phase* is mainly populated by monomers, whereas fibrils dominate at the final *plateau*. The growth phase shows similar concentrations of both species (181, 185, 186). Nuclei are the species which display the highest Gibbs energy compared to monomers and fibrils (Figure 1.8) (166, 171, 187, 188). During each phase of the process, small aggregates or oligomers, may be present at low concentration, and they are often transient and heterogeneous. Consequently, they are challenging to detect and characterize, and therefore, a detailed description of the process at the molecular level is complicated. Moreover, none of the three phases can be endorsed to a single molecular event or microscopic process, rather, a full range of microscopic processes may be active during all phases of the reaction, involving the formation of a multitude of heterogeneous molecular species (75, 128, 164, 166, 176).

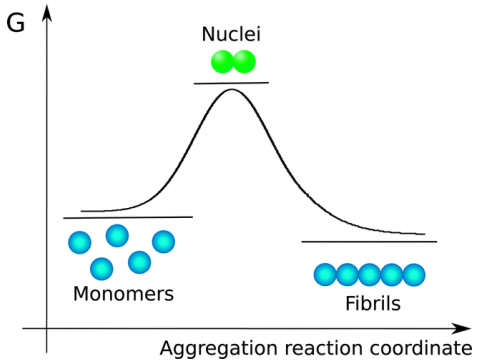


Figure 1.8 A simplistic version of the Gibbs energy diagram of the heterogeneous species which can populate the aggregation reaction. Reprinted from Phys. Chem. Chem. Phys., 2015, 17, 7606-7618 - Published by the PCCP Owner Societies.

Out of these three steps within the amyloid aggregation mechanism, the *lag phase* requires a especial attention since low molecular weight species formed during this step have revealed to be highly toxic (9, 74, 77, 119). Moreover, this *lag phase* is not only a waiting time for the nuclei to appear, but rather involves multiple parallel processes mandatory by nuclei to grow and proliferate, before the end of the lag phase is reached(171). From many studies, it has emerged that the length of the *lag phase* may be influenced by molecular factors, both intrinsic and extrinsic. Intrinsic factors include sequence variations such as mutations, truncations and extensions (189-193), whereas extrinsic include peptides and proteins (194-197), membranes (198), salts (199-202), pH (198, 203), temperature or mechanical factors such as shaking (75, 204, 205). Even subtle modifications on these factors can impact on the duration of the *lag phase* by affecting different microscopic events and, consequently, change the shape of the sigmoidal aggregation growth curve. For example, shaking might produce a decrease of the length of the *lag phase*, by favoring fragmentation processes (75).

More recent mathematical models, where several simplifying assumptions have been considered, have allowed to apply the Oosawa Model to more complex systems, suggesting the possibility of more stages of nucleus assembly, or even multiple intermediate stages of assembly (163-165). While Oosawa theorized that the actin filaments might reduce or grow in size by subtracting or adding a monomer, proposing that aggregation would occur by a primary pathway, such as *homogeneous nucleation* (161, 162), aggregates generate at a rate dependent on the concentration of monomers and independently on the existing fibril extents. From other kinetic studies and computer simulations emerged the existence of secondary pathways that generate new aggregates, at a rate dependent on the concentration of the existing fibrils (75, 167, 206). The aggregation process may consist of a number of consecutive and/or parallel stages, including partial

unfolding or refolding, oligomerization, conformational conversion, elongation, condensation, fragmentation, etc. (188, 207), which can be subdivided into monomer-independent processes, such as fragmentation, and monomer-dependent process, such as secondary nucleation (75) (Figure 1.7).

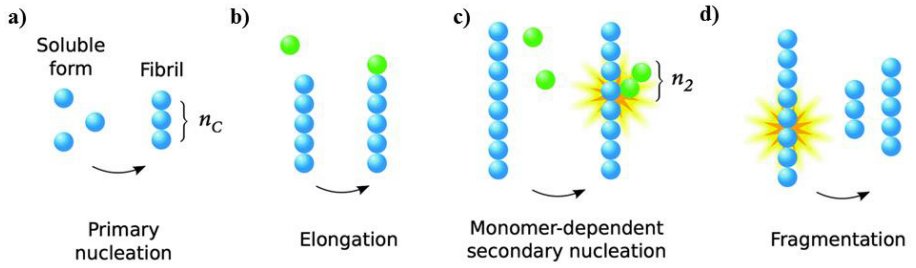


Figure 1.7 Illustration of the possible microscopic mechanisms underlying amyloid formation: primary nucleation from monomers in solution (**a**), elongation (growth) by monomers addition to existing aggregates (**b**), surface catalyzed secondary nucleation from monomers on fibril surface (**c**) and fragmentation (**d**). Readapted from Phys. Chem. Chem. Phys., 2015, 17, 7606-7618 - Published by the PCCP Owner Societies.

Thus, although the nucleation-growth mechanism proposed by Oosawa is still one of the most widely accepted mechanistic models to macroscopically describe the conversion of normally soluble peptides and proteins into amyloids fibrils, depending on the protein and on the experimental conditions, there will be predominance of one or more of these stages in the global aggregation kinetics, which will result in different kinetic behaviors (Figure 1.8) (162, 166, 185, 208-210).

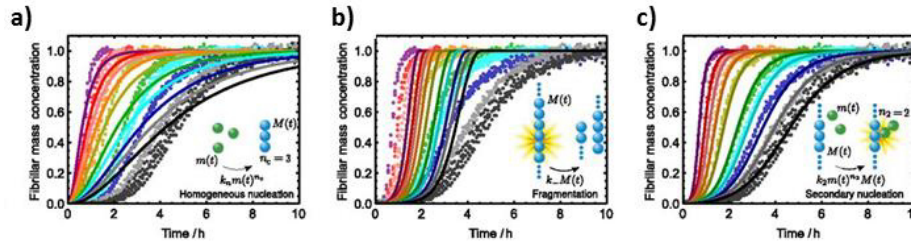


Figure 1.8 Examples of experimental kinetic profiles of Aβ42 aggregation at different initial monomer concentrations. **a:** the dominant nucleation mechanism occurs via primary nucleation, and there are no secondary pathways (161, 168), **b:** the dominant process is fragmentation, in addition to primary nucleation (167), **c:** secondary nucleation, in addition to primary nucleation. Readapted from Proc Natl Acad Sci U S A, Vol.110 9758–9763 (2013).

Moreover, from the experimental and theoretical literature has emerged other possible scenarios of nucleation: it has also been proposed that the amyloid nucleation process should take place in two phases as well, rather than in a simple one (186, 187, 211-216). One-step nucleation involves that monomeric proteins polymerize directly into fibrils, whereas in two-step nucleation soluble monomers first assemble into disordered oligomers, which subsequently convert into a β-sheet nucleus. Some experiments have revealed disordered oligomers ranging in size between dimers and micrometer-sized particles, but their role in the process of fibril formation is still undetermined (217-225). Since then strong evidences suggest that disorder oligomers may be the main cytotoxic agents in protein aggregation diseases (12, 73, 77), it has been becoming more and more important to understand their role during amyloid formation to develop intervention strategies that target these species.

1.7 SINGLE-MOLECULE TECHNIQUES

The importance of characterizing the early intermediates has significantly been increasing since the oligomers have been suggested to be more cytotoxic than the fibrils (12, 73, 77). Due to their dynamic and heterogeneous nature, the characterization of oligomeric intermediates and microscopic pathways of the aggregation process still remains elusive. Conventional methods are useful to characterize monomers and fibrils (100, 128, 226), but key states may be hidden within the ensemble, especially if they are only populated rarely or briefly.

In the last two decades, single-molecule (SM) methodologies have made realistic to monitor the direct individual molecules one at a time. These techniques have opened up new perspectives, specially, in the study of biological systems, such as protein folding, enzymatic activity, or nucleic acids structure (227-235). Moreover, the progress of the methodologies has improved the knowledge on protein oligomers, providing new insights on the mechanism of their formation, characterizing their structure and the structural elements underlying their toxicity, resolving complex structural distributions and dynamics in a straightforward manner (128, 230, 232, 234, 236-238). The two main strategies for the single-molecule studies consist of nanomechanical manipulations and fluorescence spectroscopy (228, 237, 239), or a combination of them (240). While the first includes optical and magnetic tweezers (241-244) and atomic force microscopy (AFM) (245-249), the second includes polarization, lifetime, particle tracking (233, 250-252), Fluorescence Resonance Energy Transfer (FRET) (228, 253) and intensity measurements (237, 254-257). Importantly, with the fluorescence-based SM techniques being less perturbative than the nanomechanical ones, SM fluorescence are more suitable for studying molecular systems in their natural media, as has been proved both *in vitro* and *in vivo* experiments (230, 258, 259).

1.8 SINGLE-MOLECULE FLUORESCENCE SPECTROSCOPY (SMF)

The SMF methodologies are based on measuring the time-dependent fluorescence from individual molecules, typically in the context of confocal or total internal reflection microscopy. In the last decade, these techniques and their application to biophysical and biomedical systems have undergone an impulse because of the advent of innovative optical techniques, ultrasensitive single-photon detectors and charge-coupled devices (CCDs), avalanche photodiodes for photon counting (APDs), laser sources, powerful optics and confocal microscopes (260, 261). By SMF approaches, nowadays it is possible to reveal subpopulations, competing reaction pathways and transient intermediate states, which are frequently obscured in an ensemble average (230, 232, 258, 262).

Fluorescence correlation spectroscopy (FCS) is one of the main single-molecule methods applied to detect fluorescent proteins which diffuse through a confocal volume of a microscope. This method is extremely powerful as it can provide a precise measure of diffusion, being especially useful when a small diffusing object (small molecule, peptide or small protein) can bind to a larger particle, creating a large shift in the diffusion coefficient. From the labeled protein samples, FCS permits to examine size distributions (263) and velocity of diffusion (264-266). However, FCS is not optimal when the sample is heterogeneous. The brightest events overwhelm the correlation and dominate the analysis as bursts of very large amplitude create long-range temporal correlations.

Rather than making use of the fluorescence fluctuations, the direct measure of the fluorescence signal permits to use a lower concentration range than the used in FCS, of the order of 10-100 pM. This approach allows that only one molecule or aggregate is detected at a time, as a ‘burst’ of photons, in free-diffusing samples,

or as a spot with fluorescence intensity, for immobilized samples. Burst intensity approaches are revealed extremely useful to study the presence of different subpopulations involved in this process (267-270).

Likewise, a significant advance in fluorescence microscopy technology has resulted from the adoption of FRET (271) to the investigation of a wide spectrum of biological topics at the single molecule level (272). FRET allows the quantitative estimation of the biomolecular distances in the range of a few nanometers. It involves energy transfer between an excited donor fluorophore (D) to an acceptor molecule (A), via non-radiative dipole-dipole coupling. The now excited acceptor fluorophore can return to its ground state by the emission of a photon. Consequently, the donor emission and lifetime decrease, whereas the emission intensity and lifetime of the acceptor, in case of it is fluorescent, increase. As the photons emitted by donor and acceptor are at different wavelengths, their origin can be clearly distinguished using typical SMF detection schemes.

The theoretical basis of this process was developed already in the 1940s by Förster (273). The efficiency of energy transfer, E (equation 1.1), depends strongly on the D-A distance, the spectral characteristic of the D and A, and the relative orientation of their transition dipoles (273).

$$E = \frac{1}{1 + \left(\frac{R}{R_0} \right)^6} \qquad (1.1)$$

where R is the D-A distance, and R_0 is the Förster radius, the distance at which 50% of donors are deactivated via energy transfer. R_0 can be related to the spectral properties of the fluorophores, the relative orientation of the dipoles, and the properties of the medium, as appears in equation 1.2.

$$R_0 = 0.211\left(\frac{\kappa^2 \times \Phi_D \times J}{n^4}\right)^{\frac{1}{6}}(\text{Å}) \qquad (1.2)$$

κ is the orientation factor, which depends on the relative orientation of the donor and acceptor transition dipoles, Φ_D is the quantum yield of the donor, J is the overlap integral between the donor emission and the acceptor absorption spectrum, and n is the refractive index of the medium which separates D and A (271).

The FRET efficiency is usually estimated from the fluorescence quenching produced over the donor emission. E can be obtained from the emission intensity: $E = 1 - I_{DA}/I_D$ (where I_{DA} and I_D are the fluorescence emission intensities of the D in the presence or the absence of A, respectively). Importantly, the fluorescence lifetime of the D represents an independent estimation of the FRET efficiency. The latter can be also obtained from the equation: $E = 1 - \tau_{DA}/\tau_D$ (where τ_{DA} and τ_D are the fluorescence emission intensities of the D in the presence or the absence of A, respectively).

Due to its sensitivity, FRET allows to monitor the interaction of biomolecules that are spatially located at a distance within 1–10 nm, emerging as a molecular ruler (271). Because of this molecular information, the combination of FRET with SMF has entailed one of the major advances in the biophysical field, as it has permitted to measure distances directly at the molecular level (258, 272). This has been particularly interesting in the field of protein folding, because SMF-FRET experiments directly provide structural information (272). The first clear report of SMF-FRET measurements involved monitoring SM-assembly of the DNA double helix, in 1996 (274). Three years later, Deniz and colleagues performed an initial demonstration of the distance dependence of the FRET process in SM studies (275). Since then, the SMF-FRET technique has been rapidly developed and applied to a great variety of systems, from polymers to living cells (232, 234, 236-

238, 272). Moreover, this technique has revealed new and important information on protein folding, since its pioneering experiments (276-278), and its use has been extended to protein aggregation studies, even if its application to this problem has been rare (279). The first demonstration of the use of SMF approach to the study of protein folding at the SM level was performed by Jia and colleagues, in 1999. They studied the folding and unfolding fluctuations of a variant of GCN4-P1, a peptide derived from the yeast transcription factor (GCN4) (276). GCN4-P1 was labeled with a fluorescence donor and acceptor pair and immobilized for SMF-FRET measurements on a glass surface. A two-channel confocal microscope allowed the simultaneous acquisition of the D and A fluorescence signals. These experiments provided the data of the probability distributions of energy transfer efficiency as a function of added denaturant. It permitted to deduce the conformational fluctuations from the changes in donor-acceptor distance distribution.

In 2000, Talaga and colleagues reported SMF-FRET measurements on the folding and unfolding conformational equilibrium distributions and dynamics of a variant of GCN4 (277). In a correlation analysis of their data, they observed folding dynamics on a time-scale in agreement with ensemble experiments, but also showed slow conformational fluctuations for the unfolded peptide, probably because of interactions with the surface. The first ratiometric 'unfolding' measurements at SM-resolution were performed by Deniz and colleagues (278). In this pioneering study, Deniz observed the FRET changes of chymotrypsin inhibitor 2 (CI2), a model system for protein folding (280) to study protein folding under freely diffusing conditions. Through the analysis of the folded and unfolded subpopulations of CI2 during guanidinium chloride denaturation, Deniz obtained the first SMF-denaturation curves for a single-domain protein. Furthermore, the study emerged directions and innovations necessary to manage the future studies

of protein folding by SMF-FRET techniques. In 2001, Schuler applied SMF-FRET approaches to examine the properties of the free-energy surface for folding of the cold-shock protein from the hyperthermophilic bacterium *Thermotoga maritima* (281). The results of these experiments showed bimodal distributions of FRET efficiency, similar to what was observed in the previous study by Deniz on CI2 (278). Schuler studied the system at moderate denaturant concentrations, to ensure that both the native and the unfolded protein coexist. Two FRET peaks were observed plotting FRET efficiency at various guanidinium chloride concentrations. The increase in FRET efficiency, correlated with a decrease in the distance of the unfolded forms, probably was due to a collapse of more compact denatured conformations. More complex folding behaviour has been revealed by AFM combined with SMF techniques. Rhoades et colleagues studied the folding of adenylate kinase by immobilizing individual fluorophore-labeled protein in lipid vesicles (282) and on cold-shock protein from the hyperthermophilic bacterium *Thermotoga maritima* (283). The study, without the complexity of the system, provided valuable information resulting in the first model-free demonstration of two-state protein-folding dynamics.

In 2005, Kuzmenkina and colleagues applied SMF-FRET to the folding study of ribonuclease H (RNase H) (284). The study was performed with the protein immobilized on star-polymer surfaces, allowing to observe conformational changes of individual proteins for several hundred seconds. Such approach demonstrated the two-state folding behaviour for RNase H, and provided values of folding free energies and cooperativity comparable to what was obtained in the studies of diffusing molecules. Both set of experiments, from Rhoades and colleagues and from Kuzmenkina and colleagues, demonstrated the agreement between SMF-FRET and ensemble experiments. Moreover, they showed the possibilities of the application of immobilization methods for soluble proteins to SMF-FRET folding studies.

Until then, SMF was mainly used to identify molecular distance changes and kinetics. As a matter of fact, the numerous corrections required to measure accurate FRET efficiencies within single molecules did not permit the full realization of the SMF potential. Continuous technical and methodological advances are being introduced in the latest years to extract the maximum amount of information out from the biophysical systems. For instance, the use of dual-color excitation schemes, three- and four- color detection systems, or multi-parametric detection, including anisotropy and fluorescence lifetime have been recently implemented in novel SMF approaches (285).

One of these novel SMF methods, termed two-color coincidence detection (TCCD), was developed by Li and colleagues, members of the Klenerman's laboratory (286). They modified the conventional SMF-FRET scheme introducing a second laser that directly excited the acceptor, in order to characterize the proportion and stoichiometry of interacting biomolecules. In TCCD, the confocal volume was composed of two overlapped red and blue lasers, which permits to independently illuminate two different dyes. Two-color excitation and coincident fluorescence photon detection reduced the region where the two corresponding fluorophores were simultaneously excited. Li and colleagues applied this methodology to several strands of complementary DNA, either unlabeled or labeled with a single dye. The results showed the effectiveness of this approach to detect rare molecules in complex mixtures, which was not allowed with single color detection.

Significant progresses were made by Kapanidis and colleagues, who introduced a dual-color alternating-laser excitation (ALEX) on diffusing or immobilized biomolecules to study structure, interactions and dynamics (287, 288). The ALEX excitation scheme allows the recovery of the photons emitted by the donor and the acceptor upon direct excitation, and of the photons emitted from A upon D

excitation. Using this scheme, they were able to distinguish species based on D distance (from FRET) and stoichiometry (*S*). Thus, while the FRET efficiency informs on the D-A distances, the *S* informs on the relative number of D and A dyes (stoichiometry). Combination of the FRET efficiency and the *S* on 2D histograms allows virtual sorting of single molecule, and the authors defined this method as fluorescence-aided molecule sorting.

Other technical and methodological advancements in the SMF field involve adding multidimensional information, such as fluorescence lifetime (τ) (289) or anisotropy, giving rise to more advanced techniques, termed multiparameter fluorescence detection (MFD) (290, 291). The analysis of the fluorescence lifetime distribution was demonstrated as a powerful tool to discriminate among different subpopulations, unraveling them at the SM level. A powerful methodology combines MFD with multi-color approaches, requiring lasers working in pulsed interleaved excitation scheme (PIE) (292), the so-called MFD-PIE (290). As an example, Hillger and colleagues applied a multi-parameter approach, using four detection channels and pulsed excitation for the *in vitro* detection of FRET-labeled samples of rhodanese (293). In their study, they correlated FRET efficiency with lifetime information, fluorescence intensities, lifetimes, anisotropies and distances within individual particles. Their method permitted a direct separation of folded and misfolded subpopulations.

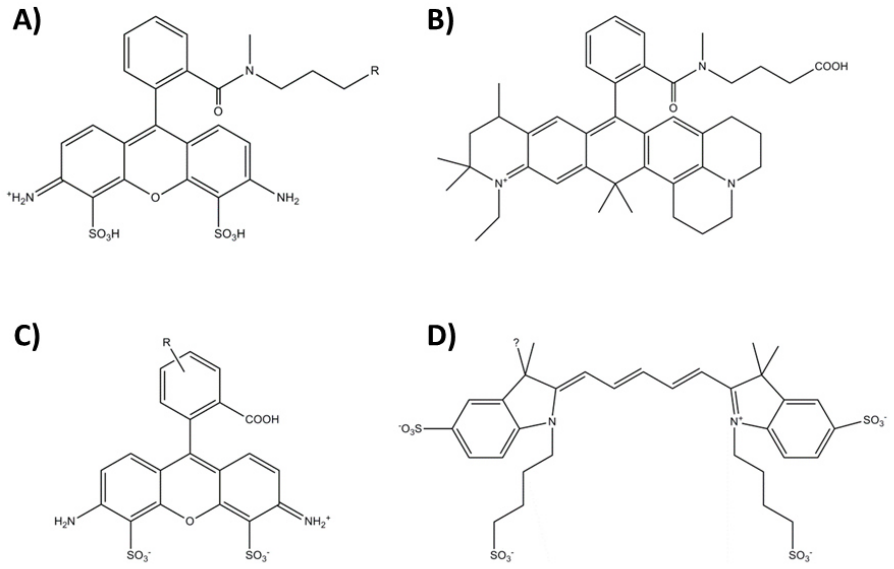
All these novel SMF techniques open up new possibilities to understand more complex biological systems. Hence, these techniques are of particular interest to study protein aggregation. This will be discussed in section 1.9.

1.8.1 Optimum properties of fluorescent dyes for SMF-FRET studies

A SMF strategy to study protein aggregation and to detect the species involved in the process usually require covalent modification of particular sites of proteins or their mutants, to permit the attachment of fluorophores. To introduce one or several dyes in a protein, it has to be taken into account the relative position of the fluorophore into its structure. This increases the technical challenging for building and validate structural models, and may also alter the aggregation behavior of the protein of interest (294). Thus, after labeling, in all these experiments it must be checked whether these modifications affect the process being investigated. Especially in the mutant proteins, the labeling should not alter either the secondary structure of the protein or the kinetics of overall aggregation process. To this end, the behavior of labeled proteins may be compared with that of unlabeled proteins using ensemble methods, for example by collecting CD spectra of both unmodified and covalently-modified protein. In addition, for FRET measurements, it also has to determine the accurate inter-dye distance.

The ideal fluorophore for SMF studies must be bright, exhibiting a maximum extinction coefficient, $\varepsilon_{MAX} > 80000\ M^{-1}cm^{-1}$ and a high fluorescence quantum yield ($\Phi > 10\%$) (272). It must be photostable, with minimal photophysical/chemical and aggregation effects, small and water soluble with sufficient forms of bioconjugation chemistries. Its fluorescence lifetime must be of the order of a few nanoseconds and its size of about 1 or 2 nanometers. Moreover, a fluorophore pair should have wide spectral separation between D and A emissions (295) and should have almost the same quantum yields and detection efficiencies (262, 272, 296). To link the fluorophore to the proteins, the common strategies involve N-terminal or lysine amine groups, cysteine thiol groups or unnatural amino acid with ketone group (272, 297). Among these, the labeling of

cysteine residues (298, 299) with maleimide based reactions constitutes the most popular labeling scheme, even though other alternatives exist (297, 300, 301). In our case, we required the labeling of thiol (-SH) group of cysteine by the maleimide derivatives of ATTO-488 and ATTO-647N (Figure 1.9). These different fluorescent probes have revealed suitable to detect and distinguish between the structural diversity of different species populated the aggregation reaction.



	Fluorophore	Absorption max (nm)	Emission max (nm)	Quantum Yield	Fluorescence lifetime (ns)	Extinction coefficient ($M^{-1}cm^{-1}$)
a)	ATTO488	501	523	0.80	4.1	90000
b)	ATTO647N	644	669	0.65	3.5	150000
c)	Alexa488	495	519	0.92	4.1	71000
d)	Alexa647	650	668	0.33	1.0	239000

Figure 1.9 Spectroscopic properties of commonly used organic dyes. Reprinted with permission from Molecules, Vol.19, 15824-15865 (2014).

Finally, in SMF experiments the labeled molecules must be present from pM- to nM-concentration. So, background signal has to be low, in order to distinguish the small signals obtained from individual molecules. This may be obtained using simple aqueous buffers, although it has been studying the effects of more complex environments, as live cells (302).

1.9 INVESTIGATING AGGREGATION BY SMF

The advent of innovative and advanced SMF techniques and their application to biophysical and biomedical systems has revolutionized the approach to the study of biomolecular systems. In sum, the main advantages of the SMF methods in aggregation study are:

- great sensitivity: both conformational and concomitantly size changes can be monitored, even if it occurred during early misfolding and assembly events as well as during the later stages of aggregation. This includes the characterization of subpopulations in a heterogeneous ensemble for separate and the direct detection of very rare events;
- low concentration ranges, which is crucial for observing the initial stages of aggregation events because of the formation of larger aggregates is generally disfavored in such conditions. Moreover, low concentrations permit to minimize the probability of multiple molecules being observed at the same time;
- site-specific attachments of dyes to protein of interest, which is a useful strategy of monitoring the emergence of oligomeric species and fibrils at different stages of the aggregation process;
- obtaining various parameters from the same probe within an ensemble of structurally diverse molecular species, as fluorescence intensity, fluorescence

anisotropy, spectral shifts, lifetimes, etc. This possibility is very important in aggregation studies, because of the major numbers of observables on the same reaction that results much convenient to tackle the large variety and parallel pathway that characterize the process;

- discriminating between subpopulations formed during the aggregation, which involves taking into account several aspects of the fluorescence signal. These include the brightness and the lifetime of the fluorophores, their spectral features, time correlations (303), polarization anisotropy (304), quenching and FRET (234).

The initial studies of protein aggregation using SMF techniques were based on the analysis of burst intensities arising from labeled proteins (305-307). Protein aggregates give rise to much brighter events than monomers do. This permits to detect, count and follow the formation of aggregates. This method has provided useful information on the aggregation of α-synuclein (308-310), prions (306), tau protein (311), and co-aggregation of different NOTCH3 mutants (312).

In a step forward, Orte and colleagues used TCCD, the dual-color excitation and coincidence detection scheme, to study the aggregation of the amyloidogenic SH3 domain of PI3 kinase (PI3-SH3) (148). In this oligomerization study, the authors labeled the PI3-SH3, separately, with Alexa Fluor 488 or Alexa Fluor 647 fluorophores. Then, they incubated equimolar mixtures of the two labeled proteins and studied the changes on the oligomeric formation as the aggregation proceeded by a burstwise analysis. The fluorescence bursts that were coincident in time on both donor and acceptor channel revealed the presence of oligomers, while no coincident fluorescence indicated single-label monomers (Figure 1.10). Moreover, the fluorescence signals were corrected for coincident events produced by chance encounters, using a criterion developed by themselves a couple of years before

(313). Applying this correction to the TCCD method, they estimated the size distribution of the fraction of oligomers of the amyloidogenic protein, and followed the oligomer–monomer equilibrium in solution.

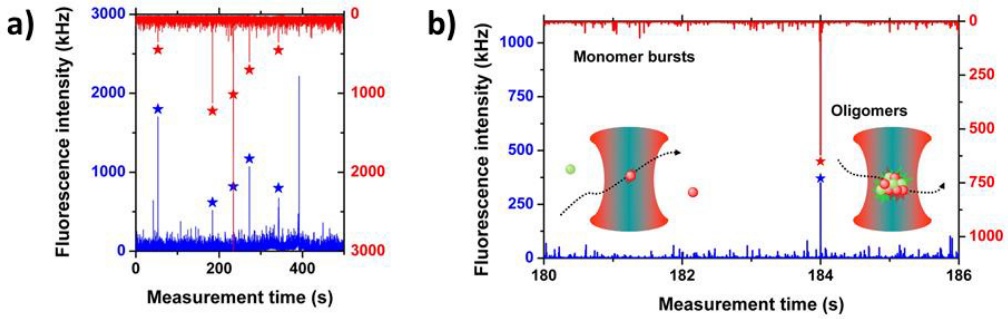


Figure 1.10 TCCD method used to detect oligomeric aggregates. **A:** The coincident fluorescent bursts on both channels show the presence of oligomers (marked asterisks). **B:** expansion of fluoresce bursts in the left image. If compared, the monomer events are not coincident and are less intense than those due to oligomers. Reprinted with permission from PNAS Vol.105, 14424-14429 (2008), copyright 2008 National Academy of Sciences, USA.

In a different work, Chiou and colleagues used the TCCD scheme to identify the species involved in the early stage of the polymerization of the neuroserpin (314). Using TCCD, they quantified the fraction of the coincident events (the association quotient, *Q*), whose value is related to the fraction of associated molecules (313). This permitted the study of the equilibrium between monomers and oligomers of neuroserpin, following the changes of the population as polymerization progressed. Their experimental results were used to test different kinetic models, to understand how neuroserpin polymerization proceeds. TCCD was also applied by Narayan and colleagues in an aggregation study of Aβ (315). They investigated the effects of oligomerization inhibitors or molecular chaperones to the heterogeneous distribution of oligomers formed during the aggregation. Using the TCCD scheme, Narayan and colleagues demonstrated the capability of the chaperon clusterin to sequestrate the Aβ intermediates and to regulate the

oligomers diffusion by fibril dissociation. A combination of FRET with TCCD methods was applied by Cremades and colleagues to monitor α-synuclein (αS) aggregation (316). The αS was tagged with Alexa Fluor 488, as the donor, and Alexa Fluor 647, as the acceptor. By this scheme, the samples were illuminated by the blue laser to excite the donor-labeled αS directly, while the acceptor dye was excited from contiguous molecules labeled with the donor, as result of energy transfer. The mentioned coincident fluorescence detection criterion (313) permitted to distinguish oligomers from monomers, depending on whether the fluorescence emission was detected as coincident bursts in both red and blue emission channels or only in the AF488 emission channel. The intensity of the fluorescence bursts of each oligomer detected was compared to the average intensity value for a single αS monomer, in order to estimate the size of the oligomeric species. They observed two different types of oligomers formed in the aggregation process of α-synuclein, exhibiting distinct inter-molecular FRET efficiency values (Figure 1.11). The oligomer type was also changing as oligomer growth proceeded. Interestingly, the oligomeric types detected showed different cytotoxicity, which suggested how conformational changes may be correlated to the capability of oligomer species to promote cell death.

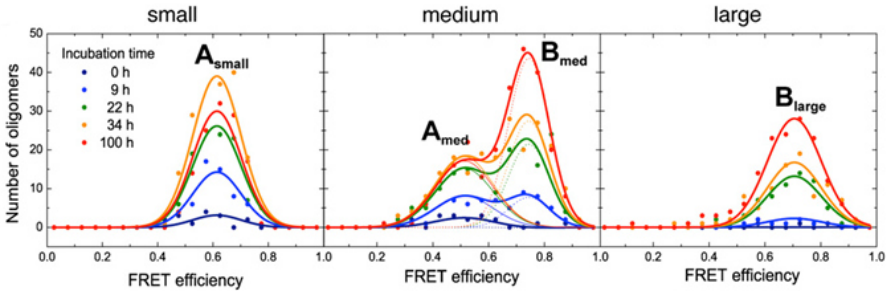


Figure 1.11 Oligomer size *versus* FRET efficiency correlations of αS-oligomers Reprinted from Cell, Vol. 149, 1048–1059 (2012) under the Creative Commons Attribution License (CC BY).

Considerable enhancements to SMF experiments came from the use of microfluidic devices and fast-flow mounted onto confocal optical systems. The main advantage of such systems relies on increasing the rate at which SM data could be acquired (317). Strikingly, data acquisition of 3 hours may be reduced to just over 5 minutes (318). Another advantage is that the flow velocity of the individual species is not dependent of their diffusivity or size, with molecules being under laminar flow. This avoids that same species, especially the large ones, can take up the confocal volume for a protracted period of time, producing multiple counting of events.

Some members of the Klenerman's laboratory exploited fast flow microfluidics in SMF measurements in diffusing molecules, to study αS and tau amyloid aggregation. Under microfluidic conditions, using SMF-FRET and kinetic analysis Shammas and colleagues studied labelled tau proteins (K18 construct) during their aggregation into filaments (319). From their analysis emerged that tau protein aggregation may proceed through monomeric assembly into small oligomers, where a slow structural conversion step precedes fibril formation. Using this approach, they compared the K18 construct with two disease-related mutants, and they were able to observe how these mutations alter the aggregation pathway. Horrocks and colleagues, in a study on the in vitro αS oligomer growth, observed the formation of low and high FRET efficiency αS oligomers (318). Working under fast-flow microfluidics, they reduced the time required to collect single-molecule data a factor of 150 fold with respect to stationary measurements. Moreover, they studied the effect of different salt conditions on the stability of these two different types of αS oligomers. Simply changing the ionic strength of the buffer, they suggested an additional way to distinguish the αS oligomers, previously characterized from FRET efficiency, susceptibility to proteinase-K digestion and cytotoxic properties (74). Tosatto and colleagues, in a study of

protein oligomerization of wild-type and three pathological mutants of αS, investigated structural properties, kinetics of formation, interconversion and accumulation for the oligomers of each αS construct (320). Comparing with the wild-type, two of the three αS mutants exhibited a preference, depending on the type and position of the mutation, for the accumulation of one of the two types of oligomers formed. On the contrary, one of the three mutants showed a different behaviour, forming a complex mixture of species. Finally, Iljina and colleagues exploited microfluidics SMF-FRET experiments on flowing molecules and kinetic analysis to characterize the early stages of αS aggregation (321). They proposed that the initial nucleation step is not an elementary step, which may involve a change in conformation from low-FRET to high-FRET αS oligomers. Their results confirmed that αS aggregation populates three distinct species: low-FRET and high-FRET αS oligomers, and the mature fibrils. This process should involve two unimolecular structural conversion steps, from the disordered to more compact oligomers, and then to fibrils.

1.10 EXPERIMENTAL APPROACH OF THIS WORK WITHIN THE CONTEXT OF THE PREVIOUS LITERATURE

It is now well established that the ability to self-assemble into stable and highly organized structures such as amyloid fibrils is not limited to the small subset of proteins or peptides that are involved in disease, but rather a property shared by most proteins (208, 322-324). The use of model proteins, even though unrelated to any disease, has become a potent approach to investigate the mechanisms of formation of amyloids fibrils. As a matter of fact, these proteins are easier to handle than those related to diseases, and often show high aggregation propensity.

One of the first model protein used to study the amyloid formation was the Src-homology region 3 (SH3) domain of bovine phosphatidylinositol-3'-kinase, which forms amyloid fibrils under acidic conditions (325), and which has been extensively used to study aggregation (148, 326-328). SH3 domains are well characterized, small globular protein domains very extended on the proteome being part of a wide variety of proteins (such as kinases, lipases, GTPases, adaptor proteins, structural proteins and viral proteins) and generally involved in intracellular signal transduction, molecular recognition or protein-protein interaction (329-332). They are composed by a single polypeptide chain composed between 60 and 85 amino acid residues (333) and share a similar tridimensional structure. The α-spectrin SH3 domain (Spc-SH3) is a small domain of 62 residues that folds into an orthogonal β-sandwich containing three β-hairpins. Several studies, including some of our research group, have studied the mechanism of amyloid fibril formation of the Spc-SH3 (201, 328, 334-337), whose wild type form as well as several mutants are rapidly converted to amyloid fibrils at mild acid pH (334, 338). Specifically, Morel and colleagues (338) found that the single mutation of asparagine 47 to alanine at the distal loop, which destabilizes the folding nucleus of the Spc-SH3, favors the rapid formation of amyloid fibrils *in vitro* under mild acidic conditions, at which the majority of the protein is in its native state. They observed that the kinetics of assembly and the final morphology of the amyloid fibrils of the single-mutant N47A of the Spc-SH3 (N47A Spc-SH3) could be strongly modulated by temperature and NaCl concentration. Different oligomers populated the early stages of the process of aggregation depending on the experimental conditions used (201). Ruzafa and colleagues (339) focused their work on the kinetics of N47A Spc-SH3 aggregation, and specifically on the characterization of the oligomers formed into the earliest stages of amyloid formation. They suggested that under conditions of rapid aggregation, in absence of lag phase, a rapid formation of amyloid oligomeric nuclei results in curly and

relatively disordered fibrillar species. This rapid aggregation obeys to a high-order irreversible kinetics, suggesting a model where a rapid conformational change within any of the oligomers and irrespective of their size could be the rate-limiting step that leads the amyloid nucleation. They proposed that these early oligomers of different size, accumulated at diverse rates depending on the conditions, must determine the final morphological and structural properties of the protofibrils and amyloid fibrils.

In a recent work, we studied the aggregation of a modified mutant of the N47A Spc-SH3 (N47A-SH3) (340) using single molecule fluorescence techniques, specifically, a dual-color PIE scheme combined with fluorescence lifetime correlation spectroscopy (FLCS) (341), an advanced modification of conventional FCS. Using FLCS-PIE we obtained the oligomer size distributions and kinetics of the soluble oligomers on-pathway of amyloid fibrillization of the N47A-SH3. This method revealed a rapid increase of the oligomer sizes during the initial hours, shifting towards larges sizes and peaking after 48 hours of incubation (Figure 1.12). Remarkably, a population of soluble oligomers remained in solution after several days of incubation, even if the sample was mainly fibrillar. These results may be associated to the dynamic nature of the aggregation process, by which the mature fibrils may be a source of small oligomers (167). This also suggested the impossibility of certain oligomers to incorporate into mature fibrils due to their inadequate structure. This can be explained proposing a mandatory structural rearrangement for the oligomers before joining into mature fibrils, to get the appropriate β-sheet conformation.

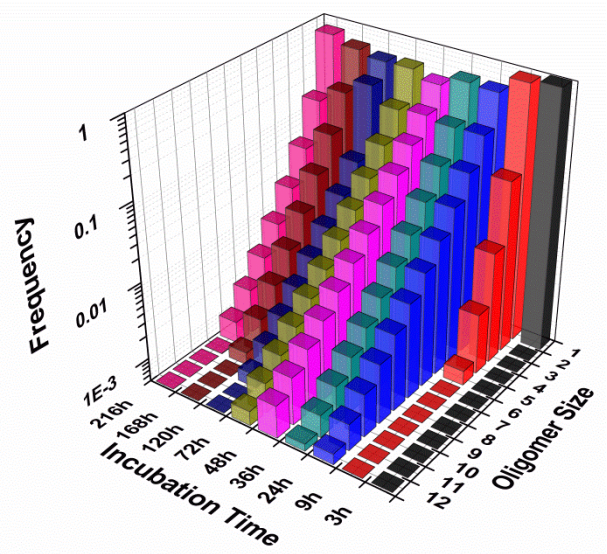


Figure 1.12 Oligomer size distribution as a function of incubation time obtained using PIE-FLCS methodology. Adapted from Paredes et al., *Int. J. Mol. Sci.*, **13**, 9400-9418 (2012) under the Creative Commons Attribution License (CC BY).

A further step in understanding the oligomerization process may be obtained using other strategies, such as applying TCCD and FRET methods for the study of intra-oligomeric energy transfer between donor and acceptor fluorophores within the same oligomer. Because of this, in this Thesis, we have thoroughly investigated the early stages of the aggregation of the N47A-SH3, using FRET and advanced SMF tools, such as PIE excitation scheme, coincidence criterion for burst picking, lifetime information, and MFD analyses. In our SMF-FRET experiments, each monomer and oligomer, which passed across the confocal excitation volume, was illuminated with two alternating lasers, detected, and characterized. The oligomeric features, such as size, arrangement, and cellular toxicity, have been scrupulously, deeply and reasonably investigated, combined with multi-disciplinary experiments.

2 – OBJECTIVES

2. OBJECTIVES

Protein aggregation has become an issue of high importance in some research fields, such as biology, chemistry, biotechnology and medicine. An emergent number of evidences show that the incorrect protein folding and its aggregation, many times in the form of amyloid fibers, lead to the formation of insoluble cellular deposits. This process is related to a group of diseases that are classified under the name of conformational diseases (342).

Initially, the toxicity associated with the formation of amyloid fibrils was attributed only to mature fibers. However, the information obtained during the last few years indicates that fibers are relatively inert and very stable biological species. In addition, intermediate species have been identified that play a key role in the mechanism of aggregation. In fact, there are substantial evidences that oligomeric intermediates formed in the early stages of aggregation are implicated in the cell death. A complete view of the aggregation process can be provided by the characterization of these intermediates. The deep understanding of the mechanism and of the elements involved in it may suggest effective strategies to prevent or cure misfolding diseases.

Despite the recent developments in the characterization of the structure and properties of amyloid fibers, the details of the mechanism are still to be determined.

Recently, the study of the protein folding has undergone a boost thanks to the new SMF techniques, which actually provides a method to study in full the energy landscapes of the proteins. This includes not only the folding process, but also the interactions with different species to form active compounds.

The use of fluorescent tags for labelling proteins related to the formation of amyloid fibers has opened new pathways to the study of the aggregation. Using a confocal microscope and extremely sensitive detection, we can obtain significant information from the protein-dye system such as the fluorescence resonance energy transfer (FRET) or the oligomer growth and size distribution, among others. Molecular subtypes, heterogeneous subpopulations or even intermediaries are just a few aspects detected thanks to this technique and they can explain by themselves the increasing significance of this methodology: SMF permits to reveal what seems "invisible", because inaccessible to other conventional techniques. These results offer new possibilities to elucidate the mechanism of the amyloid fibril process.

Nowadays, there are only few applications of SM-methods on the study of fibril amyloids. Recently, Dr. Orte developed a methodology of single molecule based on the two color coincidence spectroscopy (TCCD) (343). This scheme provided information about the oligomer size distribution, the kinetics of fibril formation and growth, as well as stability. This procedure was later applied efficaciously in the study of neuroserpine polymerization and in the interaction with the Aβ protein (344), which are the causes of familiar encephalopathy FENIB and Alzheimer's disease, respectively. Furthermore, this methodology has proved to be useful to study the attachment of Aβ oligomers caused by clusterin molecular chaperone (345). Therefore, these methodologies applied to the study of individual molecules permit to observe small biological systems (346, 347), whereas other techniques only provided information on the ensembles.

The main objectives of this Thesis are the characterization of the pre-fibrillar oligomeric intermediates involved in the early steps of aggregation and the understanding of their role in amyloid formation. In order to investigate the

oligomerization down to the single-molecule level, we combine the SM-techniques with conventional biophysical techniques.

The project is focused on the study of the pre-fibrillar nuclei formation of the modified mutant of the N47A variant of the α-spectrin SH3 domain (N47A-SH3). The N47A-SH3 system is capable of forming amyloid fibers under certain conditions, as it was widely demonstrated by previous works of our research group (348-350).

The low concentration conditions used in SMF experiments, allow to analyze the aggregation since its early stages, in order to complete the gaps in the mechanism understanding. The study proposed is multidisciplinary, combining single-molecule fluorescence spectroscopy technique with a pulsed interleaved excitation approach (SMF-PIE), together with conventional bulk techniques, such as DSC, DLS, FEE. Therefore, the aim of this work is to elucidate the relationships between fibril structures, properties, experimental conditions and the mechanism of their formation.

The aims of the Thesis include:

1. Expression, purification and labeling of the N47A variant of the α-spectrin SH3 domain engineered with a six-residue tag (N47A-SH3).

2. The biophysical characterization of the N47A-SH3 and of the N47A-SH3 labeled with ATTO dyes through conventional techniques (DSC, CD, FEE, DLS) and transmission electron microscopy (TEM), in order to check the effects of the protein modification (labelling) on the amyloid formation. Moreover, this includes identifying the techniques that we can apply to the system under study.

3. The identification and the characterization of early oligomeric species in the oligomerization process, using conventional biophysical and single-molecule techniques.

4. Performing aggregation studies of the labeled system. After elucidating the initial conditions for the study, the next step is to investigate the structure and the kinetic of the protein aggregation from the initial steps to the mature fibril formation, at the end of the process.

5. The characterization of the structural changes in oligomers by fluorescence lifetime image microscopy with a pulsed interleave excitation scheme (FLIM-PIE).

6. Once identified the crucial features in the early phase of the aggregation, we correlate the different types of pre-amyloid oligomers with their toxic effect in leve cell cultures, through cell viability and toxicity studies.

3 - EXPERIMENTAL METHODS

3 - EXPERIMENTAL METHODS

In order to study the mechanism of the aggregation of the N47A-SH3, we have employed a wide range of biophysical methodologies, including single-molecule fluorescence spectroscopy (SMF). The latter has required the label of the protein samples with SMF-ready fluorophores. For this reason, we used a construct of the N47A mutant engineered with a six-residue tag (Gly-Ser-Gly-Ser-Gly-Cys) at the C-terminus, necessary for the labelling reaction. While ensemble experiments and TEM techniques required large amounts of protein, SMF experiments only needed sample concentrations of the order of pM or nM.

3.1 PROTEIN PREPARATION

For the N47A-SH3 synthesis, the plasmid pBAT4:N47A-SH3 containing the N47A-SH3 domain clones was used to overexpress the protein in a BL21 (DE3) strain of *Escherichia coli* cells. The collected cells were lysed via sonication in a 5 mM sodium citrate buffer at pH 3.0, in the presence of 1 mM dithiothreitol (DTT) to avoid S-S interaction between free cysteine residues. The protein was recovered from the supernatant by precipitation with ammonium sulfate at 75% saturation, and then, the precipitate was solubilized in 0.050 M sodium phosphate buffer, pH 7.0, containing 0.15 M NaCl, 7 M urea, and 1 mM DTT. A size exclusion chromatography step using a Hi-Load Superdex 75 column (FPLC Hi-Load 26/60 Superdex 75 prep grade, GE Healthcare Life Sciences Corp., USA) permitted to separate the protein fractions from the rest of impurities present. Protein fractions were extensively dialyzed against pure water to remove urea and finally

lyophilized for storage until analysis. Protein purity was confirmed via SDS-PAGE electrophoresis and HRMS Q-TOF (The molecular weight of the N47A-SH3, MW, is of 7625.0005 Da).

3.2 PROTEIN LABELING WITH SMF-READY FLUOROPHORES

The N47A-SH3 was covalently labeled with maleimide-modified ATTO 488 (A488, ATTO-TEC GmbH, Siegen, Germany) or ATTO 647N (A647N, ATTO-TEC GmbH, Siegen, Germany) via the cysteine thiol moiety (Figure 3.1). The fluorophores from the ATTO series exhibit proven high quantum yield, stability, and excellent performance for SMF experiments.

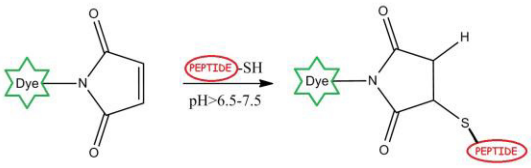


Figure 3.1 N47A-SH3 labeled with ATTO dyes via the cysteine thiol moiety.

For the labeling reaction, we dissolved the N47A-SH3 protein in 0.05 M Phosphate-Buffered Saline (PBS) solution, containing 1 mM TCEP, at pH 7.2. To enhance dye reactivity, we prepared the solution in a dye-to-protein ratio of 4:1, immediately before starting the labeling reaction. Then, we added to the protein solution freshly prepared, maleimide-modified, A488 and A647N dyes. The reaction solution was incubated at 21 ºC for 16 h, stirring and protected from light.

The two labeled proteins, N47A-SH3-A488 (A488-labeled N47A-SH3) and N47A-SH3-A647N (A647N-labeled N47A-SH3), were then collected through PD-

10 desalting columns (PD-10 Sephadex G-25M, GE Healthcare). The protein concentration and labeling efficiency were calculated using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific). The molar extinction coefficients used were ε_{503}=90,000 $M^{-1}cm^{-1}$ for the A488, ε_{650}=150,000 $M^{-1}cm^{-1}$ for the A647N (as indicated by the supplier's information), and ε_{280}=15784 $M^{-1}cm^{-1}$ for the unlabeled N47A-SH3. As both A488 and A647N dyes absorb at 280 nm, increasing the absorbance for the labeled proteins, we used a correction factor (CF) indicated by the provider (0.1 for the A488 dye and 0.05 for the A647N) (351) to *adjust t*he absorbance of the protein at 280 nm. Then, the protein samples were equilibrated through extensive dialysis against deionized water at pH 3.0, lyophilized and stored at –42 ºC to be sure to keep them stable for long-term storage. Previously studies on the N47A-SH3 performed in our laboratory have demonstrated that the lyophilization had minor effect on the structure and the thermal stability of the protein (Figure 3.2).

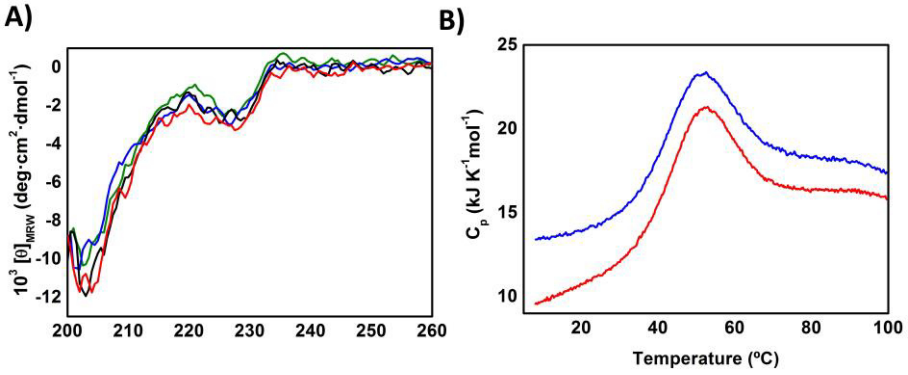


Figure 3.2 A: Far-UV CD spectra of N47A-SH3 at concentration of 0.15 mg mL^{-1} recorded at 37 ºC after dialysis in 0.1 M Gly, 0.1 M NaCl, pH 3.2, collected before (green line) and after (blue line) lyophilizing. Black and red line: far-UV CD spectra of N47A-SH3 at 0.11 mg mL^{-1} recorded at 25 ºC, after dialysis in 0.1 M Gly, 0.1 M NaCl, pH 3.2, collected, respectively, before (red line) and after (blue line) lyophilizing. **B:** DSC thermograms of N47A-SH3 at 0.94 mg mL^{-1} in 0.1 M Gly, 0.1 M NaCl buffer, at pH 3.2 and collected, respectively, before (red line) and after (blue line) lyophilizing.

A second chromatographic step using HPLC column (Water delta 600, Column: C18 VYDAC HPLC Columns, Photodiode Array detector Waters 996) was used to further separate labeled proteins from remaining unlabeled materials. The experimental conditions of the solvent gradient is shown in Table 3.1. This step permitted to reach a higher purity of the labeled proteins, which finally resulted close to 99%. This was confirmed by HRMS Q-TOF detection (MW=8337.2002 Da for the N47A-SH3-A488, MW=8392.8008 Da for the N47A-SH3-A647N). Finally, the N47A-SH3-A488 and the N47A-SH3-A647 solutions were separately lyophilized.

Table 3.1 HPLC assays. The solvents employed were: 0.1% v/v solution of Formic Acid in Water (A), 0.1% v/v solution of Formic Acid in Acetonitrile (B). Flow: 3 mL/min Retention time: 27 min for N47A-SH3, 33 min for N47A-SH3-A488, 43 min for N47A-SH3-A647N.

Time	A (%)	B (%)
0	100	0
5	100	0
10	70	30
15	70	30
75	60	40
80	0	100
90	0	100
90.10	100	0
120	100	0

For aggregation experiments, the lyophilized samples were directly dissolved in the buffer, then centrifuged for 2 min at 14000 rpm in a microcentrifuge (Hettich) to remove preexisting fibrillar species and the pH (3.2) checked and adjusted if necessary. The protein concentration was determined using UV-Vis spectroscopy by measurement of absorbance at 280, 501 and 650 nm for the N47A-SH3, N47A-SH3-A488 and N47A-SH3-A647, respectively. The spectrophotometer used was a Lambda 650 Perkin-Elmer UV Winlab or Cary-50-Bio Varian spectrophotometer.

3.3 SMF-PIE INSTRUMENTATION AND METHODS

SMFS experiments were carried out by a MicroTime 200 time-resolved confocal fluorescence microscope (PicoQuant GmbH, Germany) using dual-color pulsed interleaved excitation (PIE) (292) and Multi-parameter Fluorescence Detection (MFD) (290, 352). The excitation sources were two spatially overlapped pulsed lasers (<100 ps) alternated on a nanosecond time scale. Specifically, the two lasers employed were a PicoQuant LDH-P-C-470 system at 470 nm and a PicoQuant LDH-P-635 system at 635 nm, the outputs of which were directed to the back port of an Olympus IX-71 inverse microscope via fiber coupling. The overall excitation pulse frequency was of 20 MHz. The excitation power was 286 μW for the 470 nm and 22.7 μW for the 635 nm laser. To reach PIE, the second laser (635 nm) was delayed by 24 ns (ORTEC DB463 delay box, Ametek, USA), and with both lasers triggering alternatingly by a multichannel picosecond diode laser driver "Sepia I" from PicoQuant (Berlin, Germany). The beams were focused into 60 μL of sample solution deposited on a microscope cover slide (Thermo-Scientific, Menzel-Gläser, Brawnschweig, Germany) through a high numerical aperture oil-immersion objective (Plan Achromat 100x/1.40NA, Olympus), after passing a custom-made dual-band excitation dichroic mirror (Chroma filters). The

fluorescence photons were collected by the same objective and imaged onto a 75-μm pinhole (Melles Griot, USA). After passing through the aperture, the transmitted light was separated by a 600 DCXR dichroic beam splitter (AHF/Chroma) into green (donor) and red (acceptor) fluorescence. Then, the light was filtered with band-pass filters, 520/35 (Omega Filters) for the donor and 685/70 (Omega Filters) for the acceptor channel. Finally, the light was refocused into two single-photon avalanche photodiodes (APD) (SPCM-AQR-14, Perkin-Elmer Optoelectronics). The photon detection events were registered in time-tagged time-resolved (TTTR) mode (353) to reconstruct the fluorescence decay traces and establish the temporal time gates for filtering the burst traces (290, 292).

SMFS traces were analyzed using SymPhoTime software package (PicoQuant GmbH). The fluorescence time traces and time-resolved fluorescence data for donor emission, FRET signal, and acceptor emission were obtained from MFD-PIE measurements (290) by applying appropriate time-gates to the overall traces (Figure 3.3. The time filtering allowed for the definition of three fluorescence traces: donor (F_{A488}), FRET (F_{FRET}) and directly excited A647N (F_{A647N}). The analysis of the time-resolved information was performed by deconvolution fitting, using a mathematically reconstructed instrument response function (IRF), obtained from a control measurement at high dye concentration, and the maximum likelihood estimation (MLE) as the fitting optimization criterion (352, 354), because this method exhibits the best performance low count rates (355). Our analysis allows to determine the donor lifetime, τ_{A488}, for each individual burst (290, 356), assuming a single exponential decay of the fluorescence intensity, using the burst integrated fluorescence lifetime, BIFL, method (355).

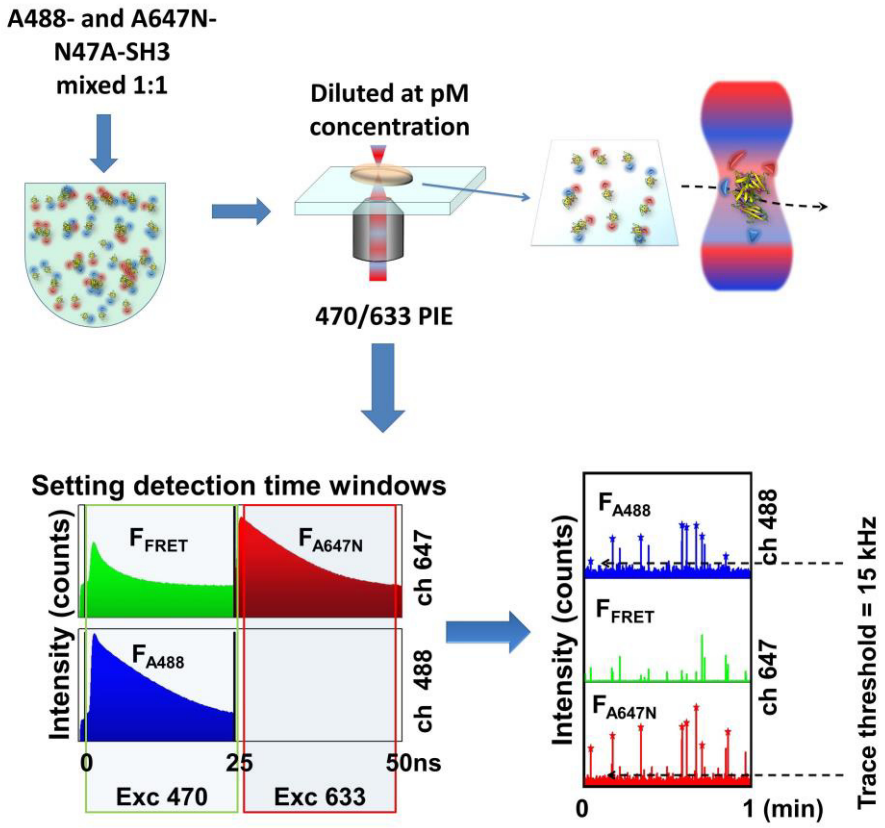


Figure 3.3 Schematics of the SMF-PIE experiments. Dual-color PIE allows to set time gates in the nanosecond timescale to filter the fluorescence signal in three distinguishable traces: F_{A488} (donor), F_{FRET}, and F_{A647N} (directly excited acceptor).

The fluorescence signals (F_{A488}, F_{FRET}, and F_{A647N}) were corrected by three correction factors: the spectral crosstalk of the donor dye in the acceptor detection channel (β), the fraction of direct excitation of the acceptor by the donor laser (α) and a factor that accounts for the different detection efficiencies of the channels and the quantum yield of the dyes (γ) (288, 290). The estimation of these correction factors were performed using 45-basepair single-stranded DNA labeled with A488 or A647N dyes, and a dual-labeled (with A488 and A647N) 45-basepair double-stranded DNA, maintaining the instrumental configuration used

for the labeled N47A-SH3 SMFS measurements. The samples of single- or double-stranded DNA were diluted to a final concentration, respectively, of 2 nM and 50 pM, in TEN buffer (10 mM TRIS, 1 mM EDTA, 0.10 M NaCl, pH 7.5) with 0.0025% TWEEN 20. The fluorescence traces were recorded during 2 min for the single-stranded DNA SMFS-measurements (to estimate the α and the β factor, which values resulted in 0.017 and 0.0174, respectively) and during 30 min for the single stranded SMFS-measurements (to estimate the γ factor, which value resulted in 0.95).

For the N47A-SH3 labelled samples, the selection and counting of single-molecule bursts were performed by setting a value of 15 kHz for both F_{A488} and F_{A647N}, to clearly distinguish a burst originating from a fluorescent particle in the confocal volume from the background signal. Coincident and FRET events are determined when a burst is detected simultaneously in F_{A488} and F_{A647N}. Using the time gating and the PIE scheme removes the influence of the ubiquitous 'zero-peak' in FRET histograms (290, 292). A general scheme of the SMFS-PIE configuration used in our SMFS experiments is represented in Figure 3.3.

The detected fluorescence bursts originated from the photons permit to obtain several different properties of the fluorescent particles. The parameters of interest were the association quotient, *Q*, defined as the fraction of coincident events (313, 314); the FRET efficiency, *E* (357), which depends strongly on the distance between the donor and the acceptor fluorophore; the ratio F_{A488}/F_{A647N} of the coincident events (275, 313), related to the stoichiometry of the interaction; and finally, the total fluorescence intensity due to the sum of fluorescence intensity of the directly excited donor and the acceptor (358).

3.3.1 Association quotient, *Q*

The association quotient, *Q*, represents the fraction of coincident events over all the detected single molecule events. A coincident event is defined according to the criterion of simultaneous detection of a burst in both channels (F_{A488} and F_{A647N}). A coincident event is produced when a molecule or aggregate containing both fluorophores crosses the overlapped excitation volume.

However, this coincidence criterion cannot discriminate the situation of two or more non-associated molecules with different fluorophores entering simultaneously in the probe volume. These events give rise to coincident bursts due to chance encounters. A more precise analysis requires removing this contribution from the total number of coincident fluorescence events. To resolve it, we employed a validated probabilistic method developed by Orte and colleagues (313). Following this method, the rate of chance coincidence events, n_C, was subtracted from the total number of coincident bursts (n_T) so that the rate of the real coincident events resulted as (n_T - n_C). For each measurement, we count the total number of events in the F_{A488} (n_{A488}) and F_{A647N} channels (n_{A647N}), so that the association quotient, *Q*, is defined as follows (313, 358):

$$Q = \frac{n_T - n_C}{n_{A488} + n_{A647N} - (n_T - n_C)} \qquad (3.1)$$

The *Q* value provides information about the fraction of aggregated species (or oligomers) in solution, with *Q* being proportional to the species which contain both fluorescence tags (313).

3.3.2. FRET efficiency, *E*

This is one of the most prominent parameter which can be obtained through SMFS experiments in solution. The FRET efficiency is conventionally defined as the fraction of photons absorbed by the donor molecules decaying by the energy transfer process (359). For SMFS measurements, including PIE excitation, and taking into account the α, β, and γ factors (see section 3.3), the FRET efficiency, *E*, is defined as follows (290):

$$E = \frac{F_{FRET} - \alpha F_{647N} - \beta F_{A488}}{F_{FRET} - \alpha F_{647N} - \beta F_{A488} + \gamma F_{A488}} \qquad (3.2)$$

Importantly, our methodology contains time-resolved fluorescence information, so the fluorescence lifetime of each fluorophore can be estimated within individual bursts by means of the BIFL method. Hence, SMF correlograms of the FRET efficiency, *E* (calculated burstwise), *versus* the donor fluorescence lifetime, τ_{A488}, can be constructed using coincident fluorescence bursts as the burst picking criterion (Figure 3.4, panel A). These correlograms permit to distinguish different species by FRET efficiency and donor lifetime. Coincidence criterion and direct acceptor excitation in PIE mode ensure that only molecules with donor and acceptor are plotted in the correlogram, removing contributions of the ubiquitous 'zero-peak' present in conventional SMFS measurements (285).

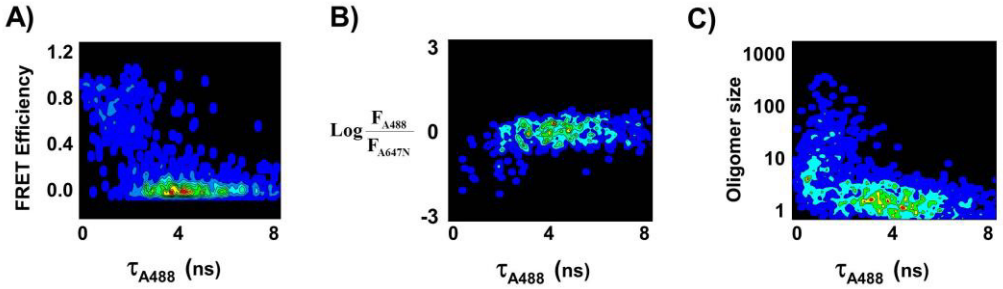


Figure 3.4 Examples of different types of single-molecule correlograms that can be obtained from SMFS-PIE experiments of incubated N47A-SH3-DA. **A:** FRET efficiency *versus* τ_{A488}; **B:** F_{A488}/F_{A647N} ratio (logarithmic scale) *versus* τ_{A488}; **C:** oligomer size *versus* τ_{A488}.

3.3.3. Fluorescence intensity ratio and oligomer size *versus* τ_{A488}

Another type of correlogram that allows to obtain structural information of the individual molecules from the fluorescence bursts is the plot of the F_{A488}/F_{A647N} ratio *versus* the fluorescence lifetime of the donor, τ_{A488}. Correlograms of F_{A488}/F_{A647N} *versus* τ_{A488} (Fig. 3.4, panel B) visualize the amount of oligomers formed at each concentration, the distribution of the relative stoichiometries, and the intra-oligomer FRET efficiency through the quenching effected over the donor's lifetime (see Chapter 5).

Moreover, the fluorescence intensities can provide additional information correlating the total fluorescence intensity (due to the fluorescence intensity of the directly excited donor and the acceptor) with respect to τ_{A488}. In previous literature using dual-color excitation schemes, the oligomer size of each individual event has been obtained as the total intensity of the burst, considering both channels, divided by the average fluorescence intensity of a monomer, according to the following (148, 360):

$$\text{Oligomer size} = \frac{F_{A488} + F_{A647N}}{F_{mon}} \qquad (3.3)$$

where F_{mon} is the value of the average intensity of a monomer, and F_{A488} and F_{A647N} are, respectively, the above defined background- and crosstalk-corrected donor and acceptor fluorescence intensity.

Nevertheless, the MFD-PIE approach has certain advantages: 1) because the two lasers are not simultaneously exciting the sample, but alternating in time, the direct excitation of the dyes is free from interactions from the other laser. For instance, it is known that a green laser produces switching off to different red dyes, such as Alexa Fluor 647 or Cy5 (361). 2) The fact that our technique has access to the dyes' fluorescence lifetime allows us to correct the total intensity of the dyes according to the degree of quenching that the dyes undergo. Indeed, processes like FRET or homo-quenching produce a decrease in the emitted fluorescence intensity and in the fluorescence lifetime. Hence, the oligomer size calculated without considering these processes will be underestimated, as the total intensity is lower than in the absence of FRET. Our approach permits to implement a method to obtain oligomer size, corrected by all the quenching processes (FRET and homo-quenching). To the best of our knowledge, it is the first time found in the literature that a correction like this is implemented for a better estimation of the oligomer size in SMFS experiments.

For all the oligomer bursts picked up, our analysis algorithm classifies the burst according to the τ_{A488} value, whether it is lower or not than the unquenched value of 4 ns, τ_0 (351). For molecules having $\tau_{A488} \geq \tau_0$, quenching and FRET processes are negligible, and consequently, the oligomer size is obtained through equation (3.3), because no other correction is necessary. For molecules exhibiting $\tau_{A488} \leq \tau_0$,

the overall intensity in the A488 channel is corrected by the amount of quenching, through the factor accounting for the lifetime ratio, τ_0/τ_{A488}.

$$\text{Oligomer size} = \left(\frac{F_{A488} \frac{\tau_0}{\tau_{A488}} + F_{A647N}}{F_{mon}} \right) \quad (3.4)$$

The estimation of the corrected oligomer size and the subsequent correlation with τ_{A488} (Figure 3.4, panel C) provides an advanced tool to link size information with FRET efficiency and structural arrangements.

3.4 FLIM-PIE

Fluorescence lifetime imaging microscopy (FLIM) is a multi-dimensional fluorescence microscopy technique that collects not only the fluorescence intensity information, but also the fluorescence lifetime in each pixel of a fluorescence image. The main advantages of using the fluorescence lifetime are that the fluorescence lifetime is independent of the local probe concentration. Thus, imprecision related to oscillations in the excitation light and heterogeneities in the optical properties of the medium can be overcome (359). FLIM is a powerful tool to directly probing aggregation and obtaining information about the properties of the labeled molecules.

Moreover, we employ herein a PIE excitation scheme using two overlapped alternating lasers (at 470 and at 633 nm). The dual-color, PIE scheme in FLIM applications provides ulterior advantages. Specifically, using FLIM-PIE, it is possible to excite the acceptor direct and independently of FRET processes, so that the contributions from incomplete FRET pairs with missing or non-fluorescing

acceptors is eliminated. As a result, we can perform, by a single FLIM-PIE experiment, a one-step reconstruction of three images: the donor dye FLIM image, FRET fluorescence image, and directly excited acceptor FLIM image, distinguishing molecules exhibiting genuine FRET from those ones which contain either donor only or acceptor only. In addition, FLIM-PIE provides lifetime and fluorescence intensity information of the labeled molecules and their colocalization. Definitely, FLIM-PIE experiments may provide important information about the features of the protein under study.

FLIM-PIE experiments were performed using the same MicroTime 200 fluorescence lifetime microscope system (PicoQuant, GmbH, Germany), described above for SMFS-PIE experiments. The microscope cover slides (Thermo Fisher Scientific, MENZEL-GLÄSER, Brawnschweig, Germany) were thoroughly cleaned using a protocol consisting of sonication for 20 min in 1M KOH, Ethanol (95% v/v) and high purity acetone, in several rounds, with every sonication step followed by thorough rinsing with deionized water (362). After the last step of sonication with deionized water, the cover slides were exposed to UV light for two hours (362) and finally stored in deionized water for a maximum of 10 days. Immediately prior to use, the slides were washed with ethanol and briefly air-dried.

To image a region, the laser beams were focused just above the surface to image the deposited samples on the slide and a wide area (between 10 and 15 μm^2) was raster-scanned with an x-y piezo-driven device (PI, Physik Instrumente, Germany) to set a region of interest (ROI). The excitation power was 6.64 μW for the 470 nm and 0.275 μW for the 635 nm laser. At least 10 images were collected from various areas for each sample with a 512 × 512 pixel resolution, a collection time of 0.60 ms/pixel and a time resolution of 116 ps per channel in the single-photon timing scale.

SymPhoTime software package was used for the analysis of the FLIM images with a similar protocol to that used in the SMFS traces, but in a pixel-by-pixel fashion. The PIE excitation scheme permitted that simultaneous imaging of two different fluorophores could be reconstructed. We performed a spatial pixel binning from 3×3 to 5×5 to the raw image to increase the number of photons per pixel, at the expense of a loss in the spatial resolution. By selecting the time windows and the detection channels for donor images and for FRET and acceptor images, the PIE scheme allowed one-step reconstruction of the donor dye FLIM image, FRET fluorescence image, and directly excited acceptor FLIM image (Figure 3.5). The selection of the molecules of interest was performed by visual inspection of those particles present in all three donor, FRET, and acceptor images (Figure 3.5). The FLIM measurement provides images with a decay curve in every pixel. For τ_{A488} measurements, the decay curves in the individual pixel were fitted with a single exponential decay function, applying the maximum likelihood estimator (MLE) as criterion for goodness of fit. For the deconvolution analysis, an instrument response function (IRF) was reconstructed from the decay of a FLIM measurement performed at high concentration of sample (from 0.5 to 2 nM), containing the same number of time channels as the data set to be analyzed. To discard the background fluorescence, we applied to the FLIM image an intensity threshold of 5 minimum photon number for pixels.

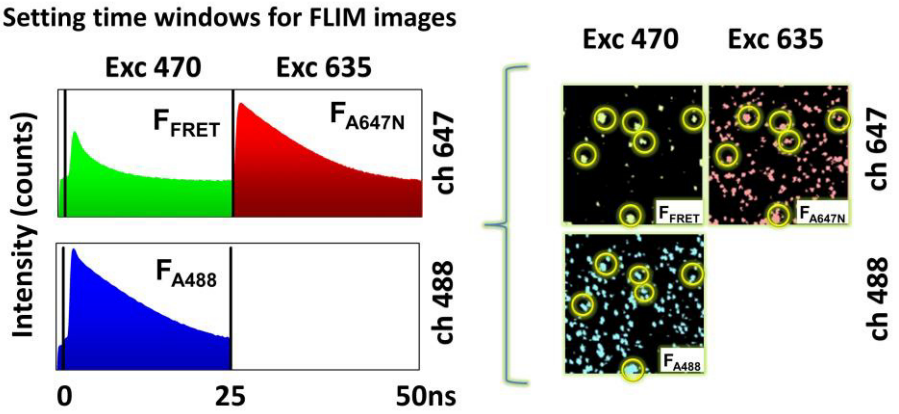


Figure 3.5 Scheme of FLIM-PIE experiments. The PIE scheme in FLIM permits the acquisition of three simultaneous images from a single scan.

Fiji is just ImageJ macro was used as an alternative analysis for the kinetic studies performed using FLIM data (363). In detail, after exporting the output matrix data of the FLIM analysis, it was realized a gaussian function for smoothing (standard deviation 1, in pixels). Then, the images were semi-automatically segmented on the basis of pixel intensity using isodata algorithm ('getAutoThreshold' method or 'make binary' function in Fiji and binary masks created, that is, 0 for background and 1 for oligomeric species). Multiplying each intensity masked channels we obtained a final image, which only contains values for coincident events. Finally, we multiplied the final masked images by the τ_{A488} to obtain a final image where each intensity-pixel is associated with the corresponding τ_{A488} (Figure 3.6).

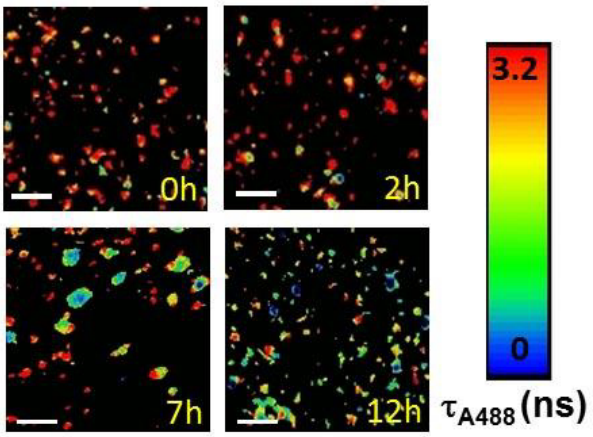


Figure 3.6 Examples of A488 FLIM images of mixture of N47A-SH3-A488 and N47A-SH3-A647N (1:1 ratio) incubated at 32 μM, obtained using *Fiji is just ImageJ* macro. The images are colored using a pseudo-color scale from 0 to 3.2 ns. White scale bars represent 2.4 μm.

3.5 DYNAMIC LIGHT SCATTERING (DLS)

Dynamic light scattering (DLS) provides information about the molecular size of the elements present in a small volume of a particle solution. This method exploits the high dependence of the intensity of scattered light with the particle size. The Brownian motions of the particles in solution produce variations in the distance between them, so that the intensity of the light dispersed is not constant but rapidly varies with time. The interferences of the scattered light by neighbor particles produce fluctuations of the intensity in the plane of detection related with these movements. The analysis of the association between the intensity fluctuations and time provides information on the translational diffusion coefficient of the particles (*D*), considering them as hypothetical spheres. The diffusion coefficient of the molecules in the sample cell can be determined from the decay of the intensity autocorrelation data provided by the DLS instrument as a correlation function of

the dispersed light intensity during a period of time. The latter is calculated by measuring the similarity between a signal acquired at a given time and after a delay (157, 364) The apparent hydrodynamic radius (R_H) of the particles is then derived from the translational diffusion coefficient using the Stokes-Einstein equation, as follows (Equation 3.5):

$$D = \frac{k_B \times T}{6 \times \pi \times \eta \times R_H} \qquad (3.5)$$

where k_B is the Boltzmann constant, T is the temperature in Kelvin, η is the viscosity of the solvent.

For our DLS measurements, we used a DynaPro MS-X instrument (Wyatt Technology Corporation, Santa Barbara, CA, USA) equipped with a semi-conductor laser of ~ 830 nm wavelength and a Peltier-based temperature control. The experiments were made using a 30 µL quartz sample cuvette. Prior to measurements, protein samples were dissolved into a buffer filtered through 0.02 µm Anotop® filters (Whatman plc, UK) and the sample solutions were then centrifuged in order to remove any aggregates and dust from the sample. To avoid possible interactions between cysteine residues of unlabeled samples, we added DTT to the buffer solution used for N47A-SH3 samples up to a final concentration of 5 mM. Dynamics software (Dynamics® V6, Wyatt Technology Corporation, Santa Barbara, CA, USA) was used to both record and process DLS data, obtaining information on hydrodynamic radius, scattering intensity and mass percentage of the samples.

3.6 CIRCULAR DICHROISM

Circular dichroism (CD) is a spectroscopy technique which is employed in the study of organic molecules with optical activity. When applied to proteins and nucleic acids in solution, this technique provides information on the secondary and tertiary structure elements, and hence, it is useful to study conformational changes (150, 365, 366). The CD technique is based on the differential absorption of right and left circularly polarized light from chiral molecules. The polarized light components interact differently with the chiral center of chromophores which either possess intrinsic chirality or are placed in chiral environments. Indeed, proteins possess a number of chromophores which can contribute to CD signal. As a consequence of the light/matter interaction, matter absorbs part of the light and, depending on the extinction coefficient of both right and left circularly polarized light, the amplitude of the light decreases. This phenomenon results in a difference in amplitude and phase of each component of the polarized light, so that the sum of the vectors of the two components rotates following a helical path proportional to the difference between the absorbance of the sample of right- and left-handed circularly polarized light (ΔAbs, λ). The ΔAbs in proteins is measured as ellipticity, θ, in units of degree. The normalized CD signal is usually expressed as *mean residue molar ellipticity*, $[\theta]_{MRW}$, in units of $deg \times cm^2 \times dmol^{-1}$, whose value is related to the chirality of the molecule as well as to its structure. The $[\theta]_{MRW}$ at a given wavelength λ, is calculated by the equation 3.6:

$$[\theta]_{MRW} = \frac{MRW \times \theta}{10 \times d \times C} \qquad (3.6)$$

where θ is the observed ellipticity, d is the pathlenght and C is the sample concentration. MRW is the mean residue weight, which is obtained by dividing the molecular mass (MW) by the number of residues in the polypeptide chain.

The CD bands observed in proteins are essentially situated in two regions of the electromagnetic spectrum. The far-UV region, between 170 and 250 nm, in which most contribution absorption bands come from the peptide bonds, gives information on the secondary structure elements in proteins: α-helix, β-sheet, β-turns, random coil (Figure 3.7) (149). More specifically, α-helices show a negative band at 222 nm whereas a negative band at about 215 nm and a positive band at about 198 nm is typical of β-strands.(151, 367, 368). The near-UV region, between 250 and 320 nm, where the observed band arises from the aromatic side chains of certain amino acids, provides information about changes in the tertiary structure of the protein. Therefore, each region gives different and complementary information related to protein structure.

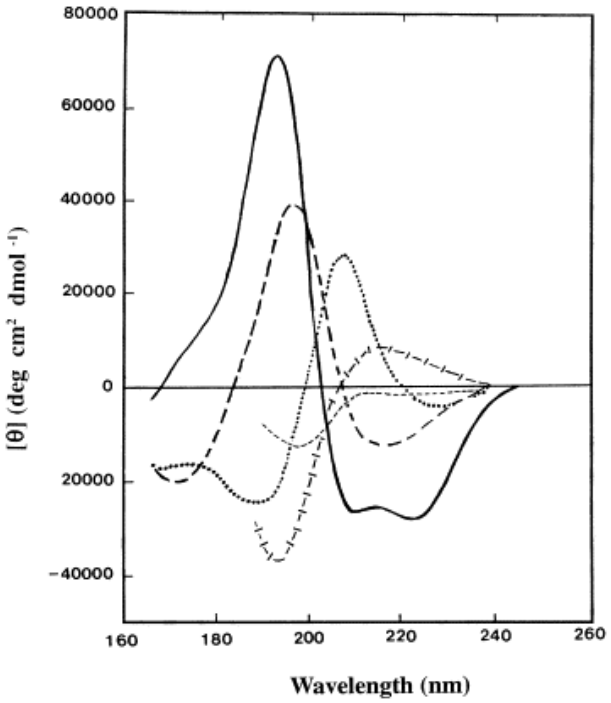


Figure 3.7 Far UV CD spectra associated with various types of secondary structure. Solid line, α-helix; long dashed line, anti-parallel β-sheet; dotted line, type I β-turn; cross dashed line, extended 31-helix or poly (Pro) II helix; short dashed line, irregular. Reprinted with permission from Kelly, S. M., T. J. Jess, et al. (2005), *Biochim Biophys Acta* **1751**: 119-139 (2005) with permission from Elsevier.

Far-UV and Near-UV CD spectra of N47A-SH3-DA and N47A-SH3 were collected using a JASCO J-715 (JASCO, Japan) spectropolarimeter, equipped with a Peltier-driven temperature control. A nitrogen flow of 5 L/min was used to purge and refrigerate the system. Spectra were collected using the following parameters: 1 mm quartz cuvette; wavelength, 260 nm to 200 nm; response time, 1 sec; scan speed, 100 nm min^{-1}; bandwidth, 1 nm. With the buffer solution containing chloride ions (which absorb strongly below 195 nm), the CD data acquisition was truncated at 200 nm. Unfortunately, the range of wavelength scanned limits the use of secondary structure prediction software (DICHROWEB (369) and SELCON (370), i.e.). Even so, we could interpret far-CD data using the CDNN algorithm with a reliable estimation (371). This tool works comparing experimental far-CD spectra with a reference spectra database of known globular proteins and gives the result output in terms of content of secondary structure (α-helix, β-sheets, turns and random coils) in an ordered table.

Near-UV CD spectra were collected as follows: 5 mm quartz cuvette; wavelength, 350 nm to 250 nm; response time, 1 sec; scan speed, 100 nm min^{-1}; bandwidth, 1 nm. The resulting spectrum was the average of the numbers of the scans, adjusted to the concentration of the sample.

For kinetic experiments, 300 µL of freshly prepared protein sample were loaded in the cuvette and incubated at constant temperature (37 ºC). The CD signal was measured at 215 nm (201, 372) as a function of time with 20 seconds of step resolution, 4 seconds of response, 1.0 nm of bandwidth**,** 100 nm min^{-1} scan speed, and one single accumulation. The acquired data were subtracted from a baseline scan of the buffer, which were obtained under identical experimental conditions. For the unlabeled N47A-SH3 experiments, we added TCEP 0.1 M to the buffer to avoid free cysteine-terminal interactions.

3.7 TRANSMISSION ELECTRON MICROSCOPY (TEM)

TEM techniques are widely used to obtain morphological information about the aggregates formed during the amyloid fibrillation process. This method uses an electron beam to visualize entities with sizes in the range of nm to µm.

TEM measurements of protein samples were performed in a Libra 120 Plus transmission electron microscope (Carl Zeiss SMT, Germany) operated at 120 kV and equipped with a LaB6 filament and an SSCCD 2k × 2k direct coupling camera. To prepare samples for TEM measurements, the proteins were diluted in the buffer and incubated at 37 ºC. During the aggregation process, two aliquots of 10 µL were taken from the aggregation sample and rapidly frozen in liquid nitrogen to stop aggregation. Before to measurements, the samples were deposited on Formvar 300-mesh copper grids (ANAME, Madrid, Spain) and left incubate for 5 minutes. One of the aliquots were diluted 2-fold with the buffer. The copper grids were washed twice with deionized water and negatively stained during 1 minute using 1% (w/v) uranyl acetate solution. Excess stain was removed with a tissue paper and the samples were finally dried out and observed under the electron microscope.

3.8 DIFFERENTIAL SCANNING CALORIMETRY (DSC)

Differential scanning calorimetry (DSC) is a powerful technique that can be used to obtain information about thermal stability and energetic changes occurred in unfolding and folding processes (373, 374). DSC measures the heat capacity of the protein as a function of temperature, constituting a direct method to obtain the enthalpy of the process and all thermodynamic parameters associated with the transition. Actually, DSC is the only technique capable of a complete thermodynamic characterization of the folding-unfolding equilibrium of a protein. The outputs of DSC experiments are usually thermograms in which the heat absorption peaks correspond to endothermic processes induced by temperature increments. The analysis of these thermograms provides the thermodynamic parameters (enthalpy, entropy, Gibbs energy, heat capacity) associated with the transition for reversible and not scan rate dependent processes (375-378).

DSC experiments were performed using an automatic VP-DSC capillary cell microcalorimeter from MicroCal (Northampton, MA, USA) (373). This calorimeter works adiabatically and monitors the difference in heat capacity between a reference cell filled with solvent and an almost identical cell containing the sample in study (the cell operating volume is 0.134 mL). The range of the temperature scans used for thermally induced processes was 5 ºC to 110 ºC. Cells were kept under a negative pressure of 60 psi to avoid degassing during the scan and boiling at high temperatures. Protein solutions were prepared for the experiments by extensive dialysis against a buffer containing 0.1 M Gly, 0.1 M NaCl, and pH 3.2, at 4 º C. The buffer of the last dialysis step was used in the reference cell of the calorimeter. All samples were centrifuged for 10 min at 14000 rpm in a table-top microcentrifuge. Before each DSC run, four baselines were registered using buffer solution in both cells to ensure a correct equilibration of the instrument, whereas the reversibility of the thermal process was determined

through a second consecutive reheating run of the protein sample.

The difference in heat capacity between the solution containing the protein and the solvent measured by DSC permits to obtain the protein partial heat capacity, as follows (378):

$$\Delta C_{p,ap}(T) = C_{p,ap}(T) \times m_p - C_{p,s}(T) \times \Delta m_s \qquad (3.7)$$

where $C_{p,ap}$ (T) and $C_{p,s}$ (T) are, respectively, the partial specific heat capacity of the protein and of the solvent at temperature T, m_p is the mass of protein which is in the calorimetric cell, and Δm_s is the mass of the solvent displaced by proteins in solution.

This difference in heat capacity is normally negative, taking into account that the heat capacity of the solution is smaller than the heat capacity of the same volume of solvent (378), so that the heat capacity of macromolecules in solution is smaller than that of the same volume of the solvent.

Moreover, considering that Δm_s is equal to:

$$\Delta m_s = m_p \frac{V_p(T)}{V_s(T)} \qquad (3.8)$$

where $V_p(T)$ and $V_s(T)$ are, respectively, the partial specific volume of the protein and of the solvent at temperature T, from equation (3.7) and (3.8) we obtain the partial specific heat capacity of protein:

$$C_{p,ap}(T) = C_{p,s}(T) \times \frac{V_p(T)}{V_s(T)} + \frac{\Delta C_{p,ap}(T)}{m_p} \qquad (3.9)$$

Since we usually work in aqueous solutions and the specific volume and the heat capacity of pure water are known with high accuracy, and assuming the average V_p = 0.73 $mL \times g^{-1}$ (the average value for *globular proteins*) for the protein under

study (379), we can obtain from the equation 3.9 the partial specific heat capacity of protein in solution.

3.9 CROSSLINKING

Chemical crosslinking involves the formations of covalent bonds between two or more molecules by using bifunctional reagents. These crosslinking reagents contain reactive ends to specific functional groups (primary amines, sulfhydryls, carbonyls, carboxylic acids etc.) present on proteins or other molecules. The availability in proteins and peptides of these functional groups that are commonly targeted for bioconjugation, make them targets for conjugation and for study using crosslinking methods. Chemical crosslinking can be used as a strategy to study protein-protein interactions *in vitro*. As two proteins physically interact with each other, they can be covalently cross-linked. The formation of crosslinks between two distinct proteins is a direct and convincing evidence of their close proximity. Before starting crosslinking experiments, it should be selected the adequate crosslinker to ensure an optimal reaction with the protein, considering features as solubility, spacer arm length, reversibility of cross-links. There are a variety of crosslinkers reagents, either homo- or hetero-bifunctional, depending on the fact that their reactive groups are identical or not. The homo-bifunctional reagents specifically react with primary amine groups and have been used extensively as they are soluble in aqueous solvents and can form stable inter- and intra-subunit covalent bonds (380). Among these, the treatment with glutaraldehyde is particularly useful to obtain preliminary information on the quaternary association of proteins (381, 382). This crosslinker should be used in buffers free from amines, and permit, once proteins are crosslinked, the direct detection of crosslinked products by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) (383).

We have applied cross-linking method to capture and analyze the interactions between protein samples, in order to study their oligomerization state. For the cross-linking experiments, the samples were dissolved in the buffer and incubated at the temperature under study (20, 40, 60 or 70 ºC) for 3 min in the presence of 25% (w/v) glutaraldehyde. Later, 2 M $NaBH_4$ in 0.1 M NaOH was added to the crosslinked samples and left incubate for 10 minutes more. Finally, sodium deoxycholate was added to the sample up to a final concentration in solution of 0.02%. Protein samples were precipitated with an aqueous 78% (w/v) trichloroacetic acid (TCA) solution and later washed thoroughly with deionized water to remove any remaining TCA in the samples. Finally, the samples were boiled in Laemmli sample buffer and analyzed by SDS-PAGE under reducing conditions.

4. COMBINING CONVENTIONAL AND SMF TECHNIQUES TO STUDY N47A-SH3 OLIGOMERIZATION

4. COMBINING CONVENTIONAL AND SMF TECHNIQUES TO STUDY N47A-SH3 OLIGOMERIZATION

4.1. LABELED AND UNLABELED N47A-SH3: INITIAL CONSIDERATIONS

Our research group has previously demonstrated that the N47A mutant of the Spc-SH3 domain (N47A Spc-SH3) form curly amyloid fibrils under mild acidic conditions (334, 338), and that experimental conditions, such as temperature, pH, and concentration of proteins and salts, affect the kinetics of nucleation considerably (201). Recently, Ruzafa and colleagues used a widely range of biophysical techniques to analyze quantitatively the formation of the N47A Spc-SH3 amyloid nuclei under conditions that permit to avoid a lag phase (339). However, the techniques used in these studies are not able to observe the precursor oligomers that have been identified playing a key role in the mechanism of aggregation (74, 129, 199). The information provided by previous investigations on N47A Spc-SH3 is therefore incomplete, since they lack details on the initial oligomers formed during the microscopic process preceding the formation of aggregates. The sensitivity of the SMF methodologies permits to work at lower concentration conditions than those used in conventional biophysical experiments, making it possible to detect individual molecules, including heterogeneous, transient, rare intermediate species. The use of fluorescence-based techniques requires the incorporation of bright and photostable dyes into the protein. In a previous study, we have used FLCS-PIE, an advanced modification of conventional FCS, to follow the oligomer growth of the N47A Spc-SH3 (340), with the protein being labeled with either Alexa Fluor 488 or Alexa Fluor 647

fluorophore. The results of this study have been thoroughly described in the introduction. To label the protein, the N47A Spc-SH3 was engineered by introducing a six-residue tag, Gly-Ser-Gly-Ser-Gly-Cys, with the Cys residue at the C-terminus. This mutant, including the six-residue tag, is what we define throughout this work as N47A-SH3, and we use it as a model system.

Here we go a step beyond previous studies by investigating the oligomeric species involved in aggregation and extending our knowledge to the initial phase of the aggregation of N47A-SH3. To do this, we have used a combination of the SMF-PIE technique, fluorescence imaging tools, and conventional biophysical methodologies. Notably, SMF-PIE methods provide the best results for the determination of FRET efficiency, because it is a multidimensional technique, which adds up orthogonal information on the oligomers. In order to apply the SMF-PIE method, the N47A-SH3 was labeled with either a donor or an acceptor fluorophore, so that intra-oligomer FRET can be detected upon interaction of individual monomers into oligomers comprising both dyes. We use a green (ATTO 488 maleimide, A488, the donor) or a red (ATTO 647N maleimide, A647N, the acceptor) fluorophores, which are suitable for labeling thiol groups of proteins. These small fluorophores (with MW between 700 and 800 Da) possess excellent water solubility and exceptional thermal and photo-stability, which make them suitable to study the N47A-SH3 model system under a wide range of concentration, pH, temperature and salt conditions. From now on, we employ the following nomenclature: N47A-SH3-D and N47A-SH3-A to refer, respectively, to the N47A-SH3 labeled with A488 and A647N, whereas N47A-SH3-DA indicates a 1:1 ratio mixture of N47A-SH3-D and N47A-SH3-A. The experimental conditions that we have used to study the N47A-SH3 aggregation are those previously studied to induce N47A Spc-SH3 mutant aggregation *in vitro* (334, 338, 339). These include a buffer solution containing 0.10 M Gly and 0.10 M NaCl at pH 3.2 (from now on, ‘the aggregation buffer’), at 37 °C. In this work, we

have also explored the effect of NaCl concentration on the aggregation, so that this concentration was varied where noted.

4.2 EFFECTS OF THE CHEMICAL MODIFICATIONS ON N47A-SH3 STRUCTURE

Several biophysical techniques were employed in order to verify whether the fluorescent tags somehow altered the structural integrity of N47A-SH3. Firstly, we carried out far- and near-UV CD experiments of N47A-SH3, N47A-SH3-D and N47A-SH3-A samples, diluted in the aggregation buffer. The CD signal could only be measured from 200 nm, due to saturation of the instrumental signal at shorter wavelengths. Both far-UV and near-UV CD spectra were similar for all samples tested, regardless the presence/absence of the fluorophore (Figure 4.1). In addition, from the far-UV CD data we were able to determine the content of secondary structure in each construct using the CDNN software (384) (Table 4.1).

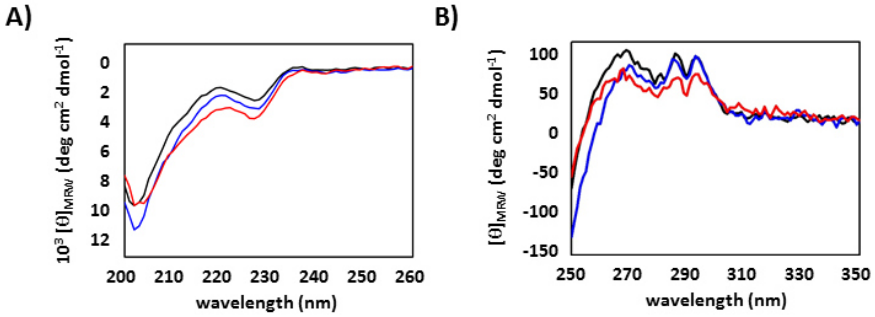


Figure 4.1 A: Far-UV CD spectra of N47A-SH3-D (blue line), N47A-SH3-A (red line) and N47A-SH3 (black line). The protein samples were diluted down to a final concentration of 15 µM in the aggregation buffer. **B:** Near-UV CD spectra of N47A-SH3-D (blue line), N47A-SH3-A (red line) and unlabeled N47A-SH3 (black line), diluted in the aggregation buffer down to a final concentration, respectively, of 30, 25 and 35 µM.

Table 4.1 Percentage of secondary structure of labeled and unlabeled N47A-SH3, as calculated from CD spectral deconvolution using CDNN software.

2nd Structure	α	β	Random
N47A-SH3	15.9	32.3	51.8
N47A-SH3-D	17.3	31.1	51.4
N47A-SH3-A	19.1	28.6	52.3

The results indicate that while the far-UV CD spectrum of N47A-SH3-D was nearly identical to that of N47A-SH3, N47A-SH3-A exhibited a slightly different spectrum, resulting mostly likely from the chiral morphology of the A647N dye itself. To endorse this assertion, we measured the CD spectra of a double-stranded DNA and the same double-stranded DNA labeled with the A647N dye (Figure 4.2). The latter exhibited a more negative ellipticity in the 210-240 nm region, which is similar to that observed in the far-UV CD spectrum of N47A-SH3-A. Therefore, being the A647N dye being optically active in this region, it may be responsible for the minor signal decrease that we observe in the far-UV CD spectrum of N47A-SH3-A construct.

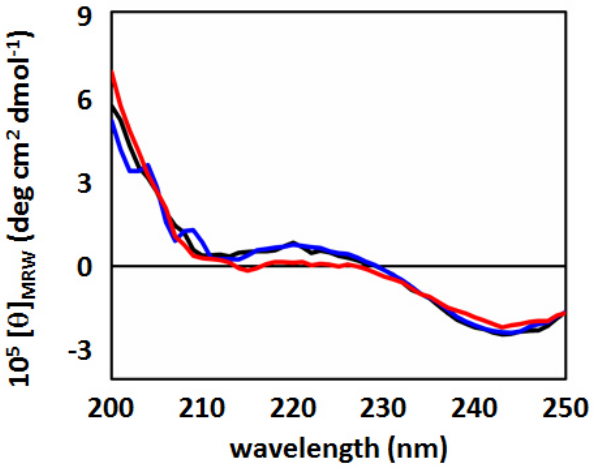


Figure 4.2 Far-UV CD spectra of A488- (blue line) and A647N-labeled double-stranded DNA (red line) and unlabeled DNA (black line).

Considering the contribution to the chirality of the A647N, we may justify the slight difference in the far-UV CD spectrum of the N47A-SH3-A compared with the N47A-SH3 and N47A-SH3-D far-UV CD spectra. A simple interpretation of these results suggests that the fluorophores do not vary the structural integrity of the N47A-SH3 construct.

4.3 EFFECTS OF THE CHEMICAL MODIFICATIONS ON THE N47A-SH3 AGGREGATION

In order to study the effect of the modifications carried out onto the N47A mutant (including those coming from the six-residue tag and the fluorescent label) on the capability of the protein to form amyloid fibrils, we observed the morphology of the aggregates formed under identical conditions of temperature and NaCl concentration. We incubated samples of N47A-SH3 and N47A-SH3-DA at 37 °C for 15 days, diluting them in the aggregation buffer down to a final concentration

of 120 and 32 μM, respectively. This difference on concentrations was due to the roughly 4-fold higher rate of aggregation observed for the labeled protein compared to that observed for the unlabeled. The enhancement of the aggregation rate caused by the fluorophores is driven by the aromaticity introduced by the dyes' moieties. Previous studies on the role of the aromatic amino acids on amyloid formation showed a strong enhancement of the process caused by the aromatic amino acid variant of the protein under study. This effect can be attributed to the high hydrophobicity, β-sheet propensity and the planar steric profile of the aromatic side-chain groups **(385)**.

Both the six-residue tag and the labeling, by visual inspection of the TEM images, seem not to alter the capability of the protein to aggregate and form amyloid fibrils (Figure 4.3). Moreover, we cannot find any significant difference on the shape or length of the fibrils formed with labeled or unlabeled protein. Comparing to the TEM images previously reported by Varela and colleagues (334), the fibrils of N47A-SH3 and N47A-SH3-DA are similar than those of N47A Spc-SH3 after 180 minutes of incubation at 37 °C. The reason for this difference in temporality (180 min for N47A Spc-SH3 *versus* 15 days for N47A-SH3 and N47A-SH3-DA) lies in the fact that we have worked with both N47A-SH3 and N47A-SH3-DA at much lower concentrations than Varela and colleagues did (they used 8 mg/mL of protein). The shape of the fibrils of both labeled and unlabeled sample is, in any case, curly. This shape is typical for the N47A Spc-SH3 amyloid fibrils when heterogeneous nucleation is the main mechanism (339), and this is the case for incubations at low concentration of proteins.

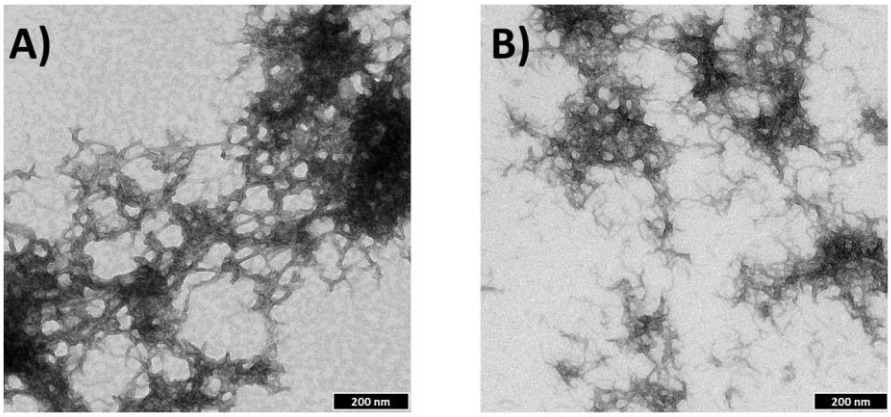


Figure 4.3 TEM images of samples of N47A-SH3 at the total protein concentration 120 μM (**A**) and N47A-SH3-DA at the total protein concentration 32 μM (**B**), after 45 days of incubation at 37 °C in the aggregation buffer. Black scale bar corresponds to 200 nm.

4.4. THE EFFECT OF THE TEMPERATURE ON THE N47A-SH3 UNDER CONDITIONS OF EQUILIBRIUM

In order to investigate the thermal stability of the N47A-SH3 constructs, we have performed several experiments at different temperatures, using a variety of biophysical techniques, such as far-UV CD, differential scanning calorimetry (DSC), and chemical crosslinking.

4.4.1 Far-UV CD

We have also used far-UV CD to estimate changes in secondary structure upon heating of N47A-SH3, N47A-SH3-D and N47A-SH3-A. Fig. 4.4 shows far-UV CD spectra measured at temperatures from 25 to 70 °C at 5 °C intervals. The behavior of all the three protein constructs practically coincide, showing a major

change in the spectrum between 40 °C and 70 °C, as the ellipticity between 210 and 230 nm decreases, indicating continuous structural changes in the protein molecule with heating. Judging from far-UV CD data, at the temperatures under study, the labeled protein shows the same changes in secondary structure with temperature as the N47A-SH3 does.

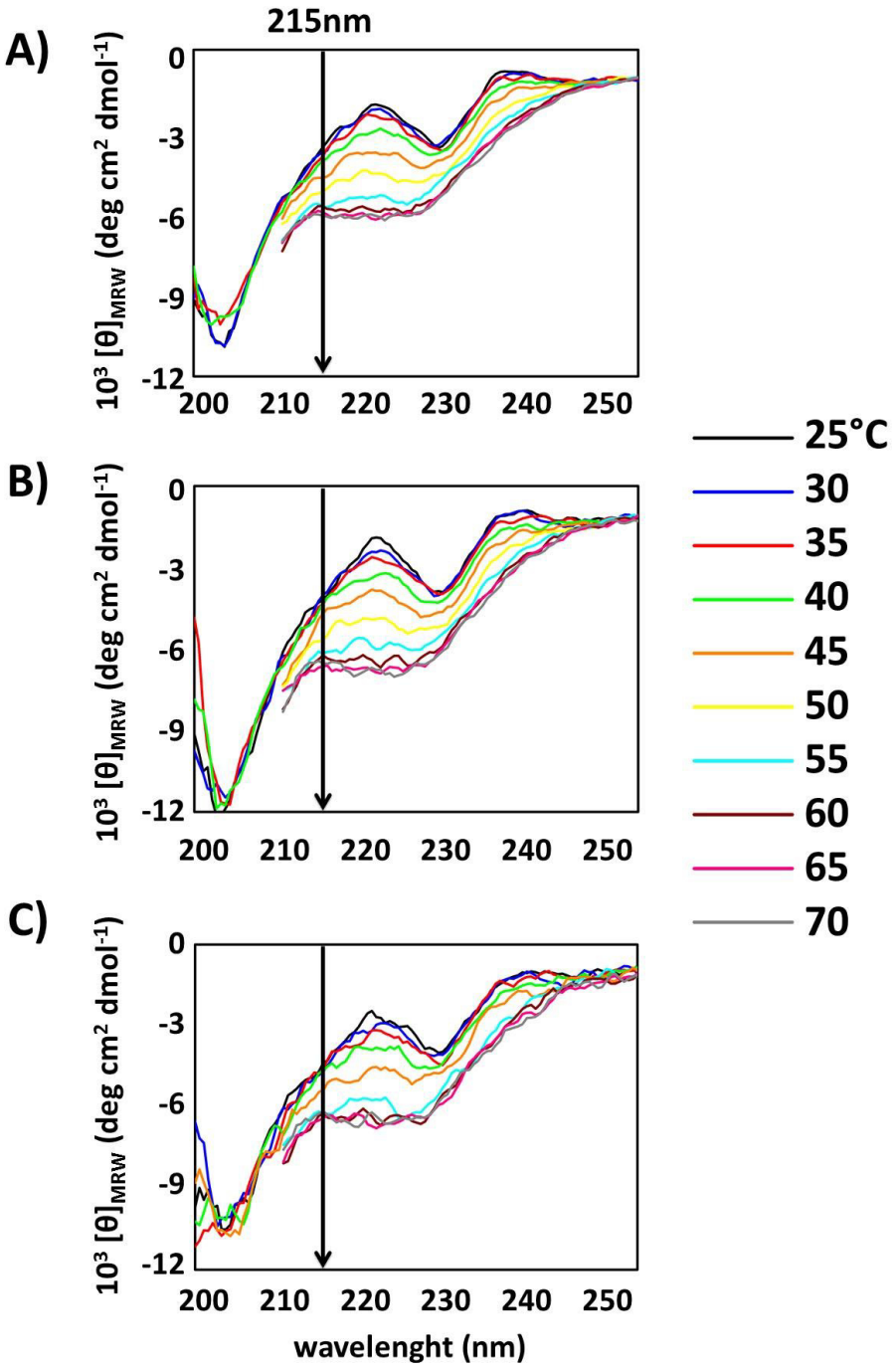


Figure 4.4 Far-UV CD temperature scans of N47A-SH3 (**A**), N47A-SH3-D (**B**) and N47A-SH3-A (**C**). samples.

4.4.2 DSC

To further investigate the role of the temperature, we carried out DSC experiments of N47A-SH3, N47A-SH3-D and N47A-SH3-A under conditions of formation of amyloid aggregates. The protein samples were diluted in the aggregation buffer down to a final concentration of 120 μM, and the temperature scan was set from 3 to 100 °C and at a scan rate of 1.5 K/min.

For the N47A-SH3, the DSC thermograms present a single transition with a maximum at ≈ 50 °C (Figure 4.5) and high reversibility in a second consecutive scan (data not shown). Comparing with the previous studies of Morel and colleagues on N47A Spc-SH3, performed at a protein concentration ranging between 120 and 450 μM (338), the N47A-SH3 at 120 μM exhibits a similar behavior, showing no evidence of protein aggregation even after heating at high temperatures. On the contrary, the DSC thermograms of both N47A-SH3-D and N47A-SH3-A (blue and red lines, in Figure 4.5) exhibited complicated shapes with additional endothermic peaks around 50, 70 and 80 °C. Comparing with the work of Morel and colleagues, the maximum of these peaks are similar to those shown by N47A Spc-SH3 at a concentration of 600 μM. In the previous study on N47A Spc-SH3, these different thermal transitions were attributed, among other reasons, to intermediate oligomers or to the formation of higher aggregates (338).

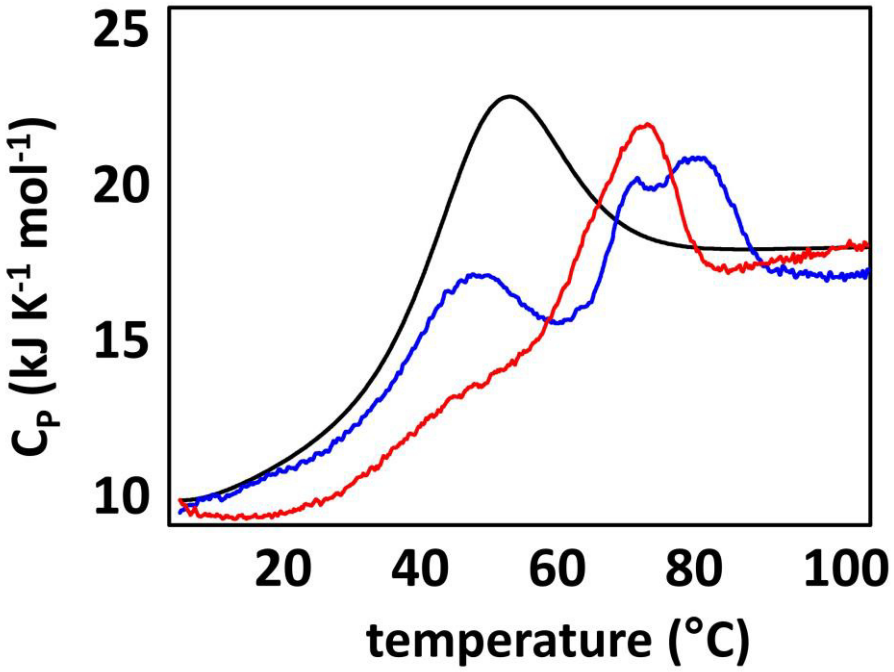


Figure 4.5 DSC thermograms of N47A-SH3 (black line), N47A-SH3-D (blue line), and N47A-SH3-A (red line).

4.4.3 Cross-linking

This test permits to catch labile oligomeric species present in a protein sample. We employed a cross-linking agent (glutaraldehyde 25% (w/v) to detect any oligomer present in N47A-SH3, N47A-SH3-D and N47A-SH3-A samples when protein concentration is kept at 60 μM in the aggregation buffer.

The experiments were made at 20 °C, at which the N47A-SH3 is in its native state (Figure 4.5), at 40 °C, a pre-transition temperature, at which the N47A-SH3 may still in the native state, and it is a temperature roughly close to the physiological (37 °C) temperatures. Besides, we also carried out this experiment at 60 and 70 °C, that are, respectively, high temperatures after the maximum (Tm = 51 °C) of the peak in the N47A-SH3 DSC curve, and at which the N47A-SH3 is all practically

unfolded., . These experiments revealed a ladder of species in which the main band corresponded to the monomeric form, with higher-order species also detected (Figure 4.6A and B). These oligomers were detected in all samples tested, with or without a fluorescent probe attached, clearly indicating a genuine oligomerization process of the protein. Interestingly, the presence of SDS resulted in the disaggregation of the oligomers when the cross-linking agent was not used. This result suggests that these associations are relatively weak and certainly non-covalent in nature. As temperature increases, we observe a shift of the band correspondent to monomers towards higher MW and, concomitantly, the formation of higher-order species (roughly of 21 kDa). At the highest temperature investigated, the equilibrium is populated by species of different sizes that can be assigned to monomers, dimers, larger oligomers and higher-order aggregates. These experiments demonstrate the existence of a genuine oligomerization of the N47A-SH3 domain, either labeled or not, which seems to be enhanced as the incubation temperature increases. In order to obtain a better level of details on the species observed in the context of cross-linking experiments, they will be studied by techniques that possess better resolution than SDS-PAGE gels have, as we will discuss in the following sections.

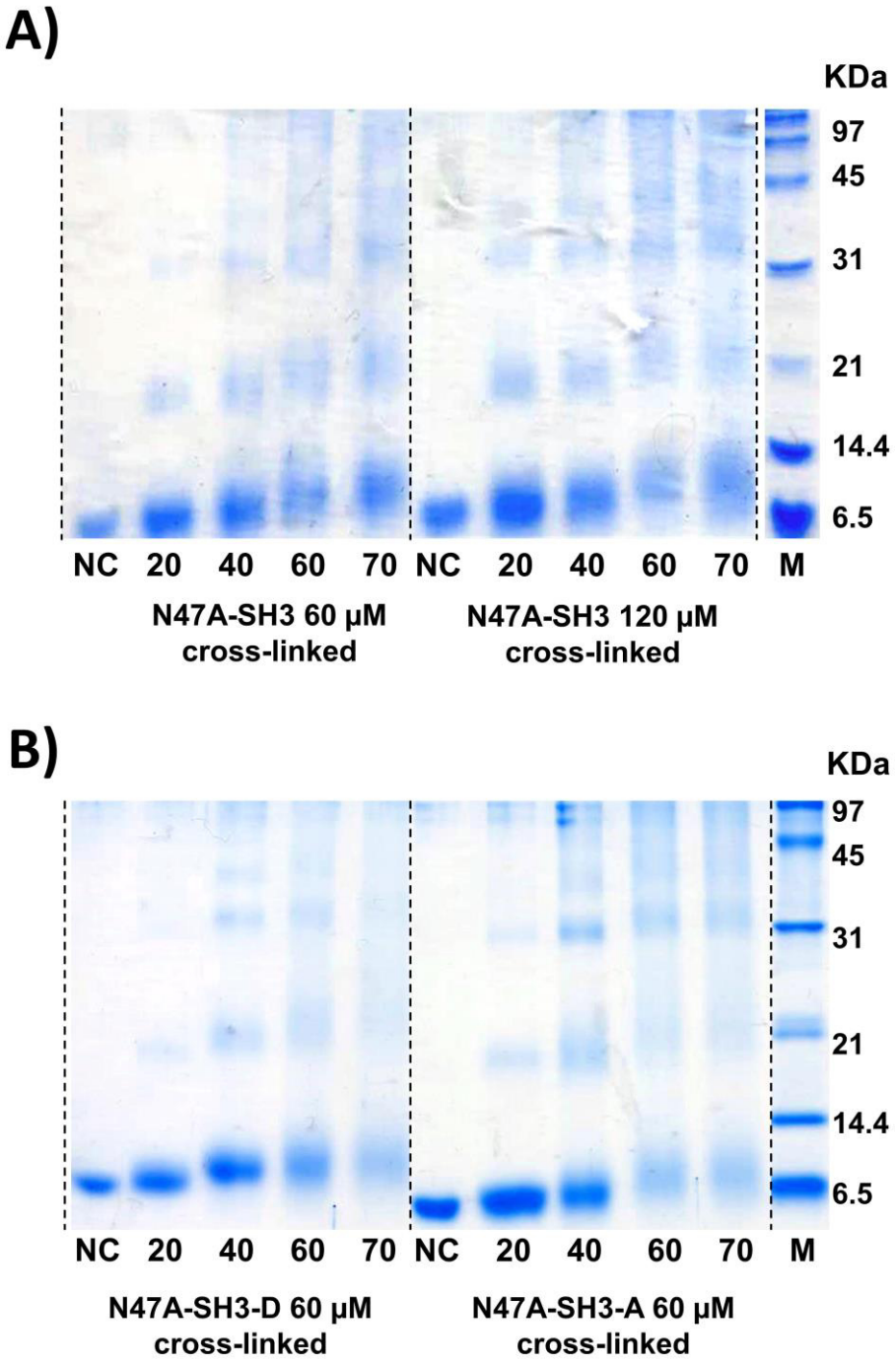


Figure 4.7 A, B: SDS-PAGE gels of the N47ASH3, N47A-SH3-D and N47A-SH3-A submitted to cross-linking with glutaraldehyde at different

temperature. **A:** N47A-SH3 samples diluted, respectively, to 60 and 120 μM in the aggregation buffer. **B:** N47ASH3-D and N47A-SH3-A samples diluted to 60 μM in the aggregation buffer. In both A and B gels the samples were crosslinked at 20, 40, 60 and 70 °C, lane NC corresponds to not cross-linked samples and lane M corresponds to molecular weight markers.

4.5. OPTIMIZATION OF THE EXPERIMENTAL CONDITIONS FOR SMF-PIE EXPERIMENTS

The formation of the oligomers of N47A-SH3-DA was analyzed using SMF-PIE, that permits to extract a variety of parameters, such as the fluorescence lifetime of individual molecules, τ, or burstwise FRET efficiency, *E*, with a much higher reliability than conventional single-molecule FRET experiments (313, 358). Interestingly, these parameters can be correlated for each individual molecule event, allowing building up multi-dimensional population histograms. The study of oligomer formation required focusing only on the aggregated species, those containing at least one N47A-SH3-D unit and one N47A-SH3-A unit, and allows detecting individual oligomers even in the presence of a large excess of monomers (358, 386). Moreover, the close proximity of the A488 and A647N fluorophores make it feasible to undergo intra-oligomer FRET, which is also accurately detected in SMF-PIE.

4.5.1 Detection of N47A-SH3-DA oligomers by FRET efficiency

In order to detect the amyloidogenic oligomers which are formed during the aggregation process of N47A-SH3 using an advanced SMF approach, stoichiometric amounts of the labeled constructs (N47A-SH3-DA, meaning 50%

of N47A-SH3-D and 50% of N47A-SH3-A) were incubated to form amyloid fibrils at 37 °C for 3 days, after carefully diluting the sample in the aggregation buffer down to a micro-molar concentration (16, 20, 24, 28, 32 and 44 μM). At concentrations above 24 μM, the appearance of an insoluble deposit suggests the formation of fibers. While incubating, 2 □L aliquots were taken out of the sample at different times. These aliquots were diluted 20000 fold, down to single molecule regime and analyzed using SMF-PIE according to the procedure described in the Materials and Methods chapter (see 3.3).

In a first analysis approach, the experimental data were represented in correlograms of *E versus* τ_{A488}, where the latter entails an orthogonal estimation of the *E* through the quenching caused to the donor (A488) fluorophore. These correlograms exhibited the presence of a population of soluble low-FRET oligomers at early times, even in the absence of incubation, and even when the samples were mixed at low concentrations (Figure 4.8). Interestingly, a distinct second population of oligomers arose with the incubation time, displaying higher values of *E*. The latter is related to the conformation and the oligomer size, and hence, these results suggest a gradual oligomeric structure organization and compactness with incubation time. The *E versus* τ_{A488} correlograms of N47A-SH3-DA after 8 hours of incubation clearly showed two different populations of oligomers, low-FRET and high-FRET. The second exhibited a prominent shift of the D-A distance (*E* value distribution centred at 0.6) and of the correspondent τ_{A488} value (centred at ≈ 1.5 ns).

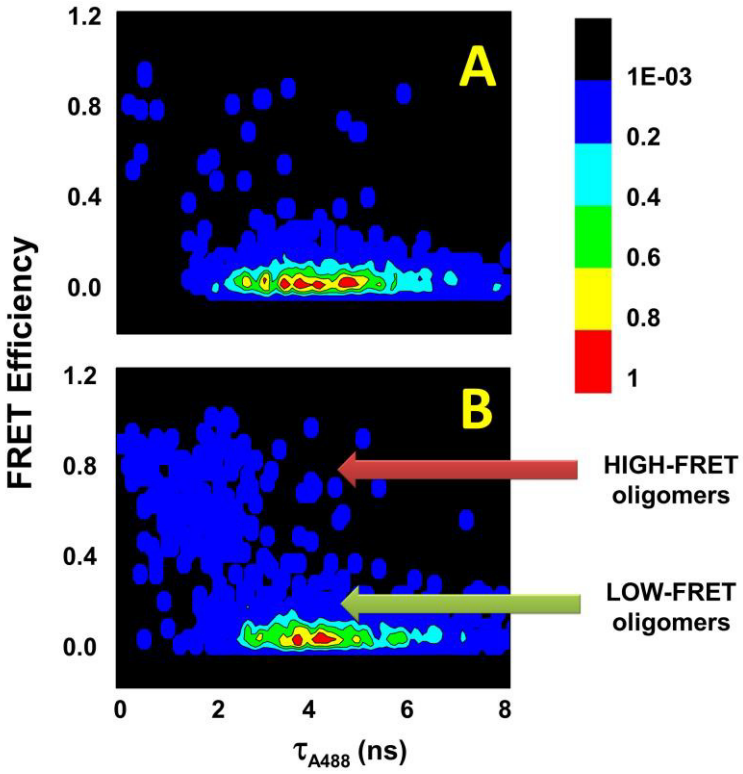


Figure 4.8 Single molecule *E versus* τ_{A488} correlograms. **A:** N47A-SH3-D and N47A-SH3-A samples mixed at a 1:1 ratio at a total concentration of 800 pM, and measured without previous incubation. **B:** N47A-SH3-DA incubated at 32 μM for 8 h.

Therefore, depending on the concentration and incubation time, we detected, within the heterogeneous complex mixture, two different types of oligomers. One of them, remarkably, is present even at low concentration and without incubating the protein samples, whereas the formation of the second distribution is concentration- and incubation time-dependent. The range of concentration that we consider more adequate for our studies was between 20 and 44 μM, because this range permits to observe the formation of the different oligomers in the early 24 hours.

As shown in Figures 4.9 and 4.10, all protein samples tested exhibited the low-FRET population at time zero). In the sample incubated at 44 µM, just after 20 minutes of incubation we already detected the formation of higher-FRET oligomers. Interestingly, the observed amount of the latter became more prominent after 1h of incubation (Figure 4.9). These oligomers appear later on during the incubation in samples incubated between 20 and 32 µM of total protein concentration (Figure 4.10), whereas in samples incubated at 16 µM we did not detect formation of high-FRET oligomers even after 24h of incubation (data not shown).

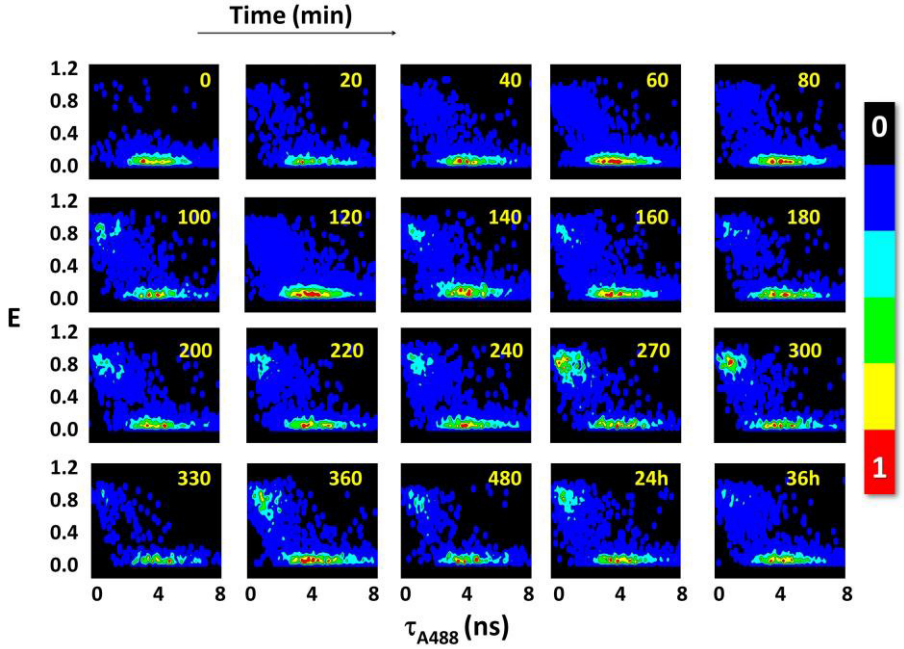


Figure 4.9 Single-molecule correlograms of the *E versus* τ_{A488} of N47A-SH3-DA samples incubated at a total concentration of 44 µM in the aggregation buffer at 37 °C. Aliquots were taken at different incubation times (expressed in minutes, except where hours are indicated).

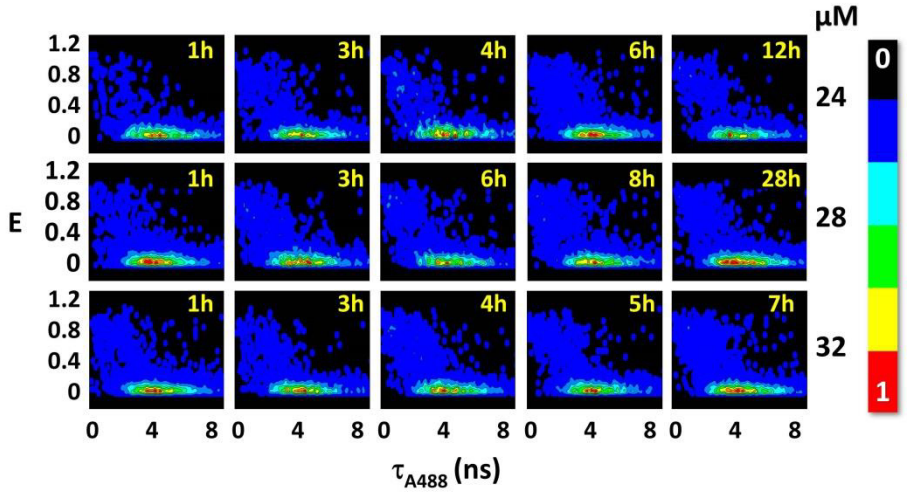


Figure 4.10 Single-molecule correlograms of the *E versus* τ_{A488} of N47A-SH3-DA samples incubated at a total concentration of 24, 28 and 32 μM in the aggregation buffer at 37 °C. Aliquots were taken at different incubation times (expressed in hours).

4.5.2 Identification of different oligomer types by oligomer size

Yet more interesting and stimulating results emerged from a different type of analysis, in which we correlated oligomer size *versus* τ_{A488}. Importantly, the access to the value of τ_{A488} permits to implement a correction on the oligomer size caused by the quenching on the A488 fluorescence (see 3.3.3). This important correction has been implemented herein for the first time, while in previous works a potential underestimation of the oligomer size systematically occurred (217, 320, 321, 387). Figure 4.11 shows oligomer size *versus* τ_{A488} correlograms from aliquots taken at different incubation times of N47A-SH3-DA samples, incubated in the aggregation buffer at different total protein concentrations.

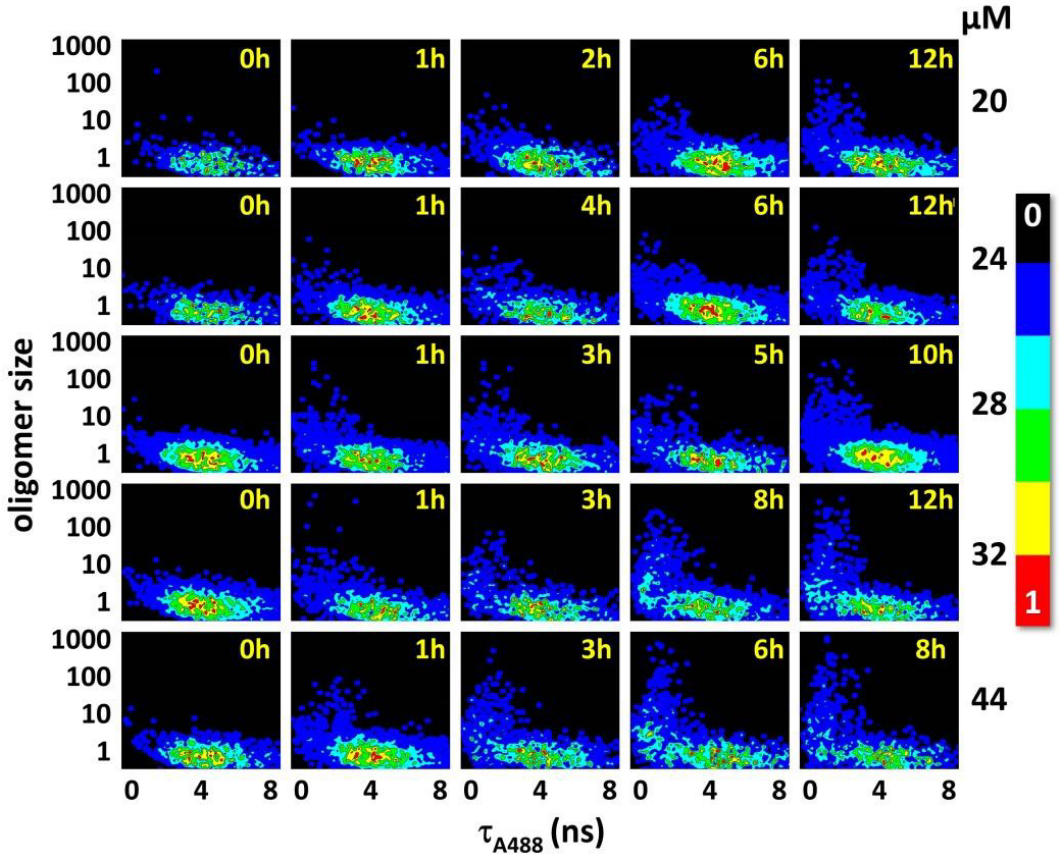


Figure 4.11 Single-molecule size *versus* correlograms of N47A-SH3-DA samples incubated at different concentrations (20, 24, 28, 32 and 44 μM) in the aggregation buffer at 37 ºC. The analyzed aliquots were taken at the indicated incubation times.

These analyses provided the unique ability of directly identifying several types of oligomers and, importantly, following their kinetic history upon aggregation. At the very early moments of aggregation, a stable population of oligomers was found, exhibiting low FRET efficiency (386). This is surprising, but well supported by the bulk experiments detailed in section 4.4. Prior to the formation of large aggregates, we detected the development of a second compact oligomer

population, whose apparent size was comparable to the initial population (from 4 to 10 monomeric units), but displaying a high FRET efficiency, as evidenced by a shortened τ_{A488} value. At longer incubation times, a third compact population of large oligomers (from 20 to over one hundred monomers), also exhibiting high FRET efficiency, clearly arose. Further details of these oligomer populations are described in section 4.7.

4.6. EFFECT OF NaCl CONCENTRATION ON THE FORMATION OF THE N47A-SH3 OLIGOMERS

Previous studies on the N47A Spc-SH3 demonstrated that the kinetics of its aggregation depends strongly on NaCl concentration (201). It was observed that at low NaCl concentration the N47A Spc-SH3 displays a prolonged lag phase. With the NaCl concentration increasing, the N47A Spc-SH3 aggregation showed a reduction of the lag phase, which may be totally removed at higher NaCl concentrations. Moreover, it was observed that NaCl concentration significantly affected the morphology of the fibrils of N47A Spc-SH3, suggesting that ionic strength lead the mechanism of fibril assembly.

We monitored the effect of the NaCl concentration on the formation of the N47A-SH3-DA oligomers. Specifically, the buffer solution was prepared at three different NaCl concentrations: 0.050, 0.10 and 0.15 M, covering the same range of the previous study performed on N47A Spc-SH3 (201).

4.6.1 Effect of NaCl concentration on the N47A-SH3 aggregation

In order to investigate possible changes in secondary structure during aggregation as a function of NaCl concentration, we followed the modifications in far-UV CD spectra of both N47A-SH3 and N47A-SH3-DA, incubated at 37 °C. In Figure 4.12 we show the far-UV CD spectra of aliquots of samples taken at different incubation times.

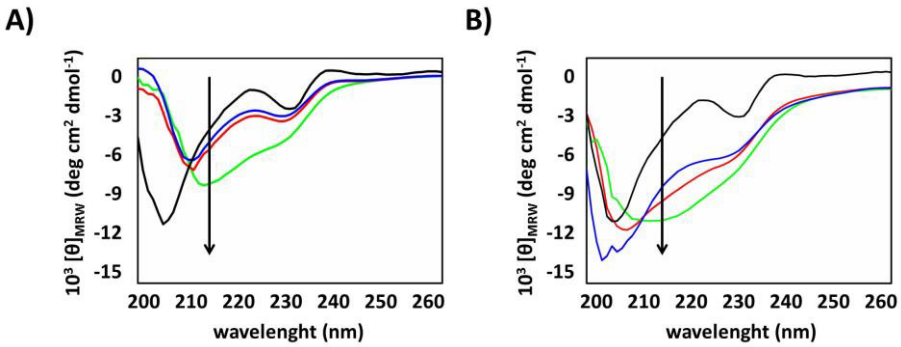


Figure 4.12 A: Far-UV CD spectra of N47A-SH3 diluted to 15 µM in 0.10 M Gly, 0.10 M NaCl at pH 3.2, at 20 °C (black line), and after incubation for 60 days at 37 °C at a protein concentration of 120 µM in 0.10 Gly, at pH 3.2, and at different concentrations of NaCl: 0.050 M (blue line), 0.10 M (red line) and 0.10 M (green line). **B:** Far-UV CD spectra of N47A-SH3-DA diluted to 15 µM in 0.10 M Gly, 0.10 M NaCl at pH 3.2, and at 20 °C (black line), and of 32 µM N47A-SH3-DA samples incubated for 60 days at 37 °C in 0.10 Gly, at pH 3.2, and at different concentrations of NaCl: 0.050 M (blue line), 0.10 M (red line) and 0.10 M (green line). Arrows indicate 215 nm, a typical band to measure β-sheet content.

The far-UV CD spectra of the proteins changed considerably upon incubation and gradually developed a negative band at ≈ 215 nm, typical of β-sheet in amyloid fibrils (151). These changes became more prominent at higher NaCl concentrations and they were drastically pronounced for both N47A-SH3 and

N47A-SH3-DA samples incubated at 0.15 M NaCl, indicating that NaCl concentrations exert a striking effect to the formation of fibrils.

We compared the morphology of the N47A-SH3 and N47A-SH3-DA aggregates formed after incubating during 60 days at 37 °C at different NaCl concentrations by using TEM. For the samples incubated at 0.050 M NaCl, we visually observed the sporadic formation of aggregates of N47A-SH3 (Figure 4.13A), whereas the N47A-SH3-DA samples incubated under the same NaCl concentration formed fibrillar structures (Figure 4.13D). At NaCl concentration of 0.10 M or 0.15 M NaCl, both labeled and unlabeled N47A-SH3 formed fibrillar structures. The N47A-SH3-DA aggregates formed at 0.10 M NaCl, as well as those of N47A-SH3 formed at 0.15 M NaCl, consisted of curved protofibrils reaching lengths of up to several micrometers (Figure 4.13C, E). Likewise, the N47A-SH3-DA at higher NaCl concentration resulted in curved protofibrillar formations, but showing a shorter length than those observed at the intermediate NaCl concentrations studied (Figure 4.13F).

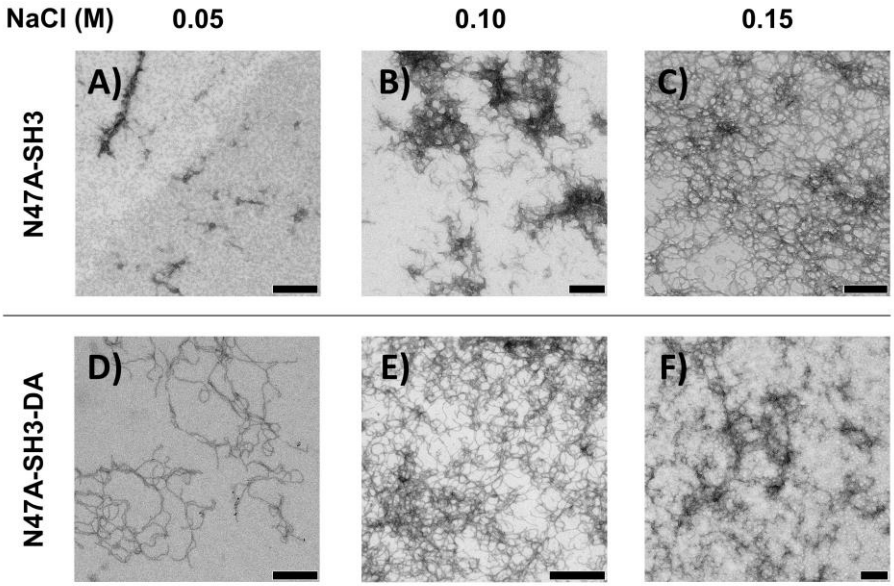


Figure 4.13 TEM images of samples of N47A-SH3 (A, B, C) and N47A-SH3-DA (D, E, F) after 60 days of incubation at 37 ºC. The N47A-SH3 and N47A-SH3-DA samples were diluted down to a final concentration of 120 µM and 32 µM, respectively, in a solution buffer containing 0.10 M glycine and different NaCl concentrations (as indicated above), at pH 3.2. Black scale bar corresponds to 200 nm.

These results were in good agreement with previous evidences on the aggregation propensity under different NaCl concentration emerged from N47A Spc-SH3 TEM images, which also showed shorter-in-length and less organized aggregates at high NaCl concentration (201). TEM evidences also justifies the net changes in secondary structure observed in the far-UV CD spectra of both N47A-SH3 and N47A-SH3-DA incubated at higher NaCl concentration, with the second showing higher aggregation propensity even at low concentration of NaCl. TEM technique however, does not allow to explain the differences observed in the morphology of N47A-SH3-DA fibers formed at 0.10 and 0.15 M NaCl, as shown in Figure 4.13E and 4.13F. Indeed, comparing the latter TEM images, NaCl concentration seem to

somehow affect fibril elongation. Previous studies on N47A Spc-SH3 revealed that at high NaCl concentration (0.20 M NaCl) the diameter of the fibers was significantly lower than that of the filaments formed at lower NaCl concentration (201). The authors suggest that NaCl concentrations mainly affect the initial phase of nucleation, where the different arrangements of the N47A Spc-SH3 molecules within the fibrils internal structures is what leads the mechanism of fibril assembly. As it will be discussed in section 4.6.3, by SMF-PIE we could notice an early appearance of highly organized oligomers at 0.15 M NaCl concentration, but the shorter lag phase that we observed at this high concentration of NaCl seems to cause the skip of the mentioned arrangements of the early oligomers. Thus, the lower extent of the N47A-SH3-DA fibrils observed in Figure 4.13F might be associated to high NaCl concentration, resulting in less efficiency in the conversion into amyloid fibrils. This may emphasize the crucial role played by the soluble N47A-SH3-DA oligomers and the significance of their initial morphological conversions leading competing pathways of the aggregation.

4.6.2 Effect of the NaCl concentration monitored by SMF-PIE

The formation of different N47A-SH3-DA oligomers at different NaCl concentration was studied using SMF-PIE techniques. Samples of N47A-SH3-DA were incubated to a final concentration of 32 μM in 0.10 M Gly, pH 3.2, at 37 °C and different NaCl concentrations: 0.05, 0.10 and 0.15 M.

The FRET efficiency correlograms of the N47A-SH3-DA samples incubated at 0.050 M NaCl did not show significant change in oligomer distributions during the incubation (Figure 4.14A). The initial low-FRET oligomers detected were the prevalent species during the course of the aggregation even in samples incubated

for up to 4 days. This result was in accord with what we observed in the oligomer size *versus* τ_D correlograms obtained from the same experiments, showing slight or no changes in oligomer growth within the initial distribution, also even after 4 days of incubation (Figure 4.14B). However, as shown by TEM images under those same NaCl concentration (Figure 4.13D), at low concentration of NaCl (0.050 M) the N47A-SH3-DA samples do form fibrils. According to what we observed by SMF-PIE experiments, low NaCl concentrations could cause aggregation to take place with an extremely slow lag phase, mainly dominated by the presence of the low-FRET, small size oligomers.

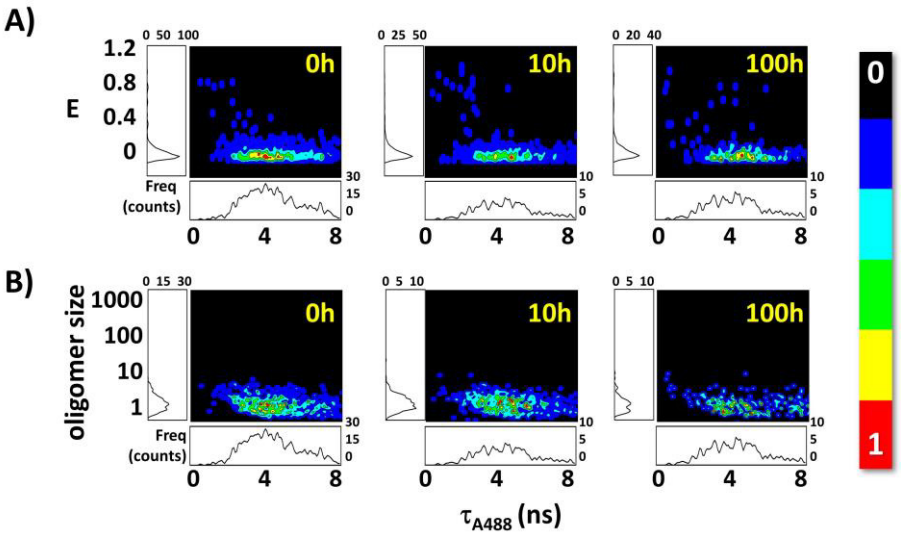


Figure 4.14 A: Single-molecule *E versus* τ_{A488} (**A**) and oligomer size *versus* τ_{A488} (**B**) correlograms of N47A-SH3-DA samples incubated at 37 °C at a total concentration of 32 μM in a buffer containing Gly 0.10 M, NaCl 0.050 M, at pH 3.2. Aliquots were taken out and analyzed after 0, 10 and 100 hours of incubation.

At 0.10 M and 0.15 M NaCl concentration, the single-molecule correlograms of *E versus* τ_{A488} and of oligomer size *versus* τ_{A488} clearly showed different distributions of oligomers. In both cases, we observed the presence of low-FRET oligomers in the early phases of the incubation that undergo different arrangements as incubation proceeds.

The SMF-PIE study on the N47A-SH3-DA oligomers incubated at 0.10 M NaCl was already described in 4.5.1 and 4.5.2 (Figures 4.10 and 4.11). Observing the changes on the single-molecule correlograms of oligomer size *versus* τ_{A488} for samples incubated at 32 μM of total protein concentration, we distinguished an initial low-FRET distribution of small-size oligomers in the first stages of the incubation. After 3 hours of incubation, together with these oligomers, it was clearly observed the formation of high-FRET oligomers, sharing the same small-size of the first ones. Just only after the appearance of these second oligomers, it was revealed the presence in solution of another oligomer type, high-FRET and of larger size than those aforementioned.

The samples incubated under 0.15 M NaCl (Figure 4.15) displayed the formation of a high-FRET distribution of small-size oligomers after only one hour of incubation. After 6 hours of incubation, the oligomer size *versus* τ_{A488} single-molecule correlograms showed a distribution of high-FRET and large size oligomers. However, the extent of these oligomers appeared less pronounced compared to the same population found in samples incubated at 0.10 M NaCl, analyzed at later incubation times (Figure 4.14B).

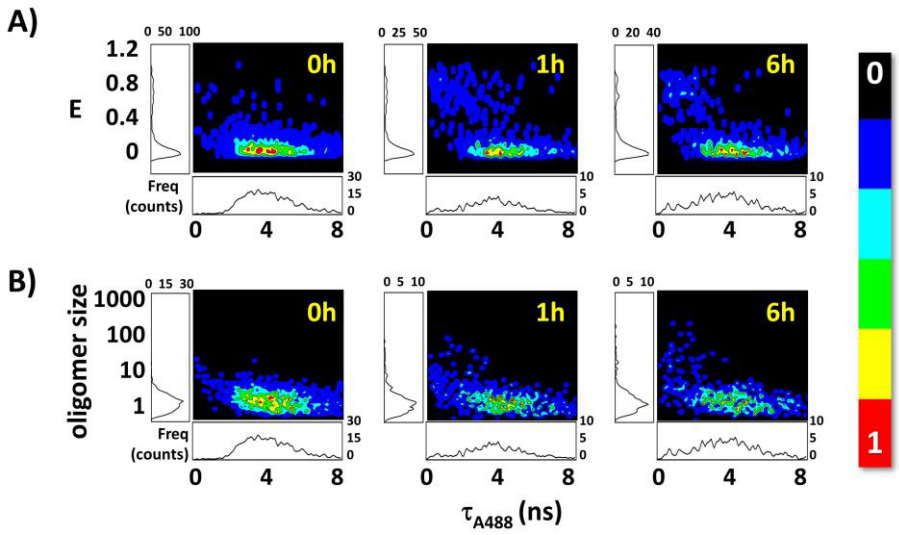


Figure 4.15 Single-molecule correlograms of *E versus* τ_{A488} **(A)** and oligomer size *versus* τ_{A488} **(B)** of N47A-SH3-DA samples incubated at 37 °C at a final protein concentration of 32 μM in a buffer containing 0.15 M NaCl, 0.10 M glycine, at pH 3.2. Aliquots were taken out and analyzed after 0, 1 and 6 hours of incubation

The observed discrepancies among the single-molecule correlograms for N47A-SH3-DA samples taken at different NaCl concentrations, suggest that under higher NaCl concentration the lag phase is neglected or considerably reduced. This might yield the skip of some crucial molecular events that take place during the initial stages of the amyloid fibril formation, as these oligomer interactions lead to the accumulation of amyloidogenic N47A-SH3-DA oligomers. The importance of nonspecific interactions for amyloid nucleation has been recently reported by Saric and colleagues using computer simulations (186). Therefore, the morphology of the N47A-SH3-DA fibers observed in Figure 4.13F, shorter and thinner than those formed at 0.10 M NaCl (Figure 4.13E), may be the result of alternative competing pathways of assembly, led by the different NaCl concentration.

4.6.3 Effect of the NaCl concentration monitored by FLIM-PIE

An advanced FLIM-PIE imaging scheme allowed us to investigate the oligomers deposited on the glass slide surface. This study is especially important to study those oligomers formed in the later times of incubation, whose presence cannot be detected in solution because of their large size and weight. This additional direct imaging information completes the screening of all the species present in the samples.

We analyzed the N47A-SH3-DA samples incubated at 37 °C at 32 μM of total protein concentration in a pH 3.2 buffer containing 0.10 M Gly and 0.050, 0.10 or 0.15 M NaCl. The samples incubated at 0.050 M NaCl, did not show the presence of N47A-SH3-DA oligomers deposited on the slide surface, even after 3 days of incubation (data not shown), whereas under the same conditions we detected and characterized in solution low-FRET and small-size oligomers (Figure 4.14A and B). This was probably due to the low MW of these oligomers that did not deposit on the slide surface, and are washed out. This confirms the slow kinetics of the aggregation under these conditions of NaCl concentration, including an extended lag phase dominated by the initial formation of very small oligomers. Nevertheless, the donor FLIM-PIE images of N47A-SH3-DA samples incubated at 0.10 and 0.15 M NaCl showed the presence of heterogeneous oligomers. These exhibited much variance in fluorescence intensities, which is related with their size and, at longer times of incubation, also in their τ_{A488} (Figures 4.16 and 4.17). For those samples incubated at 0.10 M NaCl, after 2 hours of incubation, we detected oligomers exhibiting a distribution of τ_{A488} centered at 3.58 ± 0.02 ns (values obtained from the average of 5 donor FLIM images). Given the unquenched fluorescence lifetime of the A488 dye is 4 ns, these oligomers exhibited a low FRET efficiency. Together with these low-FRET oligomers, we observed another

types of oligomers characterized by a distribution of τ_{A488} centered at 1.81 ± 0.02 ns, indicating high FRET efficiency. After 7 hours of incubation, the peak of the τ_{A488} distribution of the early oligomers, before centered at 3.58 ± 0.02 ns, shifted towards 3.12 ± 0.02 ns, suggesting a mild rearrangement to more compact oligomers. Together with the presence of these two types of oligomers detected in the first hours of incubation, we also detected the appearance of oligomers, apparently larger in size than those formed earlier, and whose distribution of τ_{A488} was centered at 0.93 ± 0.03 ns. In the later times of incubations, these oligomers characterized by high-FRET and high intensity, frequently become the predominant species detected in the FLIM-PIE images. Figure 4.16 shows some representative donor FLIM images, together with their corresponding τ_{A488} distributions.

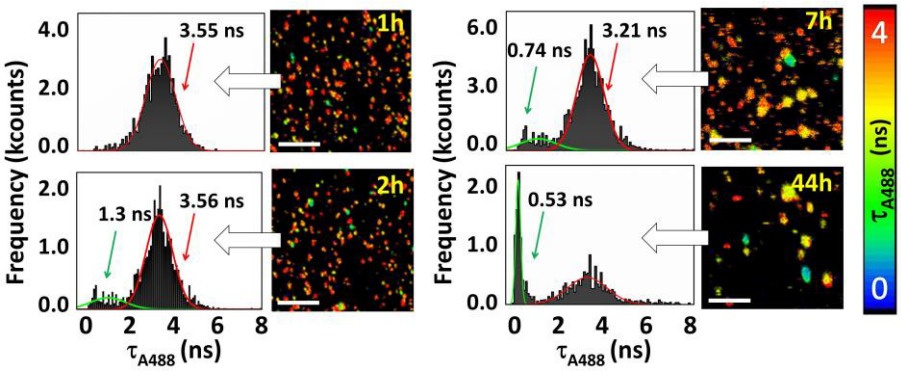


Figure 4.16: Representative donor FLIM images with their corresponding τ_{A488} distributions, from N47A-SH3-DA samples incubated at 37 °C at 32 μM total protein concentration, in a buffer containing 0.10 M NaCl, 0.10 M glycine, at pH 3.2. Aliquots were taken out the incubating sample at the times shown in the figures. White scale bars: 2.4 μm.

We also performed the FLIM-PIE analysis of N47A-SH3-DA samples incubated at 0.15 M NaCl. The FLIM images at the highest NaCl concentration tested

showed a similar tendency of these oligomers to shift towards shorter τ_{A488} as the incubation proceeds, but we did not observe any rearrangement of the early oligomers. Indeed, the τ_{A488} distribution of these low-FRET oligomers was centered at 3.5 ns, and this value did not change in the first 4 hours of incubation, where they were the only species detected. In addition, at later times, we did not observe the presence of much extent of high-FRET and high-intensity oligomers, as detected for those samples incubated at 0.10 M NaCl. These results were in agreement to what we observed in the single-molecule correlograms of oligomer size *versus* τ_{A488} under the same NaCl and protein concentration (Figure 4.15B): FLIM-PIE experiments confirmed the rapid formation of small-size oligomers, but exhibited a lag time of ≈ 6 hours to finally develop high-FRET and large size N47A-SH3-DA oligomers. Moreover, the high-intensity oligomers, that we considered apparently larger in size, were only sporadically detected in donor FLIM images. This also agrees to the less extent of high-FRET and large-size oligomers emerged from the SMF-PIE measurements in solution (Figure 4.15B).

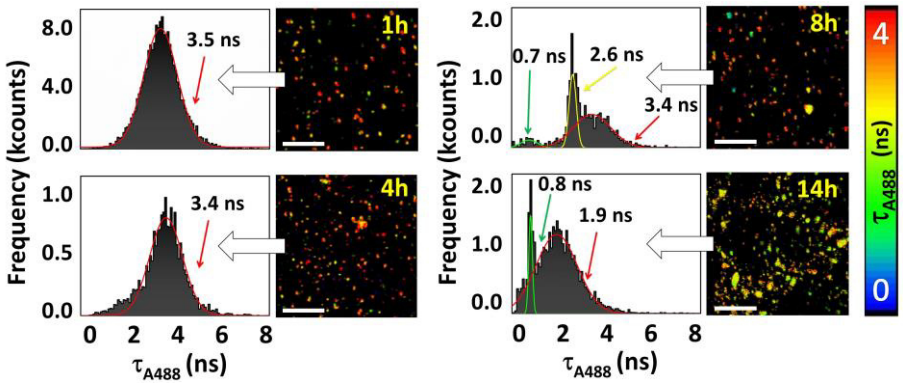


Figure 4.17: Representative donor FLIM images with their corresponding τ_{A488} distributions of N47A-SH3-DA samples, incubated at 37 °C at 32 µM total protein concentration, in a buffer containing 0.15 M NaCl, 0.10 M glycine, at pH 3.2. Aliquots were taken out the incubating sample at times shown in the figures. White scale bars: 2.4 µm.

4.7 DISCUSSION

Here we have shown that the combination of conventional biophysical techniques and novel SMF techniques permits to obtain a significant level of detail on the amyloidogenic modified mutant of the N47A Spc-SH3. This is the first time that this protein has been studied through such a widely range of techniques.

Examining the aggregation stages of the mixture of N47A-SH3 and N47A-SH3-DA by TEM, we confirm that either the six residue tag or the labeling do not alter the ability of the proteins to form mature amyloid fibrils. Far-UV CD and near-UV CD experiments demonstrated that the protein structure was not affected by the chemical modifications on the protein, in spite of a slight contribution to the CD spectra probably caused by the chirality in the molecular structure of ATTO647N. However, these modifications strongly accelerate the formation of amyloid fibrils in the presence of moderate NaCl concentrations, whereas the N47A Spc-SH3 showed a slower fibrillation under the same conditions. Moreover, DSC experiments performed in conditions favoring the formation of aggregates showed complicated thermograms for the N47A-SH3-DA, that it may be associated to the presence of intermediate oligomers or to the formation of high order aggregates. These formations were studied by cross-linking revealing early oligomers of low MW, detectably even before incubating the samples, therefore proving a genuine oligomerization in the protein.

Using SMF-PIE, the oligomers of N47A-SH3-DA were characterized at the single-molecule level. We observed the presence of early oligomers at protein concentrations lower than those used in conventional studies on amyloid aggregation. This early oligomers showed a low FRET efficiency, and were detected under all NaCl concentrations tested. Strikingly, they were observed

even when mixing N47A-SH3-D and N47A-SH3-A samples together at pM concentrations in the absence of incubation. The presence of these early aggregates was also detected by FLIM-PIE images performed at 0.10 M and 0.15 M NaCl concentration. These images showed low-FRET and low-intensity oligomers at the initial times of the incubation. An in-depth characterization of these initial oligomers and of their equilibrium will be discussed later in chapter 5.

The heterogeneity observed by bulk experiments during the aggregation of the N47A-SH3-DA was investigated by SMF-PIE and FLIM-PIE techniques, at lower concentrations than that used in bulk techniques. These techniques resulted adequate for studying the oligomers formed during the aggregation at the single molecule level and at concentrations close to the physiological ranges. The optimal conditions for the study of the N47A-SH3-DA aggregation resulted those that were already reported by Varela and colleagues (334) and Morel and colleagues (201) in previous studies on the N47A Spc-SH3. These were 0.10 M NaCl, 0.10 M Gly, at pH 3.2 and at 37 °C. In these experimental conditions, the N47A-SH3-DA aggregation exhibits a lag phase in which the amount of produced oligomers and their time evolution is compatible with our SMF-PIE experiments. In contrast, other concentrations of NaCl (0.05 or 0.15 M) did not result adequate to monitor the oligomers evolution, due to either a protracted lag phase (at 0.05 M NaCl) or to a too reduced one (at 0.15 M).

Importantly, our SMF-PIE experiments revealed the formation of different types of oligomers. The single-molecule correlograms of *E versus* τ_{A488} and especially of oligomer size *versus* τ_{A488} permitted to identify different populations of oligomers formed during the course of the incubation under conditions favoring the aggregation. As shown in Figure 4.18, we identified the different types of aggregates according to their size and on their FRET efficiency (through the τ_{A488}

distribution), by setting up specific regions in the oligomer size *versus* τ_{A488} correlograms (Figure 4.18B).

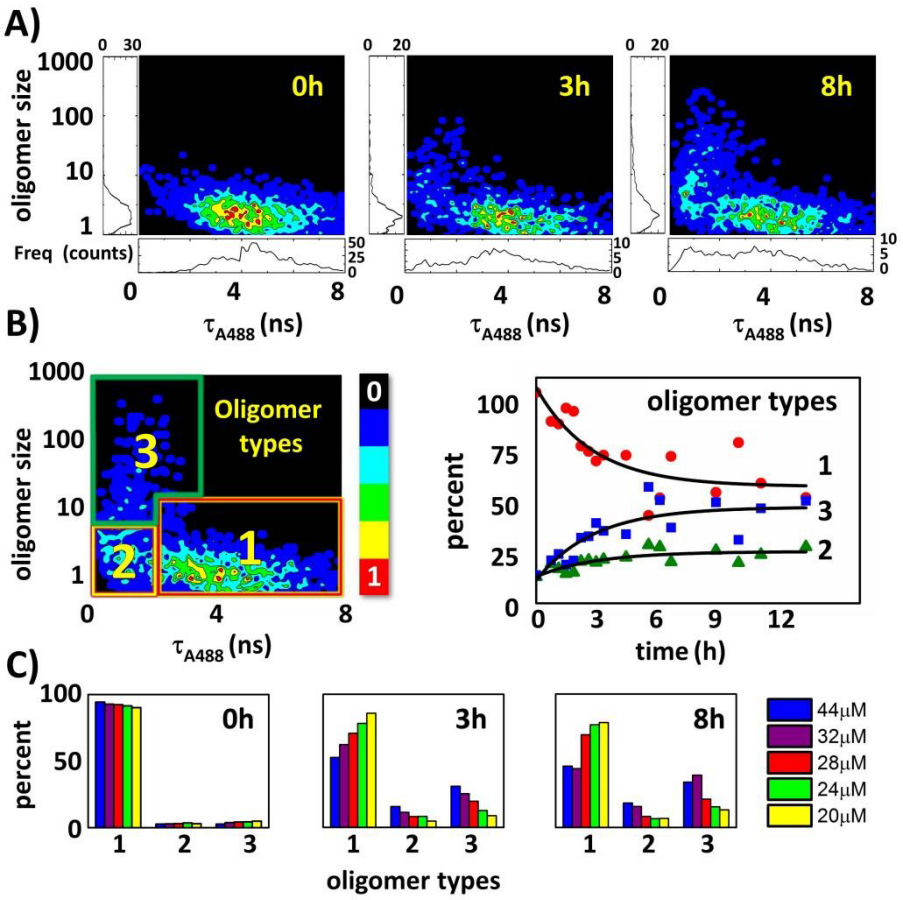


Figure 4.18 A: Oligomer size *versus* τ_{A488} single-molecule correlograms of N47A-SH3-DA samples incubated at 32 μM in the aggregation buffer for 1, 3 and 8 hours. **B:** On the left, definition of the three types of oligomers, **1, 2**, and **3**, from the oligomer size *versus* τ_{A488} correlograms. On the right, illustrative time trace of the oligomers populations from N47A-SH3-DA samples incubated at 32 μM of total protein concentration. **C:** Relative concentrations (percentage) of each type of oligomer (**1**, **2**, and **3**), monitored during the course of the aggregation reaction, after 0, 3, and 8 hours of incubation.

These analyses provided the unique ability of directly identifying three types of oligomers and, importantly, following their kinetic history upon aggregation. At the very early moments of aggregation, yet in the absence of actual incubation conditions, an initial population of low-FRET and small-size oligomers was found (type **1**). Prior to the formation of large aggregates, we detected the development of a second oligomer population (type **2**), whose size is comparable to the initial population (from 4 to 10 monomeric units), but displaying a high FRET efficiency, as evidenced by a shortened τ_{A488} value. At later incubation times, a third population of large oligomers (from 20 to over one hundred monomers), also exhibiting high FRET efficiency, clearly arose (type **3**) (Figure 4.18B). The distinction between types **2** and **3** was not detectable through *E versus* τ_{A488} correlograms (Figure 4.8), since the heterogeneity between populations could only be noticed from the different total fluorescence intensity of the events. Moreover, the addition of an orthogonal parameter, such as τ_{A488}, provides an estimation of the FRET efficiency independently from the bursts intensity, being a much more reliable approach than the estimation of both *E* and oligomer size from the same burstwise intensity, a previously employed method to study α-synuclein oligomers (320, 388).

Given our methodology is capable of following the time evolution of the three different types of aggregates (Fig. 4.18C), we have performed an in-depth kinetic study of the N47A-SH3-DA amyloid aggregation, relating it to the fundamental mechanistic model of amyloid fibril formation. Details of this study are described in chapter 6.

5. THE FIRST OLIGOMERIZATION EQUILIBRIUM

5 – THE FIRST OLIGOMERIZATION EQUILIBRIUM

5.1 THE EARLY EVENT OF THE AGGREGATION

As summarized in the introduction, the characterization of the early oligomeric structure that can precede amyloid fibril growth has been become of great importance, especially thanks to recent studies that have identified differential cytotoxicity caused by different pre-amyloidogenic oligomers (74, 76, 389).

Our previous study on the N47A-SH3 oligomers at the single molecule level has revealed the presence of various types of oligomers during the aggregation of N47A-SH3 (see chapter 4). Importantly, we have detected small oligomeric species since the early phases of the oligomerization, even prior to incubation. The detection of these oligomers was observed even at the low, physiologically relevant, protein concentrations employed. In this chapter, we focus on this surprising population of oligomers, through a series of complementary experiments to fully characterize the first step toward amyloid fibril formation of N47A-SH3.

5.2 CHARACTERIZATION OF EARLY OLIGOMERS AT THE SINGLE MOLECULE LEVEL

We hypothesize that the oligomers in this early population are both dynamic and labile in nature because they are ubiquitous, even in the absence of incubation. To investigate their stability, we performed a series of SMF-PIE experiments in which the N47A-SH3-D and N47A-SH3-A samples were kept separated in deionized water at low concentration (from 20 to 40 μM), and just immediately prior to the

measurement, they were mixed (1:1 ratio) to pM concentration in the aggregation buffer (to remind, pH 3.2, 0.10 M NaCl, 0.10 Gly). This protocol ensured that all of the oligomers that included both N47A-SH3-D and N47A-SH3-A were readily formed *de novo* within the time frame of the SMF-PIE measurement, and confirmed the dynamic nature of these oligomers.

The low protein concentration used for the SMF-PIE measurements prevents any further amyloidogenic aggregation. Specifically, our SMF-PIE experiments were performed at the room temperature between 50 and 1000 pM of total protein concentration. The detection of the formation of aggregates was confirmed using the coincidence criterion, i.e., simultaneous single-molecule fluorescent bursts detected in the A488 channel and in the A647N channel (313). SMF-PIE experiments permitted to obtain, from a single measurements, the donor fluorescence lifetime τ_{A488}, the distribution of the A488 and A647N intensity ratio (F_{A488}/F_{A647N}) and the fraction of coincident events, i.e., the association quotient, Q (313).

The single molecule correlograms of F_{A488}/F_{A647N} *versus* τ_{A488} shows the amount of oligomers formed at each concentration, the distribution of the stoichiometry, and the intra-oligomer FRET efficiency. Some of these correlograms are shown in Figure 5.1.

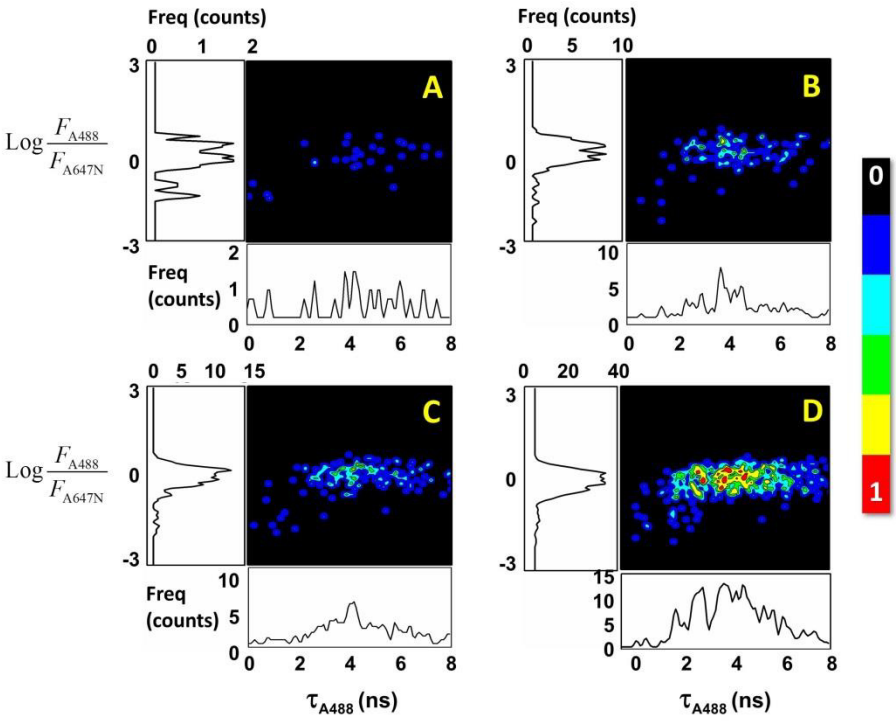


Figure 5.1 Single molecule correlograms of the F_{A488}/F_{A647N} ratio (logarithmic scale) *versus* τ_{A488} of the coincident events from mixtures of N47A-SH3-D and N47A-SH3-A (1:1 ratio) at total protein concentrations of 75 (**A**), 250 (**B**), 400 (**C**) and 700 (**D**) pM.

These correlograms reveal a formation of oligomers even at very low concentrations (Figure 5.1A). These oligomers do not show any significant difference in τ_{A488}, which is centered approximately in 4 ns. The F_{A488}/F_{A647N} distribution (logarithmic scale) is centered at zero, indicating a 1:1 stoichiometry of the -A488 and -A647N. From the average of the non-coincident donor channel and acceptor channel bursts intensities, obtained from 17 different SMF-PIE experiments performed between 75 and 1000 pM of total protein concentration, we obtained an average burst intensity value of 29.7 kHz for the N47A-SH3-D events and 31.9 kHz for the N47A-SH3-A events. The average of these values

(30.8 kHz) was considered as the bursts intensity value of one of each labeled monomer. The average burst intensity values of the donor and acceptor channel of the coincident events, i.e. the detected oligomers were 34.5 kHz and 34.7 kHz, respectively). These values range on the same intensity values of a labeled N47A-SH3 monomer, therefore, we considered that the majority of the detected oligomer species (as coincident events) must be dimers. The analysis of the single-molecule oligomer size *versus* τ_{A488} correlograms obtained from the sum of all SMF-PIE measurements, also confirmed the small size nature of these oligomers detected (Figure 5.2).

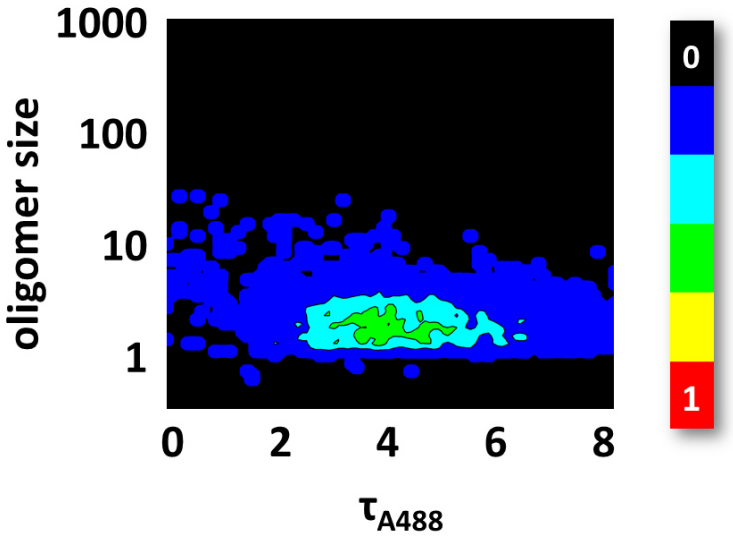


Figure 5.2 Single-molecule oligomer size *versus* τ_{A488} correlogram averaged from all the data acquired between 75 and 1000 pM (data normalized from 0 to 1).

Further evidences on the dimeric nature of the oligomers emerged considering the *Q* values, indicating the percentage of the aggregated species formed de novo. As an illustrative example of one of our measurements, with a 1 % of aggregated species, the probability of forming a trimer would be 100 times lower than forming

new dimers. Encounters to form tetramers would be even much less probable. Therefore, the *Q* values support the hypothesis that these oligomers are mostly dimers. Despite being an approximation, this assumption is however reasonable and justified.

5.3 FLIM-PIE EXPERIMENTS

In order to visually monitor the formation of the early oligomers, we performed FLIM-PIE experiments at low concentration. Firstly, we mixed samples of N47A-SH3-D and N47A-SH3-A to a total protein concentration of 32 μM into the aggregation buffer. After dilution to a final protein concentration of 280 pM (single-molecule level) and deposition onto a microscope glass slide, the single proteins were imaged. The analysis of the FLIM-PIE images (Figure 5.3) revealed several events due to A488-only and A647N-only N47A-SH3. The presence of oligomers was confirmed by the simultaneous colocalization of single molecules in the donor, the acceptor, and the FRET image (circled molecules in Figure 5.3A). Significantly, the τ_{A488} distributions within the aggregates (Figure 5.3A) did not appreciably differ from the A488-only monomers (Figure 5.3B), supporting the notion of a low FRET efficiency and therefore a loose structural arrangement.

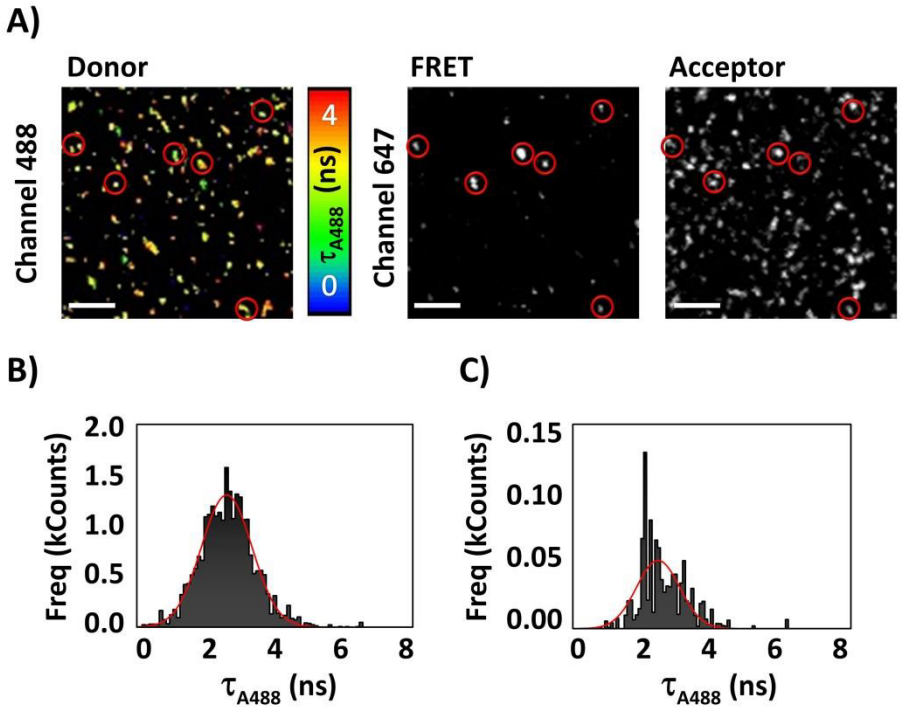


Figure 5.3 PIE-FLIM experiments. **A:** Donor (F_{A488}), FRET (F_{FRET}), and directly excited A647N (F_{A647}) images of N47A-SH3-DA diluted in the aggregation buffer to a total protein concentration of 32 µM. The circles indicate colocalized spots that contain all three events: A488, FRET, and directly excited A647N. The donor FLIM image is colored using a pseudo-color scale ranging from 0 to 4 ns. The scale bars represent 2.4 µm. **B:** τ_{A488} distribution among non-colocalized molecules (donor-only), corresponding mainly to N47A-SH3-D monomers. **C:** τ_{A488} distribution among protein aggregates (circled regions) collected over 10 different images.

In order to quantify the FRET efficiency of such oligomers, we collected images of N47A-SH3-D to obtain the τ_{A488} in the absence of acceptor (Figure 5.4). To this, N47A-SH3-D and N47A-SH3-A samples were kept separately in deionized water at pH 3, and diluted in the aggregation buffer to a total protein concentration of 375 pM. The average lifetime of N47A-SH3-D on glass slides in the absence of

acceptor was 3.40 ± 0.10 ns. The gaussian fit on Figure 5.4C on the oligomer population was centered at 2.45 ns, and hence represents an average FRET value of 0.28 (= 1 – 2.45/3.40), confirming the low-FRET nature of these oligomers. The lifetime histogram in Figure 5.3C exhibits an additional peak centered at 2.1 ns from oligomers undergoing a slightly higher FRET than the rest of the aggregates. Such population is still in the low-FRET regime, with E = 0.38. It is important to note that a larger FRET efficiency values, of around 0.7, would exhibit a τ_{A488} of approximately 1 ns. Therefore, this peak in the τ_{A488} distribution is only caused by random donor-acceptor distances in low-FRET, loosely-structured oligomers.

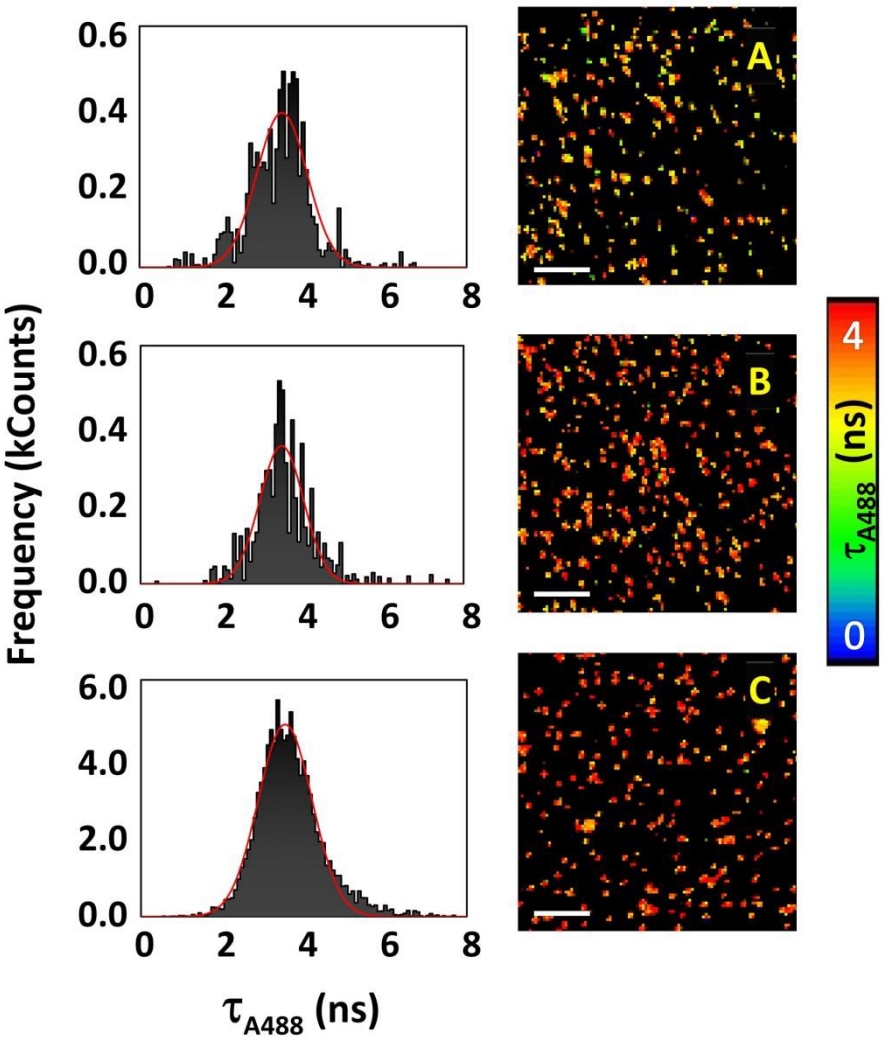


Figure 5.4 PIE-FLIM experiments. **A:** N47A-SH3-D in 10 mM Phosphate, 100 mM NaCl, pH 7.2. **B:** N47A-SH3-D in aggregation buffer. **C:** N47A-SH3-D and N47-SH3-A directly mixed in the aggregation buffer.

5.4 MODEL FOR FRET EFFICIENCY WITHIN N47A-SH3 DIMERS

In order to provide further proof on the low-FRET and dimeric nature of the dynamic, early oligomers, we developed a model for loosely packed dimers based on geometrical calculations. To this, we used the available structural information on the N47A-SH3 domain of α-Spectrin (PDB entry 1QKX, Figure 5.5 (390)). We modeled a dimer (Figure 5.6) in which the centers of the two monomers are separated a distance, *D*, equal to the full diameter of the protein (25 Å). The six-residue, unstructured tail and the dyes' linkers add a certain distance from the dyes to the protein center. The length of the flexible 6-residue random coil can be obtained from polymer dynamics as the root-mean-square end-to-end distance: $\langle rms \rangle = (1.927 \cdot n^{0.59}) \cdot 6^{0.5}$, with $n = 6$. This distance is 13.6 Å. The distance added by the fluorophores' linker was considered to be 12 Å, as typically employed in single-molecule FRET experiments (391). Therefore, the distance *L* in the model equals 38.1 Å, by adding in half of the protein diameter (12.5 Å) to the distance of the flexible tag and the fluorophores' linker (13.6 Å + 12 Å + 12.5 Å = 38.1 Å). We considered three different angles, θ_1, θ_2, and θ_3, to probe all of the potential inter-fluorophore distance space. Trigonometric calculations resulted in the following equation to assess the inter-fluorophore distance, *R*:

$$R = \sqrt{\left(L\cos\theta_2 \sin\theta_3\right)^2 + \left(D + L\left(\sin\theta_1 + \sin\theta_2\right)\right)^2 + \left(L\left(\cos\theta_1 - \cos\theta_2 \cos\theta_3\right)\right)^2} \quad \textbf{(5.1)}$$

We then determined the averaged inter-fluorophore distance by varying θ_1 and θ_2 between –20º and 200º (where 0º is the vertical position shown in Figure 5.6) considering the steric hindrance of the other unit and full rotation for θ_3 (0º-360º):

$$\langle R \rangle = \frac{\int_{-20°}^{200°}\int_{-20°}^{200°}\int_{0°}^{360°} R \cdot d\theta_1 d\theta_2 d\theta_3}{\int_{-20°}^{200°}\int_{-20°}^{200°}\int_{0°}^{360°} d\theta_1 d\theta_2 d\theta_3} \quad \textbf{(5.2)}$$

The triple integral eq. (5.2) was solved using MathCad 15.0 (PTC, Needham, MA) and resulted in $\langle R \rangle$ = 72.4 ± 18.9 (s.d.) Å. This distance corresponds to an average FRET efficiency of 0.11 ± 0.15 (with R0 = 51 Å (351)). Therefore, this model supports the evidences of low-FRET dimers emerged from the analysis of the single-molecule experiments.

.

Figure 5.5 α-spectrin Src homology 3 domain, N47A mutant in the distal loop (392) . PDB entry 1QKX.

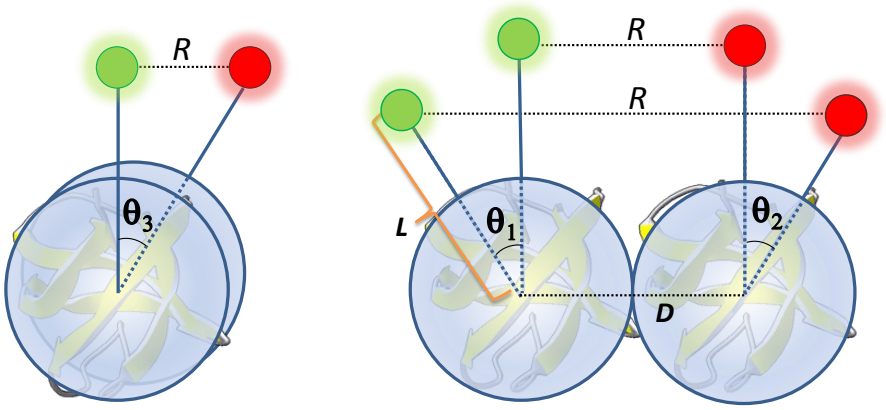


Figure 5.6 Model for loosely packed N47A-SH3 dimers, including A488- and A647N-labeled monomers. The variation of angles θ_1, θ_2, and θ_3 over the different possible values probes the inter-fluorophore distance space distribution.

5.5 ENSEMBLE EXPERIMENTS

In order to support the significance of the population of oligomers detected without the need of incubation by single molecule techniques, we also performed additional ensemble experiments, at high concentrations, to validate the presence of this type of aggregate.

5.5.1 Cross-linking experiments

Our previous cross-linking experiment revealed the presence of the N47A-SH3, N47A-SH3-D and N47A-SH3-A oligomers at different temperature (see Chapter 4). Here we focused exclusively on the N47A-SH3 and N47A-SH3-DA oligomers that characterize the early phase of the aggregation, performing the experiments at 20 °C. The samples were diluted to 60 μM in the aggregation buffer, because this concentration allowed to distinguish the bands related to each protein tested in

SDS-PAGE gel with sufficient resolution. The SDS-PAGE gel revealed a ladder of species in which the main band corresponded to the monomeric form, with higher-order species also detected (Figure 5.7).

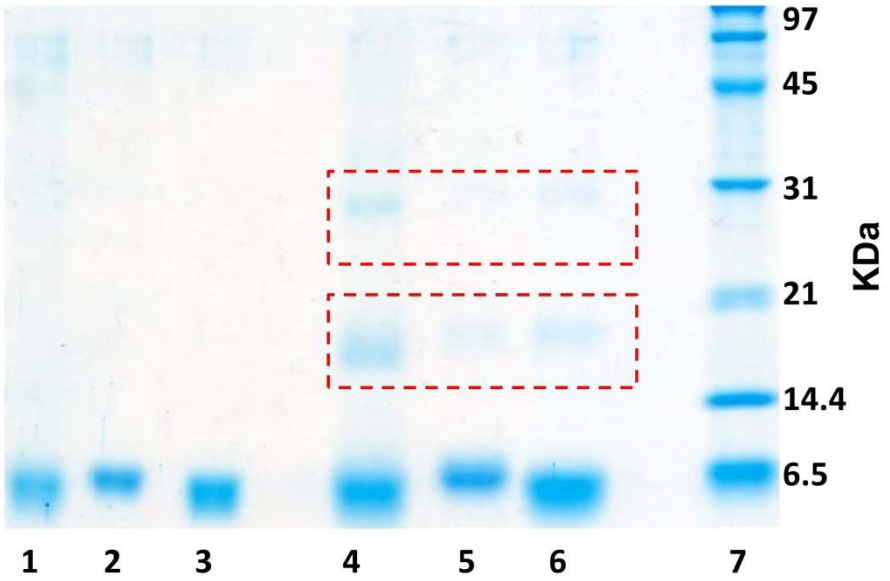


Figure 5.7 Lanes 1, 2, and 3: Non-incubated N47A-SH3, N47A-SH3-D, and N47A-SH3-A, respectively, at 60 μM in the aggregation buffer at 20 ºC without the crosslinking treatment. Lanes 4, 5, and 6: Non-incubated N47A-SH3, N47-SH3-D, and N47A-SH3-A, respectively, at 60 μM in the aggregation buffer at 20 ºC after the crosslinking protocol. Lane 7: Molecular weight marker.

These oligomers were present in all of the samples tested, with or without a fluorescent probe attached, indicating a genuine oligomerization process carried out by the protein. Interestingly, the presence of SDS resulted in the disaggregation of the oligomers when the cross-linking agent was not used. This result supports the theory that these associations are relatively weak and non-covalent in nature.

5.5.2 Dynamic light scattering

We investigated the oligomers formed from samples of N47A-SH3, N47A-SH3-D, N47A-SH3 and N47A-SH3-DA diluted to 60 and 120 μM in the aggregation buffer also performing DLS measurements at 20 °C, without prior protein incubation (Figure 5.8 and Figure 5.9). The concentration range was chosen because at concentration lower than those, it would be challenging to obtain reliable DLS results.

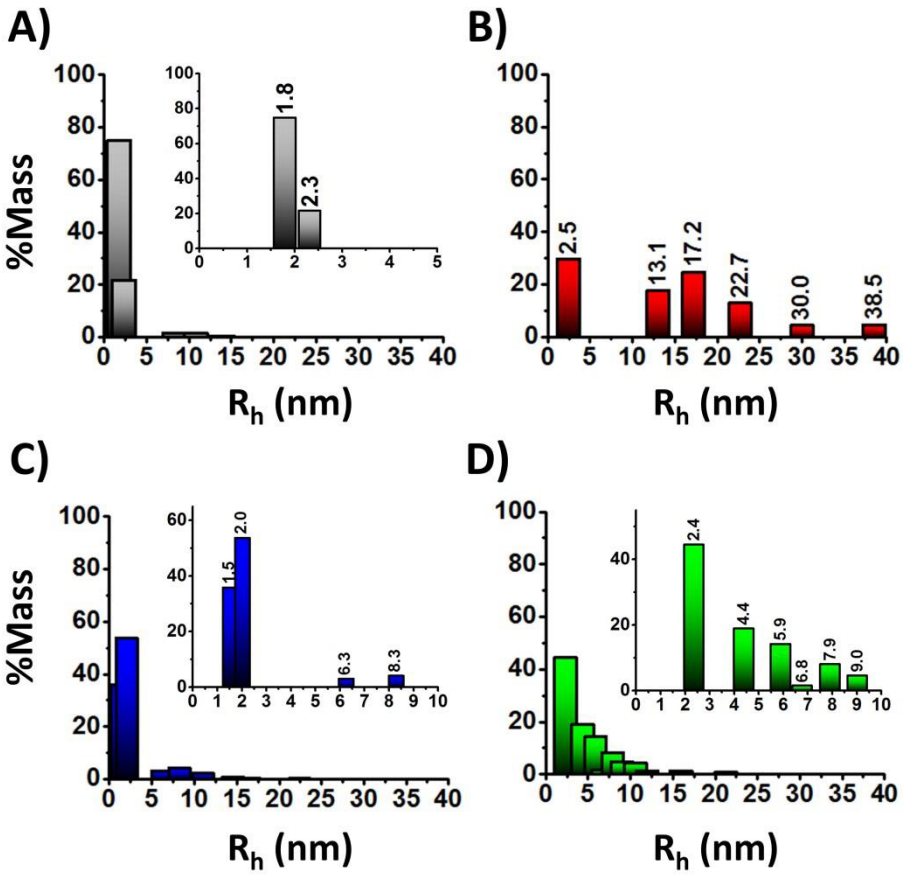


Figure 5.8 DLS histograms of N47A-SH3 (**A**), N47A-SH3-D (**B**), N47A-SH3-A (**C**), N47A-SH3-DA (**D**) diluted to 60 μM in the aggregation buffer, at 20 ºC.

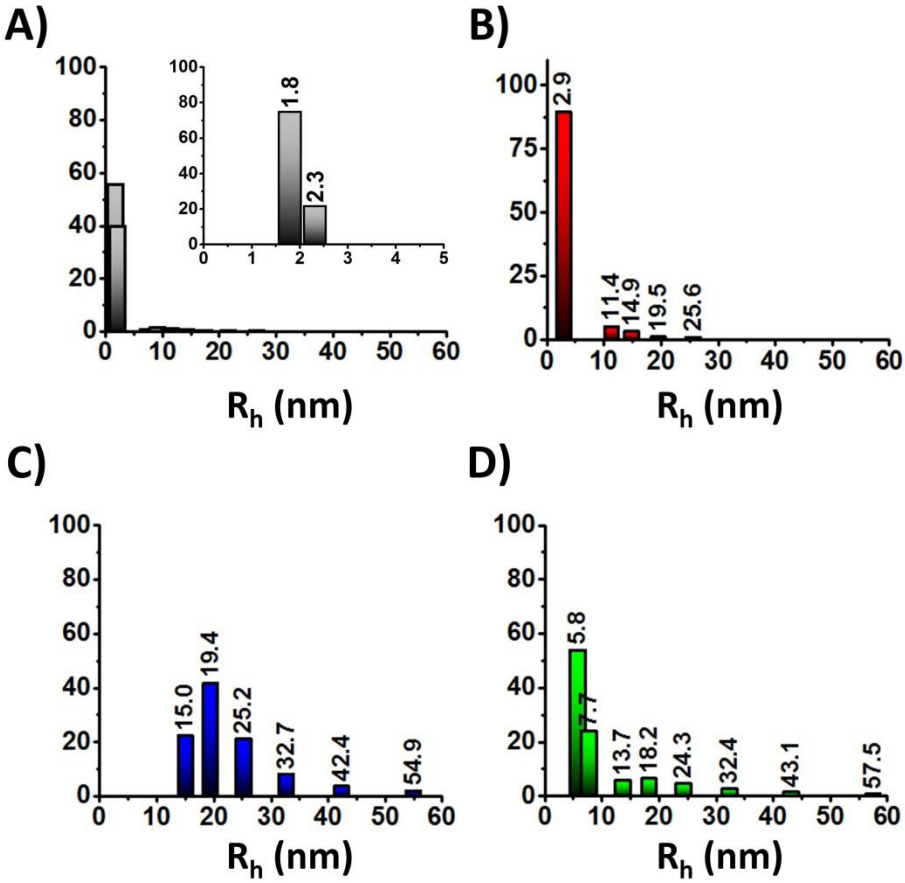


Figure 5.9 DLS histograms of N47A-SH3 (**A**), N47A-SH3-D (**B**), N47A-SH3-A (**C**), N47A-SH3-DA (**D**) diluted to 120 μM in the aggregation buffer, at 20 ºC.

Previous studies on the N47A Spc-SH3 performed at concentrations as low as 1.4 mg mL^{-1} resulted in a hydrodynamic radii (R_h) of 1.6 nm (349). This value was then assigned to a compact monomer according to Wilkins and colleagues for a 62 residue protein as the N47A Spc-SH3 domain is (393). The N47A-SH3 (-GSGSGC tagged) version used in this work is 6 residues longer, however, according to Wilkins and colleagues (394) should result again in an $R_h \approx 1.6$ nm. In our case, the unlabeled protein gives R_h values of 1.8 and 2.3 nm (Table 5.1).

According to Wilkins and colleagues, 1.8 nm could be assigned to an expanded monomeric form but not to a compact dimer ($R_h \approx 1.98$ nm), and this expanded species could be a necessary step prior oligomerization, as previously stated by Morel and colleagues (349). Following the same rationale, 2. nm cannot be assigned to either an unfolded monomer ($R_h \approx 2.45$ nm) or a compact dimer ($R_h \approx 1.97$ nm). Instead, this R_h value would be more consistent with an expanded dimeric species.

Table 5.1 Mass percentages of the different populations from DLS experiments. An apparent hydrodynamic radius (R_h) of ≈ 1.7 nm corresponds to the monomeric native protein, and Rh ≈ 2.4 refers to the dimeric form.

	Conc (μM)	Monomers	Dimers	Oligomeric forms	Larger aggregates
R_h (nm)	-	1.7	2.4	3-20	>20
N47A-SH3	60	74	26	-	-
	120	33	63.4	2.6	1
N47A-SH3-D	60	-	29.7	42.4	0.8
	120	-	-	86	14
N47A-SH3-A	60	40.8	21.5	36.8	0.9
	120	-	-	64	36
Mixed N47A-SH3-D and N47A-SH3-A (1:1 ratio)	60	-	44.4	54.1	1.5
	120	-	-	83	17

Moreover, from the DLS experiments emerged that the presence of the labeling dyes appears to accelerate the process of the oligomerization. This effect was more evident in the N47A-SH3-A construct, as inferred by the fact that the species with $R_h \approx 1.8$ nm were only present in N47A-SH3-A samples at low concentration (Figure 5.8C) but completely vanishes for the N47A-SH3-D variant or the two variants mixed at both concentrations tested (Figure 5.8B and 5.9B).

5.6 DISCUSSION

Our SMF-PIE experiments revealed small, mostly dimeric species, which exhibit a loose and dynamic molecular organization. Interestingly, the most striking feature of this result was that these low-FRET oligomers were formed at the initial stage of the oligomerization reaction, even in absence of incubation and even when the samples were mixed at low concentration. This result was surprising and important because these early oligomers may constitute the first step of amyloidogenic nucleation, and therefore, they might be a potential target in the first line of therapy against amyloidogenic nucleation. Thus, an important aspect to consider about these oligomers is their actual role on the mechanism of amyloid fibril formation, i.e., whether they are on- or off-pathway aggregates. Unlike other amyloidogenic proteins, such as transthyretin, that form oligomers in the native state, and these oligomers must be broken apart before amyloid aggregation, the N47A-SH3 is a monomer in the native state (395). We have detected the formation of these rapidly forming aggregates just when the protein is solved at mild acidic pH and high salt conditions (0.10 M NaCl). Previous aggregation studies on the N47A-SH3 mutant, although performed at higher protein concentrations than ours, have suggested that the bulk aggregation kinetics can be explained through rate-

limiting step involving a rapid pre-equilibration of the conformational opening of the monomer and subsequent oligomerization (339). Our work sheds some light into this mechanism, because our results are directly showing this suggested mechanism at the molecular level.

Moreover, a recent computational study, by Saric and colleagues (186), has highlighted the necessary two-step nucleation mechanism for amyloid fibril formation at physiological concentration regime, such as the micromolar concentrations employed herein. This protein concentration range is much lower than traditional studies on amyloid fibril formation. According to these authors, at low peptide concentrations, nucleation can only be achieved through nonspecific attractions between soluble monomers, in a two-step nucleation process. Our results importantly direct experimental support for this model, through the use of an extremely sensitive SMF technique to overcome the challenge of the low concentrations of the formed oligomers. Our technique not only provides direct detection of the oligomers, but also is able to obtain quantitative thermodynamic information on them.

After the presence of genuine oligomers prior to efficient amyloidogenic incubation was validated, more detailed and quantitative information was obtained from the SMF-PIE experiments. This dynamic population of aggregates allowed for the oligomer formation equilibrium to be studied by modifying the protein concentration as the critical parameter in the association equilibrium of the oligomers. From the relationship between Q, which is proportional to the amount of oligomers found in solution, and the total protein concentration, we are able to obtain quantitative information on the oligomerization equilibrium and these results serve as a tool to quantify the equilibrium constant. The association quotient, Q, exhibited concentration dependence (Figure 5.10), which clearly

indicates the establishment of a dynamic equilibrium between the monomers and aggregates.

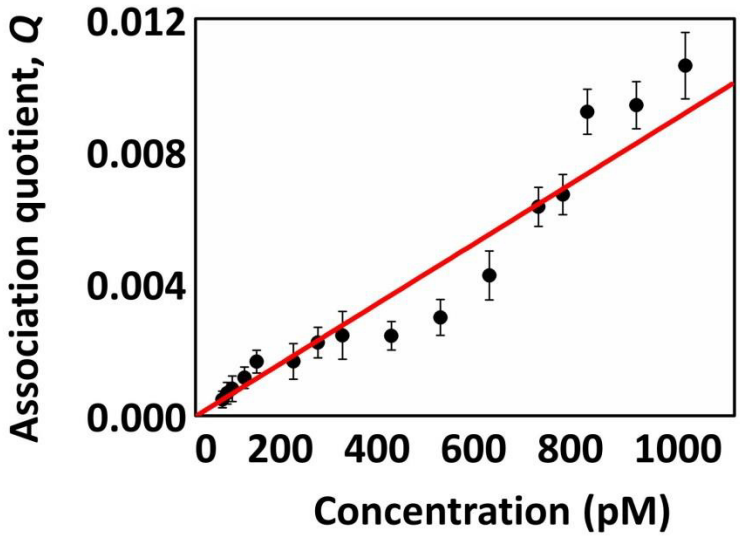


Figure 5.10 Association quotient, *Q*, *versus* the total labeled protein concentration. The error bars correspond to the propagation of the uncertainties in the parameters involved in the estimation of *Q*, and the experimental data were fitted to a straight line.

This equilibrium is extremely important for the amyloid fibril formation process because it represents the first step of amyloidogenic aggregation. Notably, the two labeled proteins had been mixed at very low concentrations; therefore, any aggregate formed that includes both fluorescent tags was produced during the time course of the measurement. Intuitively, and as discuss in 5.2, most of the dual-labeled oligomers must be dimers because the probability of forming higher-order species is very low. Therefore, using the *Q* values, we can estimate the value of the

association equilibrium constant, K_A, for the formation of the dimers, i.e., the first step of amyloidogenic aggregation. K_A can be defined as follows:

$$K_A = \frac{[D]}{[M]^2} \quad (\mathbf{5.3})$$

where [D] and [M] represent the concentrations of dimers and monomers, respectively.

As previously described, (313, 358) the association quotient, Q, is proportional to the fraction of molecules containing the two fluorescent tags. In this case, as the number of aggregates is low,

$$Q \propto \frac{[D]}{[M]} \quad (\mathbf{5.4})$$

The proportionality factor is determined by the efficiency of the instrumental detection of coincident fluorescence bursts. We determined this detection efficiency using 100% dual-labeled 45-basepair dsDNA, which yielded a detection efficiency of 18.6 ± 0.7% (as described in the Methods section). Therefore, the conversion factor to obtain the fraction of dual-labeled molecules from Q is 5.4 ± 0.2. Furthermore, another important correction must be accounted for: only dimers containing both A488 and A647N can be detected by our method. The probabilistic binomial coefficients indicate that these dimers are only 50% of the total dimers, as there is a 25% probability of finding a dimer with two A488-labeled monomers and a 25% probability of finding a dimer with two A647N-labeled monomers. Therefore, we must also consider the fraction of dimers that cannot be detected by adding a factor of 2. Considering both of these corrections, the proportionality factor can be inserted into equation (5.4), as follows:

$$\frac{[\mathrm{D}]}{[\mathrm{M}]} = 2 \cdot 5.4 \cdot Q \qquad (\mathbf{5.5})$$

Using equations (5.3) and (5.5), we can then establish the following relationship between Q and K_A:

$$Q = \frac{K_A \cdot [M]}{2 \cdot 5.4} \qquad (5.6)$$

Equation (5.6) suggests that the plot of Q *versus* the total monomer concentration should follow a straight line. Because our experimental results fit this model very well (Figure 5.10), the apparent association equilibrium constant K_A can be recovered from the slope of the linear fit. The slope of the plot in Figure 5.10 is $(9.2 \pm 0.7) \times 10^{-6}$ pM^{-1}, which provides a value for K_A of $(9.9 \pm 0.8) \times 10^7$ M^{-1} (the associated error was obtained from the error propagation of the uncertainties in the slope and the detection efficiency factor).

It is sometimes more intuitive to consider the dissociation equilibrium (the inverse reaction), which has a corresponding equilibrium constant, K_D, defined as $1/K_A$. The K_D value obtained was 10.1 ± 0.9 nM, which clearly indicates that the protein tends to aggregate at concentrations above 10 nM. This result accounts for the formation of larger oligomer aggregates observed using macroscopic techniques and reveals that the aggregates formed in the first step of amyloid aggregation are highly stable. The value for K_A that we found is much larger (2 orders of magnitude) than that reported by Ruzafa and colleagues (339) for the N47A-SH3 Spc-SH3. These authors analyzed aggregation kinetics from classic bulk experiments (ThT and ANS fluorescence, CD, DLS,…) through a simple mechanistic model involving equilibrium between the native, the unfolded, and a partially folded intermediate monomer, along with an aggregation/elongation step, governed by the equilibrium constant K_A. Although fitting of the experimental data

to this model was sometimes overparameterized, and impossible to converge without fixing some of the equilibrium constants involved, as mentioned by the authors, they managed to recover a value for K_A of (2.5 ± 0.5) × 10^5 M^{-1}. The reasons behind this apparent disagreement between parameters that ought to be comparable may lie on the differences in the concentration regimes employed in either one of the studies. At the high protein concentrations used by Ruzafa and colleagues (around 1 mM) the association events may be dominated by amyloid-nucleation, in a one-step nucleation mechanism. In fact, Ruzafa and colleagues proposed a one-step nucleation kinetic model for fitting the experimental data. Therefore, the value for K_A estimated by these authors may represent the association equilibrium when aggregation is governed by single-step nucleation. In our work, we focus on individual association events that take place at sub-nM concentrations, detected at the molecular level. We cannot rule out that the presence of the dyes contributes to enhancing the value of K_A, as the number of high-order species is larger for the labeled constructs than is for the unlabeled protein, as seen in the DLS experiments (Table 5.1). Nevertheless, the detection of such aggregation events at extremely low concentrations leading to subsequent fibril formation highlights the importance of these contacts in the amyloid aggregation process at physiological concentration regimes, and supports a two-step nucleation mechanism.

Despite performing at different concentration ranges, the results obtained with the ensemble techniques (SDS-PAGE and DLS) were fully consistent with those obtained at the single-molecule level. Importantly, whereas dimeric aggregates were primarily observed in the SM experiments, larger aggregates are also observed when using macroscopic methods, i.e., at higher concentrations. This is consistent with the concentration dependence of the aggregation mechanism, as

expected for any amyloidogenic process and it will be studied in depth in the next chapter.

6. KINETIC STUDY OF THE AMYLOIDOGENIC OLIGOMERS

6 – KINETIC STUDY OF THE AMYLOIDOGENIC OLIGOMERS

In chapter 4 we have observed that the first stages of the N47A-SH3 aggregation are populated by heterogeneous oligomers. Through SMF-PIE experiments at physiologically relevant protein concentrations, we were able to characterize different oligomer types at the single-molecule level. According to their sizes and compactness, we have identified three different oligomer types, and we have classified them as type **1**, **2** and **3**. In chapter 5 we have thermodynamically characterized the rapidly formed type **1** oligomers, finding evidence of non-specific interactions yet in the absence of incubation (386). Herein we focus on the study of the time evolution of all oligomer types detected. Using SMF-PIE techniques, well supported by other bulk biophysical techniques, the oligomers will be quantified in order to perform a full kinetic study of their formation and interconversion reactions. This will improve our molecular-level knowledge of the oligomers populating the early stages of the aggregation and the mechanisms of the molecular processes in which they are involved.

6.1 SMF CHARACTERIZATION OF AMYLOIDOGENIC OLIGOMERS

In order to study the formation of the N47A-SH3-DA oligomers at the single molecule level, we incubated N47A-SH3-DA samples at concentrations ranging from 20 to 44 µM, at 37 ºC in the aggregation buffer (to remind, pH 3.2, 0.10 M NaCl, 0.10 M Gly); aliquots at different time points were extracted and analyzed by SMF-PIE. Our study of oligomer formation focuses only on the aggregated

species, those containing at least one A488- and one A647N-labeled monomer. By applying adequate thresholds to the single molecule fluorescence traces, individual oligomers can be very sensitively detected, even in the presence of a large excess of monomers (148). As explained in chapter 4, the single-molecule correlograms of oligomer size *versus* τ_{A488} (Figure 6.1A) permitted to identify different populations of oligomers. The apparent oligomer size was obtained for each oligomer event through a simple approach. The total fluorescence intensity, due to the sum of fluorescence intensity of the directly excited donor and the acceptor, was converted to the number of monomer molecules, through the normalization with the average fluorescence intensity of a monomer (148) (see Methods). Importantly, the analysis of the apparent oligomer size, corrected for the first time by the quenching on the donor fluorescence, allowed to quantify the relative contribution of each type of oligomers formed at different incubation times. From these correlograms, we classified the different single-molecule events in three regions, and set up oligomer types **1**, **2**, and **3** (Figure 6.1.B). Using the number of events under each region, the relative amount of each oligomer type can be estimated and follow their variation with the incubation time (Figure 6.1C). Given the FRET efficiency values of the detected oligomers, we can hypothesize that type **1** corresponds to loosely packed aggregates (386), whereas type **2** oligomers are more compact, and hence, they exhibit a high FRET efficiency, as donor and acceptor dyes remain close to each other. Type **3** oligomers represent the natural growth of type **2** into higher order species.

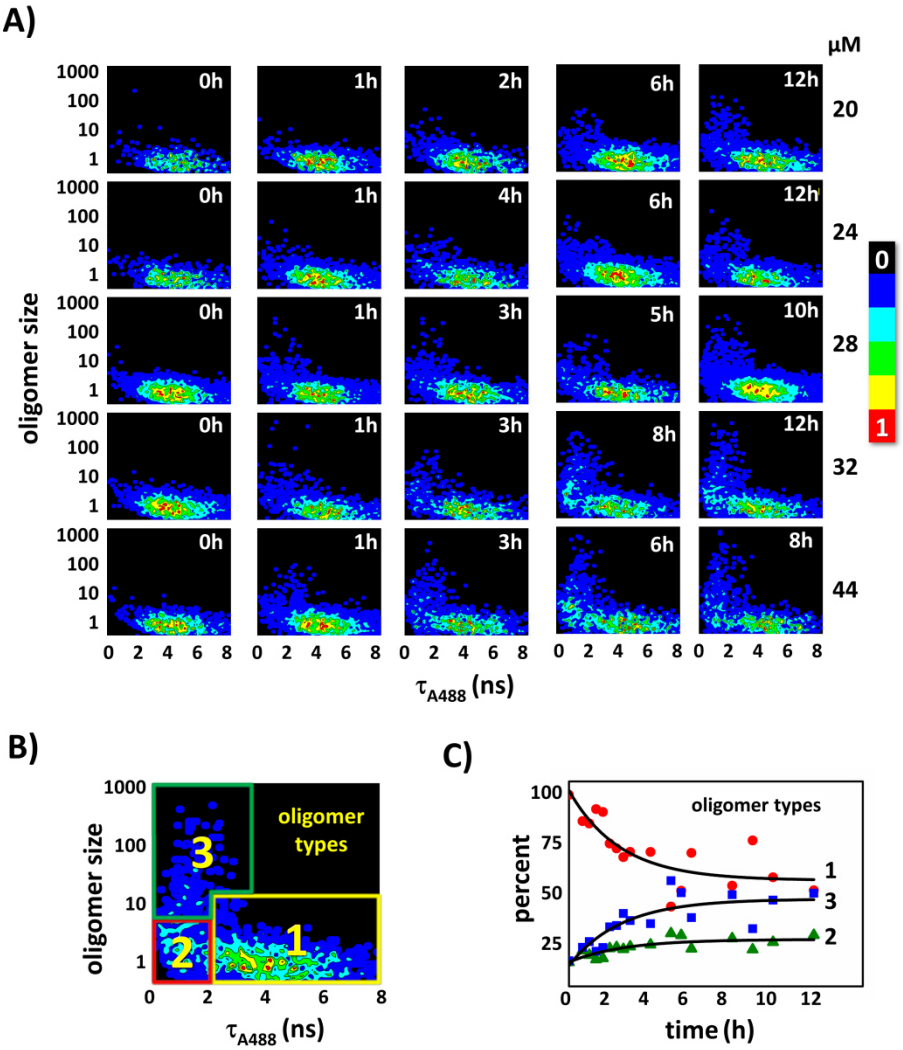


Figure 6.1 A: Single-molecule size *versus* τ_{A488} correlograms of N47A-SH3-DA samples incubated at different concentrations (20, 24, 28, 32 and 44 μM) in the aggregation buffer at 37 ºC. Aliquots were taken out and analyzed at the indicated time points. **B:** Definition of the three types of oligomers, **1, 2**, and **3**. **C:** Time trace of the relative populations of each type of oligomer (**1**, **2**, and **3**) from incubated samples at 32 μM.

At the highest concentrations studied (32 and 44 μM), the appearance of types **2** and **3** was evident even after 1 h of incubation. For this samples, the oligomerization process is so fast that, after 8 h of incubation, the predominant forms detected were mostly highly organized (Figure 6.1A). The samples incubated at lower concentrations (20 and 24 μM) exhibited a slower growth of the compact oligomer forms. The reduced development of type **2** and **3** oligomers at these concentrations may indicate the need for active encounters with type **1** aggregates to form the compact aggregates, in a two-step nucleation mechanism, in clear contrast with a conformation rearrangement mechanism **1** → **2**. Our SMF technique provides an excellent tool to follow the kinetics of the different type of aggregates (Figure 6.1C), and hence, we can extract unique information, infer molecular changes occurring during amyloid growth, and unravel critical mechanistic steps of amyloid aggregation, as discussed in the next section.

6.2 KINETIC ANALYSIS OF THE TWO-STEP AMYLOID AGGREGATION

We performed the kinetic analysis of the reactions of disappearance of the type **1** oligomers and formation of types **2** and **3** at the single molecule level by using the method of initial rates. This method permits to infer the order of reaction by analyzing the variation the change in the initial rate of reaction at different protein concentrations. By fitting the experimental data to single exponential decay functions (Figure 6.2A), we obtained the initial rate of the type **1** disappearance and of the types **2** and **3** formation as the derivative of the relative population of oligomer with respect to time, at time = 0 ($v_0 = \left| d\left[i\right]/dt \right|_{t=0}$), where [*i*] is the concentration of each type of oligomer, **1**, **2**, or **3**) (314).

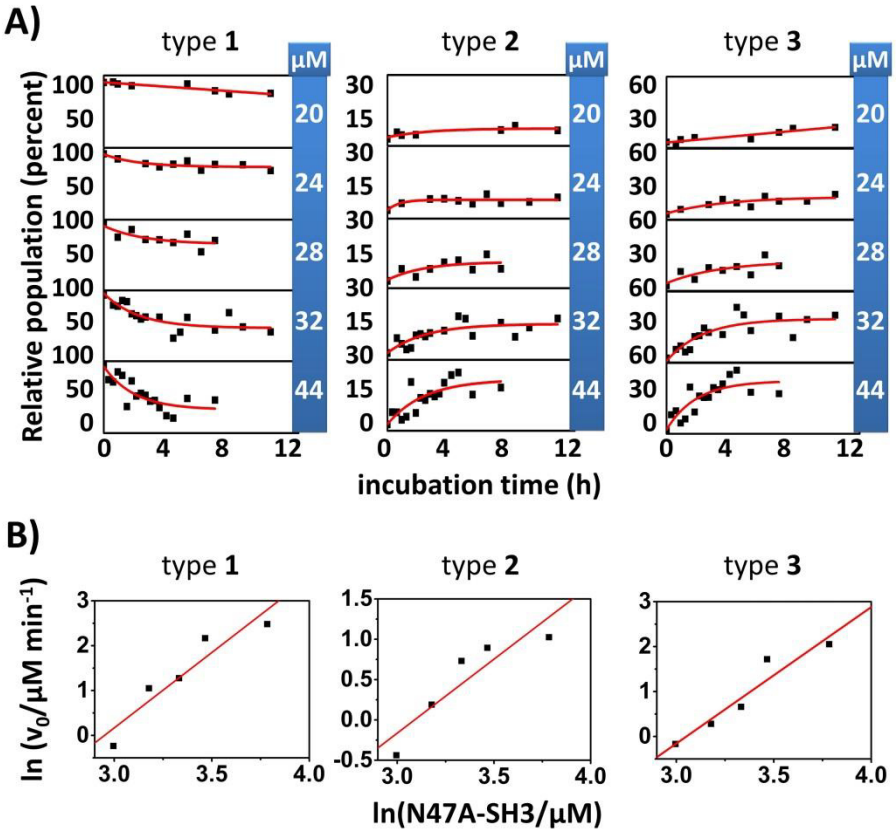


Figure 6.2 Kinetic study of each type of oligomer by the initial rates method. **A:** Time evolution of the relative population of type **1**, **2**, and **3**, obtained from the SMF-PIE data. Red lines represent the fits using single exponential functions. **B:** Initial rates, v_0, of disappearance (type **1**) and formation (types **2** and **3**) of pre-amyloid oligomers as a function of the total protein concentration, in logarithmic scale, to determine the order of reaction. Red lines represent the linear fits.

The apparent order of reaction was obtained from the slope of the plot, in logarithmic scale, of the initial rates *versus* the total protein concentration ($\ln v_0$ *versus* ln[N47A-SH3-DA], Figure 6.2B). The initial rates and the apparent values of the reaction order are compiled in Table 6.1.

Table 6.1 Initial rates (absolute values) of disappearance (type **1**) and formation (types **2** and **3**) of the different types of oligomers, and the determined reaction order calculated by employing the initial rates method.

Oligomer type		Concentration of N47A-SH3-DA (μM)					Apparent reaction order
		44	32	28	24	20	
1	Initial rate ($\mu M\ min^{-1}$)	0.51	0.33	0.28	0.12	0.045	**3**
2		0.11	0.081	0.065	0.051	0.022	**1.8**
3		0.36	0.23	0.18	0.055	0.032	**3**

The apparent value of the order of reaction was 2.9 ± 0.65 for the disappearance of the type **1** oligomers, whereas it was 1.8 ± 0.4 and 3.2 ± 0.7 for the formation of types **2** and **3**, respectively. The effect of concentration is much marked in the disappearance of oligomers **1** than in the formation of type **2**. This clearly indicates that these two events are not totally correlated, and hence type **1** can also disappear by formation of other species (type **3)**. The formation of type **2** oligomers shows a reaction order between 1 and 2. This indicates certain contribution of a conformation rearrangement, which would be order 1, but also the need for second order encounter of different species. Importantly, the formation of oligomers type **3** exhibited a marked concentration-dependence. This illustrates the importance of the nucleation step in the growth of pre-fibrillar aggregates, and will be discuss later.

6.3 FLIM-PIE EXPERIMENTS

Our FLIM-PIE experiments were focused on the study of the formation of the three types of oligomers, its time evolution, and its dependence with protein concentration. Using the donor lifetime images, the dynamics of the N47A-SH3-DA aggregation at different incubation times and concentrations were studied. The A488 FLIM images provided information about the FRET efficiency within individual aggregates, i.e. a decrease in the donor lifetime implies an increase in the FRET efficiency due to the presence of bigger and more compact aggregates. Significantly, the A488 FLIM images clearly showed two important features: the presence of small, high-FRET aggregates, and the growth into large aggregates. The presence of heterogeneous oligomers is more evident as incubation proceeds, and as expected, is protein concentration dependent. At the lowest concentration tested (Figure 6.3A), we only observe the presence of low-FRET oligomers, which shared similar fluorescence intensities even at longer incubation times. This lowest concentration of protein seems to favor a prolonged lag phase, where small-size aggregates are the predominant species formed, so that the probability of focusing on higher-size aggregates would be less. At 24 μM of total protein concentration (Figure 6.3B), together with the low-FRET oligomers detected at early times of incubation, we observe the presence of another type of oligomers after longer times. These oligomers have different features than the first ones detected. Specifically, even though there are not much relevant changes in their fluorescence intensities, these oligomers show lower values of τ_{A488} than those present at the early times of incubation. These low-FRET and low-intensity oligomers might be related to the type **2** oligomers. Interestingly, at the higher protein concentration studied (Figure 6.3C, D and E), we observed the formation of a third types of

oligomers. They exhibit high-intensity and τ_{A488} < 1 ns, which reveals their compactness, so that we cold associate them with the type **3** oligomers.

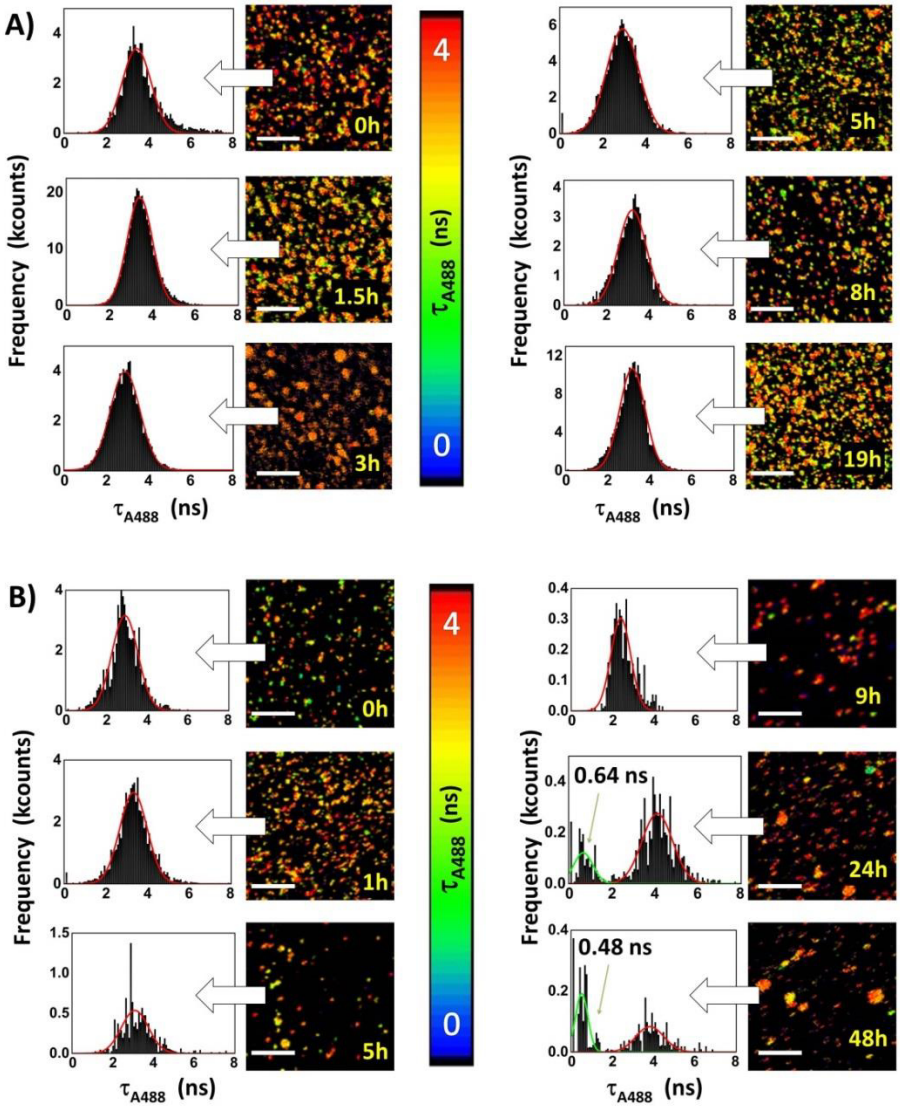


Figure 6.3 A, B: Donor FLIM images and its corresponding lifetime distributions of N47A-SH3-DA incubated at 20 (**A**) and 24 (**B**) µM in the aggregation buffer at 37 °C. White scale bars: 2.4 µm. Color scale: 0 to 4 ns.

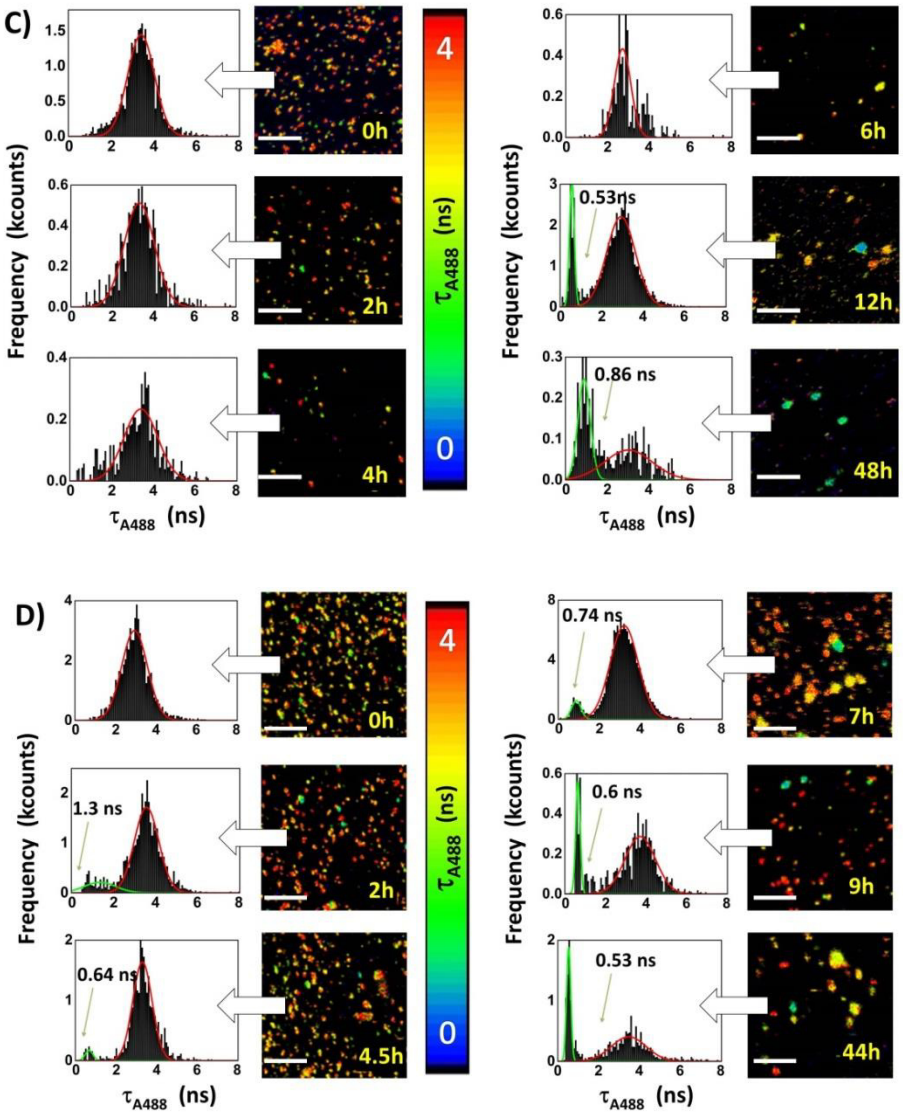


Figure 6.3 C, D: Donor FLIM images and its corresponding lifetime distributions of N47A-SH3-DA incubated at 28 (**C**) and 32 (**D**) μM in the aggregation buffer at 37 °C. White scale bars: 2.4 μm. Color scale: 0 to 4 ns.

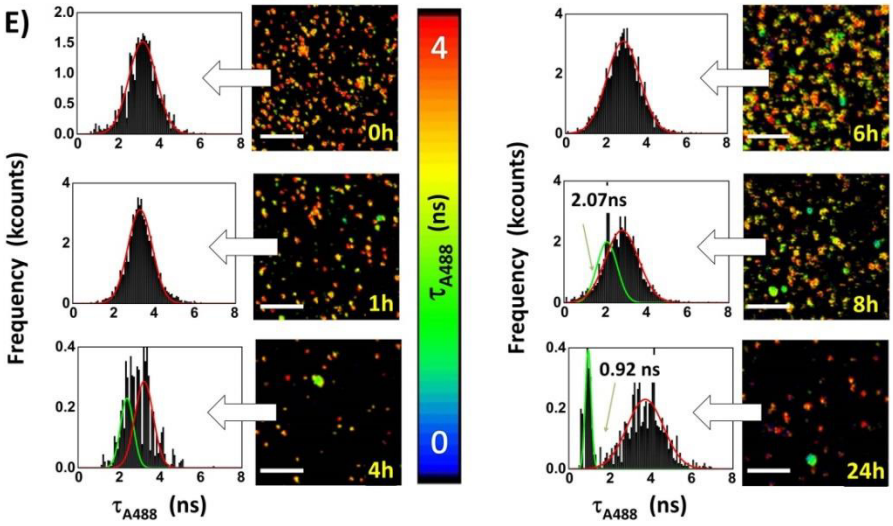


Figure 6.3E Donor FLIM images and its corresponding lifetime distributions of N47A-SH3-DA incubated at 44 µM in the aggregation buffer at 37 °C. White scale bars: 2.4 µm. Color scale: 0 to 4 ns.

We also explored the time evolution of the τ_{A488} within the oligomers detected in the FLIM images (Figure 6.4). For this analysis we obtained the average τ_{A488} of all the pixels containing oligomer events (simultaneously detected in the A488, FRET and A647N images). The final τ_{A488} was obtained as an average from at least 10 different images. The analysis of the time evolution of the average τ_{A488} of the oligomers showed a concentration-dependent decrease with the incubation time. The time dependence of τ_{A488} in Figure 6.4 was fitted to single-exponential decay curves to extract the apparent decay times and rate constants. The decay times were $(3.0 \pm 0.7) \times 10^3$ min, $(1.1 \pm 0.3) \times 10^3$ min, and $(0.5 \pm 0.2) \times 10^3$ min, for incubations at 24, 28 and 32 µM, respectively. At 20 µM, the decrease in τ_{A488} was such slow that a single-exponential decay curve fit was not possible.

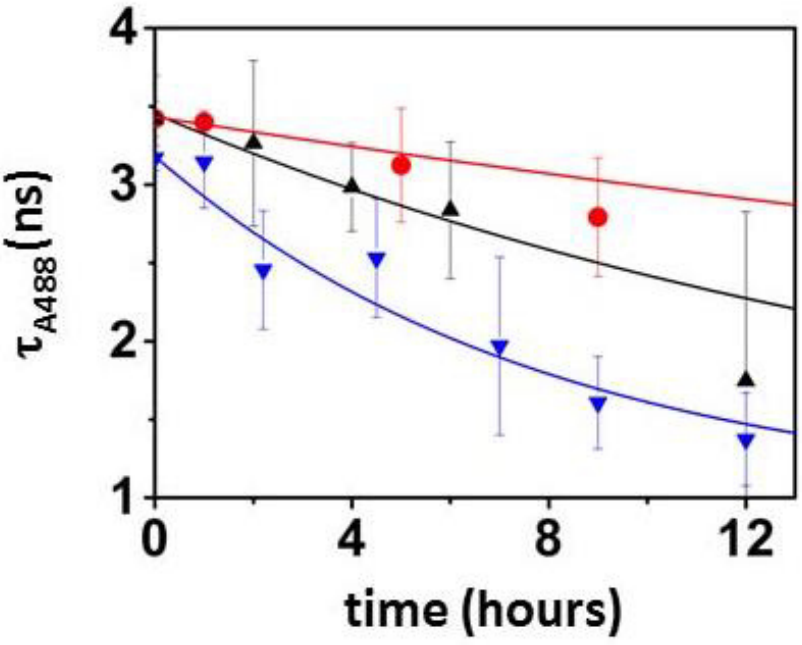


Figure 6.4 Average τ_{A488} values in the oligomers *versus* incubation time from FLIM images of aliquots of the N47A-SH3-DA samples at 24 (red), 28 (black), and 32 μM (blue). Error bars represent the standard error of the τ_{A488} values obtained from at least 10 different FLIM images. Lines represent the fits to a single-exponential function.

6.4 ENSEMBLE TECHNIQUES

In order to link the molecular level with the macroscopic level, the kinetics of the N47A-SH3-DA oligomerization was also studied using conventional biophysical techniques. Even though the molecular information that we have obtained through SMF-PIE and FLIM imaging are unique compared to conventional bulk techniques, the latter permit to compare the behaviour of N47A-SH3-DA and the N47A-SH3 amongst themselves.

6.4.1 Far-UV circular dichroism

We studied the changes occurred in the far-UV CD with the aggregation of N47A-SH3-DA (Figure 6.5A), and also of the unlabeled N47A-SH3 (Figure 6.5C). The develop of negative bands at ≈ 215 nm is typical of the formation of β-sheet structures (367), and hence, this feature is frequently detected in the formation of amyloid fibrils. Previous studies have monitored the changes in N47A Spc-SH3 secondary structure, by examining the CD signal at 215 nm during the time course of the aggregation at 37 ºC (201, 339). By far-UV CD we monitored the molar ellipticity at 215 nm (Figure 6.5A, C), in order to estimate the rate of β-sheet formation (Figure 6.5B, D) (339), and hence, to obtain information on the rate of formation of structured amyloid aggregates. The time dependence of the CD signal at 215 nm during incubation at 37 ºC contains information on the rate of formation of structured amyloid aggregates. We estimated the rate of β-sheet formation by fitting this CD time trace to a single-exponential decay function, from which we obtained the apparent decay constant at 215 nm, k_{215}, for both labeled and unlabeled N47A-SH3 (Figure 6.5B, D). The N47A-SH3-DA (Figure 6.5B), incubated at 32 μM of total protein concentration, exhibited a k_{215} of (9.6 ± 0.1)×10^{-4} min^{-1}, whereas the rate of increase of the CD signal at 215 nm (Figure

6.5D) of the N47A-SH3 was of $(2.7 \pm 0.1)\times10^{-4}$ min^{-1}. This lower value of the k_{215} obtained for the N47A-SH3, compared to that obtained for the N47A-SH3-DA, is in agreement with the enhancement of the aggregation rate for the N47A-SH3-DA observed by TEM experiments (see chapter 4), and it was related to the aromaticity introduced by the dyes' moieties (396).

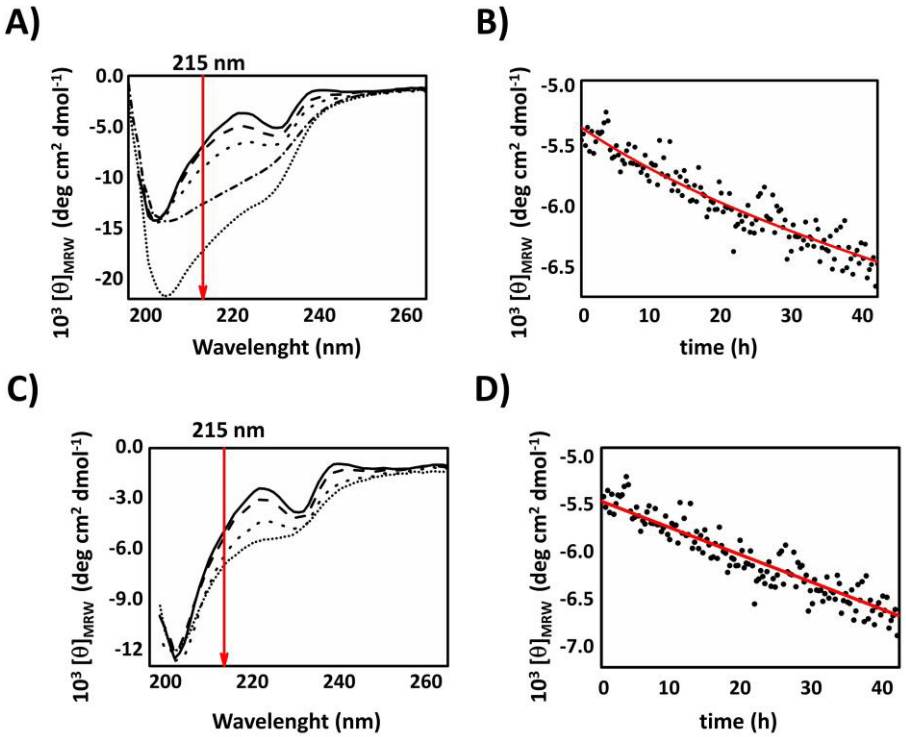


Figure 6.5 A: Far-UV CD spectra of N47A-SH3-DA incubated at 32 μM in the aggregation buffer: at 20 °C (solid line), at 37 °C at "time zero" (dashed line), and after 12 h (dotted line), 48 h (dash-dotted line), and 45 days (short dashed line) of incubation. **B:** Time dependence of the far-UV CD signal at 215 nm. **C:** Far-UV CD spectra of N47A-SH3 incubated at 32 μM in the aggregation buffer: at 20 °C (solid line), and at 37 °C at "time zero" (dashed line), after 12 h (dotted line) and 60 h (short dotted line) of incubation. **D:** Time dependence of the far-UV CD signal at 215 nm.

6.4.2 Dynamic Light Scattering

Our DLS kinetic experiments provided information on the change of the oligomer size distribution during the aggregation of the N47A-SH3-DA mixture at 32 μM. The initial hydrodynamic radius (R_h) value was around 30 nm, significantly larger than the estimated radius of the N47A-SH3 domain compact monomer (1.6 nm) (338, 394), indicating the rapid formation of aggregates type **1**, yet at time 0 (386). The values of R_h and the size distributions (Figure 6.6A, B) displayed a slow growth in the initial hours (up to 6 h), but a cooperative and rapid growth afterwards. These features are in agreement with the fact that type **1** and **2** oligomers share similar sizes, but growth seems to occur only when type **3** aggregates are formed.

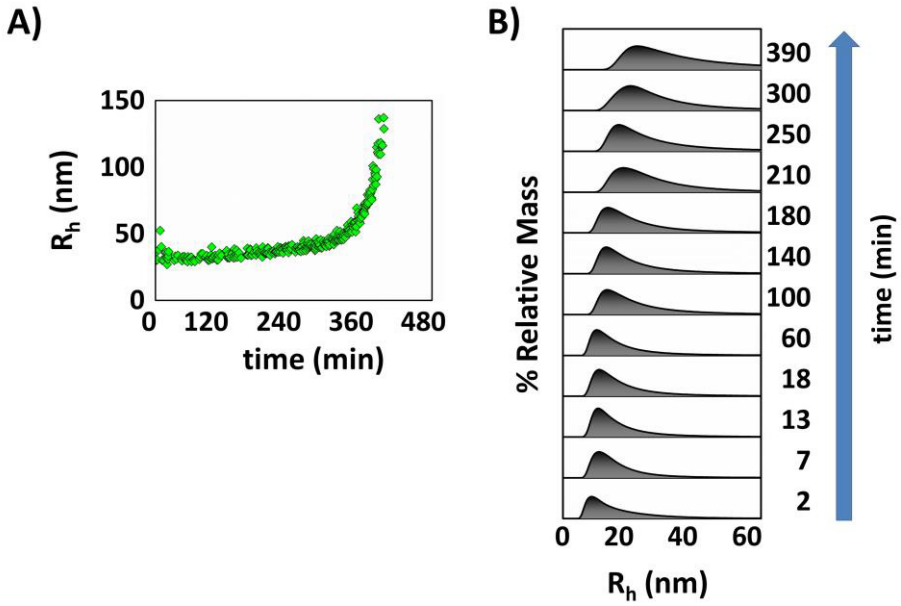


Figure 6.6 A: DLS experiments of N47A-SH3-DA incubated at 37 ºC in the aggregation buffer at 32 μM and its **s**ize distributions (**B**), as a function of time.

DLS experiments were also performed to compare the aggregation process of the unlabeled N47A-SH3 at 60 and 120 μM, whose initial measurements resulted in R_h values for these samples of ≈ 1.8 nm and 6 nm, respectively. At protein concentration of 60 μM, the N47ASH3 showed a prolonged lag phase, of over 24 h (Figure 6.7A). At protein concentration of 120 μM (Figure 6.7A), we distinguish the presence of two different species (Figure 6.7B). Together with smaller oligomers, with R_h values starting around 1.7 nm, which corresponding to a native monomer (338, 386, 394), the solution was populated by large oligomers, which size distribution started from $R_h \approx 7$ nm.

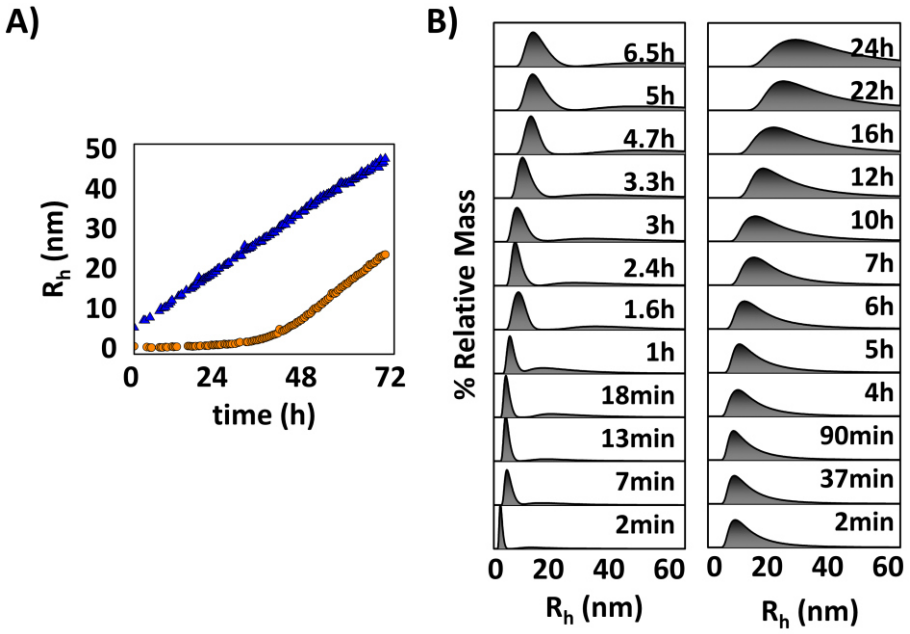


Figure 6.7 DLS experiments of samples of N47A-SH3 incubated at 37 ºC in the aggregation buffer. **A:** Kinetics experiments of N47A-SH3 at 120 (blue diamond) and 60 μM (orange circle). **B:** Size growth of the two aggregate formations visible in the mass distribution of N47A-SH3 at 120 μM.

6.5 OLIGOMER TOXICITY

In order to test the toxicity of the different types of the detected oligomers, we performed cell viability experiment, using 143B culture cells. The human osteosarcoma cell line 143B was obtained from the American Type Culture Collection (ATCC; CRL-8303). Cells were cultured with high glucose Dulbecco's modified Eagle's medium (DMEM), Glutamax^{TM}-I (4 mM), and sodium pyruvate (1 mM) (GIBCO, Life Sciences), supplemented with 1% penicillin/streptomycin (GIBCO, Life Sciences), and 10% fetal bovine serum (FBS) (GIBCO, Life Sciences).

Cell subculture was performed each 3-4 days at 2.4 or 5.6×10^3 cells/cm^2, and media renewal the day before subculture. The impact of protein aggregation on cell viability was studied by using CellTiter Blue™ viability assay (Promega). Cell quadruplicates were plated in cell culture-treated black 96 well optical flat bottom plates at 1.0×10^3 cells/well. After 48h of cell culture, aliquots of 20 μL of protein sample (either N47A-SH3 or N47A-SH3-DA incubated in aggregation buffer) were added directly to the wells. After 24 h of incubation of the cells with the amyloid oligomers, 20% v/v of CellTiter-Blue™ (Promega) reagent was added to the wells, incubated for 2 h at 37 °C, and then fluorescence was directly read at 525/580-640 nm in a Glomax®-Multidetection System (Promega). Untreated cell controls, and wells with reagents only as background controls, were run together with treated cells. The absolute fluorescence arbitrary units were recorded and subsequently the data were expressed at percentage relative to untreated control cells. At least seven independent repetitions of each data point were carried out. Statistical analysis of the cell viability populations were compared against the 100% value of untreated control cells using the Wilcoxon signed-rank test (using Origin 8.5, OriginLab Corp., MA), in order to avoid the requirement of normally

distributed populations. The cell viability experiments were carried out by Dr Miguel Martin and Mar Roldan (GENYO, Pfizer-University of Granada-Junta de Andalucia Centre for Genomics and Oncological Research), and the statistical analysis was performed by Dr Angel Orte (Department of Physical Chemistry, Faculty of Pharmacy, University of Granada). Oligomers collected at different incubation times of N47A-SH3-DA were added for 24 h to the cell cultures, and cell proliferation was measured. For N47A-SH3-DA, incubated at 32 μM and 37 °C in aggregation buffer, aliquots were taken at 0, 1, 2, 4, and 8 h of incubation. For unlabeled N47A-SH3, aliquots were taken after 0, 2, 8, and 24 h of incubation, as the kinetics of aggregation of the unlabeled variant are slower than that of the labeled domains. Figure 6.8A shows that the type **1** oligomers, dynamically formed at time 0, are essentially nontoxic. Once the type **2** oligomers are formed cell viability is decreased, which confirms that this type of aggregates entails the most cytotoxic species. We did not find statistically significant difference in cytotoxicity at later times, when type **3** oligomers are detected. Similar trends in cytotoxicity are found with the oligomers of the unlabeled N47A-SH3 (Figure 6.8B).

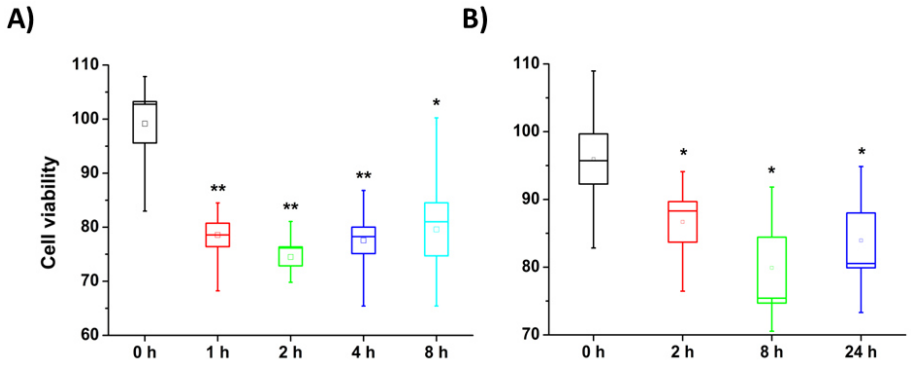


Figure 6.8 Cell viability assays of N47A-SH3-DA (**A**) and N47A-SH3 (**B**) oligomers incubated at 32 μM and 37 ºC. Aliquots of incubated samples were collected after different incubation times and were added to 143B culture cells for 24 h. Cell proliferation was tested, as compared to untreated control cells. Boxes indicate average values ± 1 (standard error), and whiskers indicate the minimum and maximum values of all the repetitions. * Different from untreated control cells with 95% confidence. ** Different from untreated control cells with 99% confidence.

6.6 DISCUSSION

In this chapter we have proposed to elucidate the molecular events that take place during the early stages of the N47A-SH3 aggregation and the role played by the different oligomers that populated those initial phases, aiming to shed some light into the amyloid fibrillation mechanism. Most of the conventional biophysical techniques have the limitation of not being able to examine the entire oligomerization process that leads to the formation of amyloid fibers. This is due to the concentration requirements of the conventional techniques, and this limitation leads to lack of an important step in the study of the amyloid cascade, which is crucial if we consider that the prefibrillar protein oligomer are believed the main toxic agents (397). Likewise, the concentration requirement normally

leads to studies of aggregation of proteins at very high concentrations, usually far from physiologically relevant ranges. Recently, Saric and colleagues demonstrated through *in silico* experiments the essential role of non-specific interactions to promote amyloidogenic aggregation at low, physiological concentrations (186). In this two-step nucleation model for amyloid aggregation at low concentrations, the first step involves formation of contacts for early-oligomers that will act as scaffold for the second, amyloid-prone nucleation. We have exploited the superior features of SMF-PIE and FLIM-PIE experiments for an improved characterization of the intra-oligomer FRET efficiency. The use of the τ_{A488} provides an orthogonal assessment of the FRET efficiency, and permits the unique correction of the oligomer size, taking into account the quenching caused onto the fluorescence of the donor dyes. This advanced characterization of the oligomers has allowed us to unequivocally detect three different types of oligomers, 1, 2, and 3 (Figure 6.1B). Type 1 exhibit a very low FRET efficiency (with $\tau_{A488} \approx$ 3-4 ns, E = 0.00-0.25) and a small size (4-10 monomer units). In chapter 5, we estimated the apparent dissociation equilibrium constant for this type 1 oligomers to be 10.1 ± 0.9 nM (386), demonstrating their high stability at these experimental conditions. The hydrophobicity-driven formation of type 1 is essential for the subsequent rearrangement into type 2 oligomers, which are of similar size than type 1 are, but exhibiting a clear increase in the intra-oligomer FRET efficiency ($\tau_{A488} \approx$ 1-2 ns, E = 0.50-0.75), which supports a more compact structure of the oligomers. Finally, the type 2 oligomers would seem to undergo a subsequent growth into type 3, characterized by large size and high FRET efficiency. The presence of these three populations was also confirmed using surface FLIM imaging (Figure 6.3).

Our study carried out on N47A-SH3-DA completes the work done by Ruzafa and colleagues (19), on the N47A Spc-SH3 aggregate cascade. In their study, the protein concentration was kept rather high (> 0.36 mM) so that the lag phase was

overcome and, kinetically, only a single nucleation step was detected. However, these authors proposed different pathways for homogeneous and heterogeneous nucleation, the latter being the preeminent process at high salt concentration (0.30 M NaCl). In our experiments, the N47A-SH3-DA is in the heterogeneous nucleation regime because of a) aggregation at low concentrations require the hydrophobically-driven first nucleation step; and b) the aromaticity conferred by the pair of fluorophores to the N47A-SH3 is known to enhance aggregation (385, 398), indeed, we have confirmed that these two effects favor the initial formation of type 1 oligomers. Moreover, the heterogeneous nucleation results in more curly fibrils (201, 334), and we found these in both the labeled and the unlabeled versions of the aggregated protein (Figure 4.13) (386)

Importantly, we have obtained quantitative kinetic information on the N47A-SH3-DA aggregation. We have quantified, at the single-molecule level, the rate of formation of the oligomers 2 and 3, and the rate of disappearance of oligomers 1 (Figure 6.2A), and hence, we have obtained valuable kinetic information on the two-step nucleation mechanism. The disappearance of the type 1 exhibits an important concentration-dependent kinetics, corresponding to an apparent third-order reaction. This value could be the result of the cooperative combination of different processes sinking type 1 oligomers, with a rapid disappearance rate and large concentration dependence. The limiting step among these processes is the formation of the type 2, which is comprised of oligomers of high FRET even though their size is comparable to the type 1. The apparent reaction order of the formation of type 2 is between 1 and 2, respect to the total protein concentration (see Table 6.1). This suggests the contribution of a conformation rearrangement such as the spontaneous formation of β-sheet structures, following a first order kinetics, but also the need for second order encounter of different species. This reaction order agrees with the results reported in the kinetic study by Ruzafa and

colleagues (339), using far-UV CD at a protein concentration range of 0.36 – 1.11 mM. Interestingly, we obtained additional kinetic results from other techniques that complement the vision on the formation of type 2 aggregates: on one hand, the kinetics of appearance of high-FRET species followed by FLIM imaging (Figure 6.4); on the other hand, the kinetic results obtained by far-UV CD on the β-sheet structures formation (Figure 6.5B). The value of the apparent decay rate constant for the appearance of β-sheet structures in N47A-SH3-DA aggregation at 32 μM obtained from CD experiments (k_{215} = 9.6×10^{-4} min^{-1}) was of the same order than that of the τ_{A488} decreases obtained by FLIM (k = 20.0×10^{-4} min^{-1} for samples incubated at 32 μM, and k = 9.1×10^{-4} min^{-1} for samples incubated at 28 μM). These kinetics are also in the same time regime than the formation of type 2 oligomers. Significantly, this confirms the correlation between the formation of compact oligomers, and hence high-FRET, and the formation of β-sheet-rich species. Meaningfully, the powerful of SMF-PIE techniques have allowed us to differentiate types 1 and 2 oligomers, despite having the same size, obtaining individual details on their features that bulk techniques are not able to infer.

The formation of type **3** oligomers, characterized by a large size and high FRET, also showed remarkable concentration dependence. The type **3** oligomers are the last one to appear but it has a more cooperative formation rate than the type **2** has. This means that the formation of big oligomers, during the growth phase of the sigmoidal kinetic aggregation process, is not the limiting step, and proceeds easily, once nucleation is finished and sufficient protein concentration is present. The kinetics of this formation corresponds to a third order, indicating a high degree of cooperativity and the presence of various and complex reactions which involve the interaction of species of different size, besides incorporating the other types of oligomers.

The mechanistic model for pre-amyloid aggregation and growth, based on our results, endorses the requirement for two types of nuclei, both type **1** and type **2** oligomers. This sequential two-step model for amyloidogenic nucleation (Figure 6.9) supports with experimental evidence the mechanism suggested by Cohen and colleagues for aggregation at low concentrations, proposed through computer simulations (75). However, from the observation of the initial rate values and reaction orders (Table 6.1) emerged that the kinetic behavior at which type **1** and **3** respectively disappeared and formed, are similar. This controversial result supports the potential existence of a concomitant reaction **1** → **3**, by which the type **3** oligomers may also act as templating agents to recruit type **1** into large aggregates, skipping any conformation rearrangement (Figure 6.10). Due to simultaneousness of these microscopic processes, we cannot affirm whether the type **2** must be already formed or not (Figure 6.10). Further investigation, *i.e.* studying the effect of the inhibition of the **1** → **2** steps on the type **3** formation, might resolve this open question.

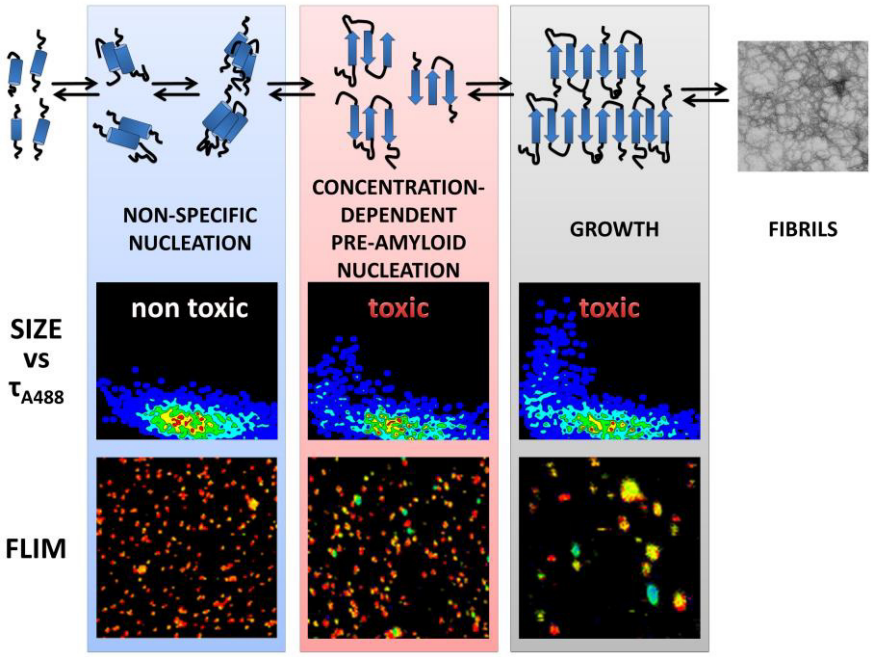


Figure 6.9 Proposed model for two-step nucleation mechanism for amyloid aggregation, correlated with the single-molecule (SMF-PIE and FLIM-PIE) evidence supporting it.

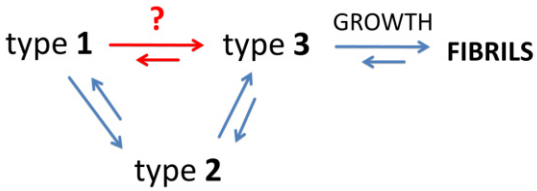


Figure 6.10 Alternative mechanistic scheme for the formation of type **3** aggregates.

With our results, we can now compare the proposed model with previous studies performed at the single molecule level, especially with disease-related proteins. The **1** → **2** conversion step that we observed for the N47A-SH3 oligomers leading the formation of bigger oligomers is also in agreement with the kinetic model for αS aggregation reported by Iljina and colleagues (321). Their model is based on

the master equations for amyloid aggregation proposed by Knowles and colleagues, starting from pure monomers (167, 168, 170, 399). They proposed a first-order reaction for the conversion from low-FRET efficiency into high-FRET efficiency oligomers; so that the initial nucleation step necessarily includes a change in protein conformation. Moreover, Iljina and colleagues also proposed that after conversion, single-monomer units can be recruited by more ordered oligomers, in a succession of elongation step. However, the apparent reaction order that we have obtained for the formation of the type **3** oligomers may include other processes occurring simultaneously to the growth of the type **2** oligomers, *i.e.* the oligomer growth by addition of type **1** to type **3** oligomers, in agreement with the low-FRET oligomers recruitment reported by Iljina and colleagues.

In summary, our results share common mechanistic features with other proposed models for disease-related proteins (75, 186, 321). This evinces the use of N47A-SH3, a protein easy to handle and to produce in large quantities, as a reliable alternative to study and draw conclusions on the general mechanisms of protein aggregation (400). Our results shed some light into the molecular conversions and the growth steps that happen during the macroscopic lag phase, an obscure process for conventional bulk techniques, and represent direct experimental evidence of the two-step nucleation mechanism suggested by Saric and colleagues (186). Likewise, our results confirm the importance of non-specific interactions (type **1** oligomers) for amyloid growth at low concentrations. Finally, cellular toxicity experiments highlighted that type **1** oligomers are essentially nontoxic, whereas toxicity arises in the initial 2 h of aggregation, once the type **2** oligomers appear. No major increase in toxicity was detected with type **3** oligomers, so we can conclude that type **2** oligomers are the most cytotoxic forms. These important findings open up potential targets for inhibition and lessening the toxic effects in disease-related aggregating proteins.

7 – CONCLUSIONS

7 – CONCLUSIONS

Finally, we summarize here the most relevant conclusions of this Thesis project:

1) The addition of the six-residue tail and the fluorescent dye to the N47A-SH3 does not change the native state of the protein and does not alter the ability of the protein to form amyloid fibrils. The shape of these fibrils is curly, like the amyloid fibrils formed by the N47A mutant of the SH3 domain when heterogeneous nucleation is the main mechanism, and this is the case for incubations at low protein concentrations. However, the six-residue tail and the fluorescent dye make the process of aggregation 4 fold faster. This is related to the aromaticity and hydrophobicity introduced by the dyes' moieties.

2) The optimal conditions for the study of the N47A-SH3-DA aggregation were those already reported in previous studies on the N47A Spc-SH3. These were 0.10 M NaCl, 0.10 M Gly, at pH 3.2 and at 37 °C. Under these conditions the protein aggregation exhibits a lag phase populated by different types of oligomers detectably by either conventional or single-molecule fluorescence techniques.

3) The N47A-SH3 oligomers were characterized by single-molecule fluorescence techniques, following the intra-oligomer FRET between donor- and acceptor-labeled monomers and the oligomer size. The use of the donor fluorescence lifetime permitted the unique correction of the oligomer size, taking into account the quenching caused onto the fluorescence of the donor dyes. To the best of our knowledge, this correction has been applied herein for the first time. We unequivocally distinguish three different types of oligomers involved in the earlier phases of the aggregation: the type **1** corresponds to loosely packed, small aggregates; type **2** oligomers are more compact, exhibiting a high FRET efficiency

and small size; and type **3** oligomers represent the natural growth of type **2** into higher order and high-FRET species.

4) We have studied the first dynamic equilibrium of the protein aggregation involving monomers and dimers, with the latter rapidly formed without the need of incubation at high temperature. The value of the equilibrium dissociation constant between monomers and dimers indicates that the protein tends to aggregate at concentrations above 10 nM, revealing that the oligomers formed in the first step of amyloid aggregation are highly stable. This result is an experimental support of the *in silico* experiments that suggested the essential role of non-specific interactions between monomers and small-size, disordered oligomers, in a two-step mechanism for fibril growth at low, physiological concentrations.

5) We have quantified, at the single-molecule level, the rate of formation of the oligomer type **2** and **3**, and the rate of disappearance of oligomers type **1**. The analysis of the initial rates allowed us to extract the order of the oligomer types' formation. For the type **2** oligomers, the reaction order is between first and second order, indicating a combination of conformational conversion but also the necessity of encounters to lead to amyloid nucleation. For type **1** and type **3** oligomers, the reaction order is 3, suggesting a large extent of cooperativity with protein concentration. This correlation may entail that the type **3** oligomers may also act as templating agents to recruit type **1** into large aggregates, skipping any conformation rearrangement. These results shed some light into the molecular conversions and the growth steps that happen during the macroscopic lag phase.

6) The value of the apparent decay rate constant for the appearance of β-sheet structures in N47A-SH3-DA aggregation at 32 μM obtained from CD experiments is of the same order than that of the donor fluorescence lifetime decreases obtained by FLIM. These kinetics are also in the same time regime than the formation of type **2** oligomers. This confirms the correlation between the formation of compact oligomers, and hence high-FRET, and the formation of β-sheet-rich species.

7) Finally, the cellular toxicity experiments showed that the type **1** oligomers are essentially nontoxic, whereas once the type **2** oligomers are formed cell viability is decreased, which confirms that this type of aggregates entails the most cytotoxic species. We did not find statistically significant difference in cytotoxicity at later times, when type **3** oligomers are detected.

FUTURE DIRECTIONS

All our experimental evidence remarks the crucial role of nonspecific interactions for the first nucleation event, and the powerful of advanced single-molecule fluorescence techniques to directly detected this first step, elusive to conventional bulk techniques because of the low concentration regime that governs this process. The use of model proteins, such as N47A-SH3, which are easy to handle and to produce in large quantities, represents an important approach to hypothesize and draw conclusions on the general mechanisms of amyloid aggregation. The multidisciplinary approach and the accurate determination of the oligomer sizes developed on this study of the N47A-SH3 aggregation can be expanded to other disease-related proteins.

The possibility of detecting different types of oligomers involved in the early phases of the aggregation, showing different toxicity features, opens up new ways to identify potential targets for aggregation inhibition and lessening the toxic effects in disease-related proteins. The effect of chaperones, chemical inhibitors, promoters of aggregation, and conformational changes on the monomers or the oligomers can be elegantly obtained using SMFs techniques.

ACKNOWLEDGMENTS

This Thesis project has presented the work carried out during the last 4 years as holder of a PhD studentship within the grant P10-FQM-6154 awarded by the Consejería de Innovación, Ciencia y Empresa (Junta de Andalucía).

COMPLEMENTARY TRAINING ACTIVITIES

A three-month period (September-November 2014) was spent at the University of Florence (Italia) in the research group of Professor Fabrizio Chiti. The main research topic was centered on learning the purification, preparation protocols of mutants of HypF-N containing one cysteine residue, and the labeling of the protein with appropriate fluorophores for FRET analysis. We have labelled the HypF-N mutants obtained with maleimide-modified ATTO 488 (A488, ATTO-TEC GmbH, Siegen, Germany) or ATTO 647N (A647N, ATTO-TEC GmbH, Siegen, Germany), reaching a quantity of protein suitable for FRET analysis, and also carried out a preliminary analysis with the two labelled proteins attained.

RESUMEN

Las amiloidosis y enfermedades relacionadas, como el Alzheimer y el Parkinson, se caracterizan por la presencia en los tejidos de depósitos de agregados insolubles de naturaleza proteica, en forma de fibras amiloides (10). El mecanismo de agregación amiloidea no se encuentra el todo explicado, aunque el más aceptado para explicar la formación de las fibras amiloides consiste en un proceso de nucleación-extensión, con una inicial formación de oligomeros, cuyos crecimiento termina con la formación de fibras maduras (179). Además, experimentos *in vitro* han demostrado que la formación de fibras amiloides es un fenómeno general de cualquier proteína, no solo de aquellas relacionadas con enfermedades, justificando el uso de proteínas modelos en el estudios de la agregación amiloidea (10).

Este proyecto de tesis se centra en el estudio, mediante metodologías de fluorescencia de moleculas individuales (SMF), de las etapas iniciales de formación de las especies oligoméricas pre-fibrilares del mutante N47A del dominio Spc-SH3 (338, 339). El empleo de metodologías de SMF al estudios de las estructuras, de los procesos y de la heterogeneidad de estas especies oligoméricas, ha permitido en los últimos años de afrontar las dificultades ínsitas a sus caracterización frente a técnicas que solo permiten estudiar las moleculas a nivel de conjunto. Debido el desarrollo solo reciente de estas nuevas técnicas, las aplicaciones de SMF en estudios de fibras amiloides, hasta la fecha, son pocas. En esta Tesis nos hemos propuesto un estudio multidisciplinar a los estudios de los oligomeros formados durante la agregación. Este ha incluido el uso de novedosas metodología de fluorescencia de moléculas individuales con un esquema de excitación pulsada alternante (SMF-PIE) (292) y técnica de imágenes de tiempos

de vida de fluorescencia (FLIM-PIE), corroborando y complementando los resultados obtenidos en los estudios de moleculas individuales mediante técnicas biofísicas convencionales como calorimetría diferencial de barrido (DSC), dicroísmo circular (CD), dispersión dinámica de luz (DLS).

La espectroscopia de fluorescencia de moleculas individuales permite la detección ultra-sensible y cuantificación de interacciones entre las moleculas, permitiendo acceder a informaciones sobre la cinética de agregación, el tamaño de oligomeros, y determinaciones de poblaciones relativas de los distintos tamaños. Por primera vez la cuantificación del tamaño de los oligómeros ha sido obtenida considerando la influencia del quenching a la intensidad de fluorescencia del dador, que constituye un avance en el análisis respecto a estudios anteriores (74, 148). Estas informaciones precisas de tamaño de oligomeros y FRET nos han permitido distinguir 3 distintos tipos de oligomeros formados durante la agregación de la proteína etiquetada. Hemos diferenciado un tipo **1** de oligomeros, de baja-FRET y tamaño entre 2-8 monómeros, un tipo **2** comprendiente oligomeros de alta-FRET y tamaño comparable al tipo **1**, y un tipo **3**, cronológicamente posterior a la formación del tipo **2**, oligomeros de alta-FRET y mayor tamaño.

Se ha caracterizado termodinámicamente el primer equilibrio de oligomerización entre monómeros y dímeros, estudiados sin incubar las muestras y realizando los experimentos abaja concentración (pM). El valor de la constante de equilibrio de disociación de los dímeros resultó ser de 10 nM, indicando la estabilidad de estos primeros oligomeros (386). Estos resultados han constituido una consolidación experimental de los modelos teorizados mediante estudio computacionales que han demostrado la importancia de las interacciones no específicas en el mecanismo de formación de fibras amiloides a concentraciones de proteínas fisiológicamente relevantes (186).

El desarrollo del estudio se ha centrado posteriormente en la cinética de oligomerización y se han estudiado los procesos cruciales del mecanismo de formación de los tres tipos de oligomeros identificados. La nueva corrección aportada al cálculo del tamaño aparente de los oligomeros por el tiempo de vida de fluorescencia de los oligomeros, ha permitido la cuantificación individual de los oligomeros, el seguimiento de sus evoluciones con el tiempo y extraer los órdenes de reacción de los procesos claves de la formación de las distintas especies. La cinética de formación de los oligomeros de tipo **2** se ha justificado con la existencia de un cambio conformacional junto con la necesidad de encuentros bimoleculares que precede la evolución hacía especies oligoméricas más organizadas. La correlación de las cinéticas de los oligomeros de tipo **1** y tipo **3**, resultando ser para ambos de tercer orden, abre la posibilidad para los oligomeros de tipo **1** de obviar el cambio conformacional hacia el tipo **2**, actuando los de tipo **3** como plantilla para el crecimiento hacia especies más grande y organizadas. Estudios de citotoxicidad efectuados para los tres tipos de oligomeros han demostrado que las especies más tóxicas son los oligomeros de tipo **2 y 3**, identificando en estos oligomeros un *target* farmacológico para prevenir, al mismo tiempo, la agregación amiloidea y disminuir la toxicidad *per se* de estos oligomeros (74, 401).

En su conjunto, esta Tesis presenta un avance significativo, tanto en la metodología de cuantificación de los oligomeros como en la multidisciplinariedad utilizada en un estudio de formación de fibras amiloides. La identificación de los potenciales agentes implicado en el proceso de agregación y las técnicas de análisis desarrolladas para su precisa caracterización podrían tener importantes aplicaciones en otras proteínas relacionadas con las enfermedades y en el estudio de inhibidores de las primeras etapas del proceso de agregación.

REFERENCES

1. Chen Y*, et al.* (2008) Protein folding: then and now. *Arch Biochem Biophys* 469(1):4-19.

2. Dobson CM (2001) The structural basis of protein folding and its links with human disease. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 356(1406):133-145.

3. Dill KA, Ozkan SB, Shell MS, & Weikl TR (2008) The protein folding problem. *Annu Rev Biophys* 37:289-316.

4. Dobson CM SA, Karplus M (1998) Protein folding: a perspective from theory and experiment. *Angew. Chem. Int. Ed.* 37:868-893.

5. Onuchic JN, Luthey-Schulten Z, & Wolynes PG (1997) Theory of protein folding: the energy landscape perspective. *Annual review of physical chemistry* 48:545-600.

6. Dill KA CH (1997) From Levinthal to pathways to funnels. *Nat Struct Biol* 4:10-19.

7. Knowles TP, Vendruscolo M, & Dobson CM (2014) The amyloid state and its association with protein misfolding diseases. *Nat Rev Mol Cell Biol* 15(6):384-396.

8. Dobson CM (2003) Protein folding and misfolding. *Nature* 426(6968):884-890.

9. Hardy J & Selkoe DJ (2002) The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. *Science* 297(5580):353-356.

10. Chiti F & Dobson CM (2006) Protein misfolding, functional amyloid, and human disease. *Annu Rev Biochem* 75(1):333-366.

11. Eisenberg D, Jucker, M (2012) The amyloid state of proteins in human diseases. *Cell* 148:1188-1203.

12. Haass C & Selkoe DJ (2007) Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer's amyloid beta-peptide. *Nat Rev Mol Cell Biol* 8(2):101-112.

13. Dill KA & MacCallum JL (2012) The protein-folding problem, 50 years on. *Science* 338(6110):1042-1046.

14. Karplus M (1997) The Levinthal paradox: yesterday and today. *Folding & design* 2:69-75.

15. Englander SW & Mayne L (2014) The nature of protein folding pathways. *Proc Natl Acad Sci U S A* 111(45):15873-15880.

16. Anfinsen CB (1973) Principles that govern the folding of protein chains. *Science* 181:223-230.

17. Callender RH, Dyer RB, Gilmanshin R, & Woodruff WH (1998) Fast events in protein folding: the time evolution of primary processes. *Annual review of physical chemistry* 49:173-202.

18. Plaxco KW, Simons KT, & Baker D (1998) Contact order, transition state placement and the refolding rates of single domain proteins. *J. Mol. Biol.* 277:985-994.

19. Levinthal C (1968) Are there pathways for protein folding? *Journal de Chimie Physique et de Physico-Chimie Biologique* 65(1):44-45.

20. Shea JE & Brooks CL (2001) From folding theories to folding proteins: a review and assessment of simulation studies of protein folding and unfolding. *Annu. Rev. Phys. Chem.* 52:499-535.

21. Dill KA & Chan HS (1997) From Levinthal to pathways to funnels. *Nat Struct Biol* 4(1):10-19.

22. Sali A SE, Karplus M (1994) How does a protein fold? *Nature* 369:248-251.

23. Bryngelson JD, Onuchic JN, Socci ND, & Wolynes PG (1995) Funnels, pathways, and the energy landscape of protein folding: a synthesis. *Proteins* 21(3):167-195.

24. Plotkin SS OJ (2002) Understanding protein folding with energy landscape theory. Part I: Basic concepts. *Quarterly reviews of biophysics* 35(2):111-167.

25. Kussell E SJ, Shakhnovich EI (2003) Side-chain dynamics and protein folding. *Proteins* 52:303-321.

26. Onuchic JN NH, García AE, Chahine J, Socci ND (2000) The energy landscape theory of protein folding: Insights into folding mechanisms and scenarios. *Advances in protein chemistry* 53:87-152.

27. Plotkin SS OJ (2002) Understanding protein folding with energy landscape theory. Part II: Quantitative aspects. *Quarterly reviews of biophysics* 35(3):205-286.

28. Oliveberg M WP (2005) The experimental survey of protein-folding energy landscapes. *Quarterly reviews of biophysics* 38:245-288.

29. Leopold PE MM, Onuchic JN (1992) Protein folding funnels: A kinetic approach to the sequence-structure relationship. *Proc Natl Acad Sci USA* 89:8721–8725.

30. Wolynes PG, Onuchic JN, & Thirumalai D (1995) Navigating the folding routes. *Science* 267(5204):1619-1623.

31. Tóth-Petróczy AT, D.S. (2014) The robustness and innovability of protein folds. *Curr. Opin. Struct. Biol.* 26:131-138.

32. Ji XLL, S.Q. (2011) Is stoichiometry-driven protein folding getting out of thermodynamic control? . *J Biomol Struct Dyn* 28:621-623.

33. Dill KA (1999) Polymer principles and protein folding. *Protein Sci* 8(6):1166-1180.

34. Dill KAO, S.B.; Weikl, T.R.; Chodera, J.D. & Voelz, V.A. (2007) The protein folding problem: when will it be solved? . *Current opinion in structural biology* 17:342-346.

35. Griko YV (2000) Energetic basis of structural stability in the molten globule state: alfa-lactalbumin. *J Mol Biol* 297:1259-1268.

36. Baldwin RL (2008) The search for folding intermediates and the mechanism of protein folding. *Annu Rev Biophys* 37:1-21.

37. Kim PS & Baldwin RL (1982) Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding. *Annu Rev Biochem* 51:459-489.

38. al. HMBe (2000) *Nucleic Acids Res.* 28.

39. al. AGMe (1995) *J. Mol. Biol.* 247.

40. al. FAe (2001) *IBM Syst. J.* 40.

41. M. Shirts VSP (2000) *Science* 290.

42. al. CDSe (2002) *Nat. Struct. Mol. Biol.* 420.

43. al. DESe (2008) *Commun. ACM* 51.

44. al. SKe (1993) *Science* 262.

45. Dill KA (1985) *Biochemistry* 24.

46. Moult J (2005) *Curr. Opin. Struct. Biol.* 15.

47. al. JMe (2011) *Proteins* 79.

48. Englander SW & Mayne L (1992) Protein folding studied using hydrogen-exchange labeling and two-dimensional NMR. *Annu Rev Biophys Biomol Struct* 21:243-265.

49. Fersht AR & Daggett V (2002) Protein folding and unfolding at atomic resolution. *Cell* 108(4):573-582.

50. Dinner AR, Sali A, Smith LJ, Dobson CM, & Karplus M (2000) Understanding protein folding via free-energy surfaces from theory and experiment. *Trends Biochem Sci* 25(7):331-339.

51. Duan Y & Kollman PA (1998) Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. *Science* 282(5389):740-744.

52. Cavalli A, Ferrara P, & Caflisch A (2002) Weak temperature dependence of the free energy surface and folding pathways of structured peptides. *Proteins* 47(3):305-314.

53. Fersht AR (1999) Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding. (Nature Publishing Group).

54. Dobson CM, Sali A, & Karplus M (1998) Protein folding: a perspective from theory and experiment. *Angew. Chem. Int. Ed. Eng.* 37:868-893.

55. Hartl FU, Hayer-Hartl, Manajit (2009) Converging concepts of protein folding in vitro and in vivo. *Nat Struct Mol Biol* 16(6):574-581.

56. Hartl FU, Bracher A, & Hayer-Hartl M (2011) Molecular chaperones in protein folding and proteostasis. *Nature* 475(7356):324-332.

57. Hartl FU & Hayer-Hartl M (2002) Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science* 295(5561):1852-1858.

58. J. F (2001) Folding of newly translated proteins in vivo: the role of molecular chaperones. *Annu Rev Biochem* 70:603-647.

59. O'Brien EP CJ, Vendruscolo M, Dobson CM (2011) New scenarios of protein folding can occur on the ribosome. *J Am Chem Soc*:513-526.

60. Gershenson A, Gierasch LM, Pastore A, & Radford SE (2014) Energy landscapes of functional proteins are inherently risky. *Nat Chem Biol* 10(11):884-891.

61. Gershenson A GL (2011) Protein folding in the cell: challenges and progress. *Current opinion in structural biology* 21:32-41.

62. Ebbinghaus S GM (2011) Protein folding landscapes in the living cell. *J Phys Chem Lett* 2:314-319.

63. Bukau B, Weissman J, & Horwich A (2006) Molecular chaperones and protein quality control. *Cell* 125(3):443-451.

64. Hardesty B & Kramer G (2001) Folding of a nascent peptide on the ribosome. *Prog Nucleic Acid Res Mol Biol* 66:41-66.

65. Gething MJ & Sambrook J (1992) Protein folding in the cell. *Nature* 355(6355):33-45.

66. Ellis RJ & Hartl FU (1999) Principles of protein folding in the cellular environment. *Current opinion in structural biology* 9(1):102-110.

67. Koga H KS, Cuervo AM (2011) Protein homeostasis and aging: the importance of exquisite quality control. *Ageing Res Rev* 10:205-215.

68. Kaganovich D KR, Frydman J (2008) Misfolded proteins partition between two distinct quality control compartments. *Nature* 454:1088-1036.

69. Tyedmers J MA, Bukau B (2010) Cellular strategies for controlling protein aggregation. *Nat Rev Mol Cell Biol* 11:777-788.

70. CM D (2004) Principles of protein folding, misfolding and aggregation. *Seminars in Cell & Developmental Biology* 15:3–16.

71. Capaldi AP, Kleanthous C, & Radford SE (2002) Im7 folding mechanism: misfolding on a path to the native state. *Nat Struct Biol* 9(3):209-216.

72. Dobson CM (2003) Protein folding and disease: a view from the first Horizon Symposium. *Nat Rev Drug Discov* 2(2):154-160.

73. Walsh DM*, et al.* (2002) Naturally secreted oligomers of amyloid beta protein potently inhibit hippocampal long-term potentiation in vivo. *Nature* 416(6880):535-539.

74. Cremades N*, et al.* (2012) Direct observation of the interconversion of normal and toxic forms of alpha-synuclein. *Cell* 149(5):1048-1059.

75. Cohen SI*, et al.* (2013) Proliferation of amyloid-beta42 aggregates occurs through a secondary nucleation mechanism. *Proc Natl Acad Sci U S A* 110(24):9758-9763.

76. Campioni S*, et al.* (2010) A causative link between the structure of aberrant protein oligomers and their toxicity. *Nature Chem. Biol.* 6(2):140-147.

77. Bucciantini M*, et al.* (2002) Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature* 416(6880):507-511.

78. Caughey B & Lansbury PT (2003) Protofibrils, pores, fibrils, and neurodegeneration: Separating the responsible protein aggregates from the innocent bystanders. *Annu. Rev. Neurosci.* 26:267-298.

79. Borysik AJ RS, Ashcroft AE (2004) Co-populated conformational ensembles of beta-microglobulin uncovered quantitatively by electrospray ionisation mass spectrometry. *The Journal of biological chemistry* 279:27069-27077.

80. Radford SE GW, Platt GW (2005) Towards an understanding of the structural molecular mechanism of beta-microglobulin amyloid formation in vitro. *Biochim Biophys Acta* 1753:51-63.

81. Eakin CM, Berman AJ, & Miranker AD (2006) A native to amyloidogenic transition regulated by a backbone trigger. *Nat Struct Mol Biol* 13(3):202-208.

82. Surewicz WK JE, Apetri AC (2006) The emerging principles of mammalian prion propagation and transmissibility barriers: insight from study in vitro *Acc Chem Res* 39:654-662.

83. Jahn TR, Parker MJ, Homans SW, & Radford SE (2006) Amyloid formation under physiological conditions proceeds via a native-like folding intermediate. *Nat. Struct. Mol. Biol.* 13:195-201.

84. Bemporad F*, et al.* (2006) Sequence and structural determinants of amyloid fibril formation. *Acc Chem Res* 39(9):620-627.

85. Jahn TR & Radford SE (2008) Folding versus aggregation: polypeptide conformations on competing pathways. *Arch Biochem Biophys* 469(1):100-117.

86. Campioni S ME, Chiti F (2010) Why Protein misfold. *Protein misfolding diseases*, ed Ramirez-Alvarado M KJ, Dobson CM (John Wiley & Sons, Inc, Hoboken, New Jersey), pp 3-20.

87. Walsh DM, Tseng BP, Rydel RE, Podlisny MB, & Selkoe DJ (2000) The oligomerization of amyloid beta-protein begins intracellularly in cells derived from human brain. *Biochemistry* 39(35):10831-10839.

88. Hartley DM*, et al.* (1999) Protofibrillar intermediates of amyloid beta-protein induce acute electrophysiological changes and progressive neurotoxicity in cortical neurons. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 19(20):8876-8884.

89. Xu G, Gonzales V, & Borchelt DR (2002) Rapid detection of protein aggregates in the brains of Alzheimer patients and transgenic mouse models of amyloidosis. *Alzheimer Dis Assoc Disord* 16(3):191-195.

90. Baba M*, et al.* (1998) Aggregation of alpha-synuclein in Lewy bodies of sporadic Parkinson's disease and dementia with Lewy bodies. *Am J Pathol* 152(4):879-884.

91. Conway KA, Harper JD, & Lansbury PT, Jr. (2000) Fibrils formed in vitro from alpha-synuclein and two mutant forms linked to Parkinson's disease are typical amyloid. *Biochemistry* 39(10):2552-2563.

92. Gong Y*, et al.* (2003) Alzheimer's disease-affected brain: presence of oligomeric A beta ligands (ADDLs) suggests a molecular basis for reversible memory loss. *Proc Natl Acad Sci U S A* 100(18):10417-10422.

93. Lee HJ SJ, Bae EJ and Lee SJ (1997) Clearance and deposition of extracellular alpha-synuclein aggregates in microglia. *Biochemical and biophysical research communications* 372:423-428.

94. DJ WDaS (2004) Oligomers on the brain: the emerging role of soluble protein aggregates in neurodegeneration. *Protein Pept Lett* 11:213-228.

95. Knowles TPJ, Vendruscolo M, & Dobson CM (2014) The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell Biol.* 15(6):384-396.

96. Selkoe DJ (2004) Cell biology of protein misfolding: the examples of Alzheimer's and Parkinson's diseases. *Nat. Cell Biol.* 6:1054-1061.

97. Serpell LC (2000) Alzheimer's amyloid fibrils: structure and assembly. *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1502:16-30.

98. Stromer T S, L C (2005) Structure and morphology of the Alzheimer's amyloid fibril. *Microscopy research and technique* 67:210-217.

99. Goldsbury C BU, Simon M N, Steven A C, Engel A, Wall J S, Aebi U, Muller S A (2011) Amyloid structure and assembly: insights from scanning transmission electron microscopy. *Journal of structural biology* 173:1-13.

100. Nilsson MR (2004) technique to study amyloid fibril formation in vitro. *Methods* 34:151-160.

101. Sunde M & Blake C (1997) The Structure of Amyloid Fibrils by Electron Microscopy and X-Ray Diffraction. *Advances in protein chemistry*, eds Frederic M. Richards DSE & Peter SK (Academic Press), Vol Volume 50, pp 123-159.

102. Comellas G RCM (2013) Protein structure determination by magic-angle spinning solid-state NMR, and insights into the formation, structure, and stability of amyloid fibrils. *Annual review of biophysics* 42:515-536.

103. Heise H*, et al.* (2005) Molecular-level secondary structure, polymorphism, and dynamics of full-length alpha-synuclein fibrils studied by solid-state NMR. *Proc Natl Acad Sci U S A* 102(44):15871-15876.

104. Heise H (2008) Solid-state NMR spectroscopy of amyloid proteins. *Chembiochem* 9(2):179-189.

105. Petkova AT (2002) A structural model for Alzheimer's [beta]-amyloid peptide fibrils based on experimental constraints from solid-state NMR spectroscopy. *Proc. Natl. Acad. Sci. USA* 99:16742-16747.

106. Sawaya MR, Sambashivan S, Nelson R, Ivanova M I, Sievers, S. A.; Apostol, M. I.; Thompson M J, Balbirnie, M.; Wiltzius, J. J.; McFarlane, H. T.; Madsen, A. O.; Riekel, C.; Eisenberg, D (2007) Atomic structures of amyloid cross-beta spines reveal varied steric zippers. *Nature* 447:453-457.

107. Nelson R*, et al.* (2005) Structure of the cross-beta spine of amyloid-like fibrils. *Nature* 435(7043):773-778.

108. Carulla N CGL, Hall D R, Zurdo J, Gairi M, Feliz M, Giralt E, Robinson C V, Dobson C M (2005) Molecular recycling within amyloid fibrils. *Nature* 436:554-558.

109. Del Mar C GEA, Mayne L, Englander S W, Woods V L Jr (2005) Structure and properties of alpha-synuclein and other amyloids determined at the amino acid level. *P Natl Acad Sci USA* 102:15477-15482.

110. Ma B & Nussinov R (2011) Polymorphic triple beta-sheet structures contribute to amide hydrogen/deuterium (H/D) exchange protection in the Alzheimer amyloid beta42 peptide. *The Journal of biological chemistry* 286(39):34244-34253.

111. Whittemore N A MR, Kheterpal I, Williams A D, Wetzel R, Serpersu E H (2005) Hydrogen-deuterium (H/D) exchange mapping of Abeta 1-40 amyloid fibril secondary structure using nuclear magnetic resonance spectroscopy. *Biochemistry* 44:4434-4441.

112. AWP F (2013) Atomic structure and hierarchical assembly of a cross-beta amylid fibril. *Proc. Natl Acad. Sci. USA* 110:5468-5473.

113. Tomic JL, Pensalfini A, Head E, & Glabe CG (2009) Soluble fibrillar oligomer levels are elevated in Alzheimer's disease brain and correlate with cognitive dysfunction. *Neurobiol Dis* 35(3):352-358.

114. Selkoe DJ & Abraham CR (1986) Isolation of paired helical filaments and amyloid fibers from human brain. *Methods in enzymology* 134:388-404.

115. Sipe JD (1992) Amyloidosis. *Annu. Rev. Biochem.* 61:947-975.

116. Sipe JD (1994) *Amyloidosis* (CRC Press, Boca Raton.).

117. J K (1998) Alternative conformation of amyloidogenic proteins and their multi-step assembly pathways. *Current opinion in structural biology* 8:101-106.

118. MM R (1998) Genetically determined neuropathies. *J Neurol* 245:6-13.

119. DJ S (2003) Folding proteins in fatal ways. *Nature* 426:900-904.

120. Greenwald J & Riek R (2010) Biology of Amyloid: Structure, Function, and Regulation. *Structure* 18(10):1244-1260.

121. Chapman MR, *et al.* (2002) Role of Escherichia coli curli operons in directing amyloid fiber formation. *Science* 295(5556):851-855.

122. Fowler DM, Koulov AV, Balch WE, & Kelly JW (2007) Functional amyloid--from bacteria to humans. *Trends Biochem Sci* 32(5):217-224.

123. Fowler DM, *et al.* (2006) Functional amyloid formation within mammalian tissue. *PLoS Biol* 4(1):e6.

124. Maji SK, *et al.* (2009) Functional Amyloids As Natural Storage of Peptide Hormones in Pituitary Secretory Granules. *Science* 325(5938):328-332.

125. Mannini B, *et al.* (2014) Toxicity of protein oligomers is rationalized by a function combining size and surface hydrophobicity. *ACS Chem Biol* 9(10):2309-2317.

126. Kayed R, *et al.* (2003) Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis. *Science* 300(5618):486-489.

127. Glabe CG (2009) Amyloid Oligomer Structures and Toxicity. *The Open Biology Journal* 2:222-227.

128. Bemporad F & Chiti F (2012) Protein misfolded oligomers: experimental approaches, mechanism of formation, and structure-toxicity relationships. *Chem Biol* 19(3):315-327.

129. Fandrich M (2012) Oligomeric intermediates in amyloid formation: structure determination and mechanisms of toxicity. *J Mol Biol* 421(4-5):427-440.

130. Kayed R & Lasagna-Reeves CA (2013) Molecular mechanisms of amyloid oligomers toxicity. *Journal of Alzheimer's disease : JAD* 33 Suppl 1(0):S67-78.

131. Jan A, *et al.* (2011) Aβ42 Neurotoxicity Is Mediated by Ongoing Nucleated Polymerization Process Rather than by Discrete Aβ42 Species. *The Journal of biological chemistry* 286(10):8585-8596.

132. Gosal WS MS, Radford SE, Thomson NH (2006) Amyloid under the atomic force microscope. *Protein Pept Lett* 13:261–270.

133. Goldsbury CSWSM, S.A.; Sunderji, S. (2000) Studies on the in vitro assembly of Ab 1-40: implications for the search for Ab fibril formation inhibitors. *J. Struct. Biol.* 130:217-231.

134. Walsh DM*, et al.* (1999) Amyloid beta-protein fibrillogenesis. Structure and biological activity of protofibrillar intermediates. *The Journal of biological chemistry* 274(36):25945-25952.

135. Sipe JD & Cohen AS (2000) Review: history of the amyloid fibril. *J Struct Biol* 130(2-3):88-98.

136. Habicht G*, et al.* (2007) Directed selection of a conformational antibody domain that prevents mature amyloid fibril formation by stabilizing Abeta protofibrils. *Proc Natl Acad Sci U S A* 104(49):19232-19237.

137. Eckert A*, et al.* (2008) Oligomeric and fibrillar species of beta-amyloid (A beta 42) both impair mitochondrial function in P301L tau transgenic mice. *J Mol Med (Berl)* 86(11):1255-1267.

138. Barghorn S, Nimmrich, V., Striebinger, A., Krantz, C., Keller, P., Janson, B. et al. (2005) Globular amyloid β-peptide1–42 oligomer—a homogenous and stable neuropathological protein in Alzheimer's disease. *J. Neurochem.* 95:834–847.

139. Broersen K, Rousseau F, & Schymkowitz J (2010) The culprit behind amyloid beta peptide related neurotoxicity in Alzheimer's disease: oligomer size or conformation? *Alzheimers Res Ther* 2(4):12.

140. Haupt C, Leppert, J., Rönnike, R., Meinhardt, J., Yadav, J. K., Ramachandran, R. et al. (2011) Oligomer structure provides a basis for preventing Aβ-dependent synaptic dysfunctions. *Angew. Chem., Int. Ed. Eng.*).

141. Habicht G*, et al.* (2007) Directed selection of a conformational antibody domain that prevents mature amyloid fibril formation by stabilizing Aβ protofibrils. *Proc Natl Acad Sci U S A* 104:19232–19237.

142. Sunde M & Blake CC (1998) From the globular to the fibrous state: protein structure and structural conversion in amyloid formation. *Quarterly reviews of biophysics* 31(1):1-39.

143. Chimon S*, et al.* (2007) Evidence of fibril-like beta-sheet structures in a neurotoxic amyloid intermediate of Alzheimer's beta-amyloid. *Nat Struct Mol Biol* 14(12):1157-1164.

144. Ahmed M, *et al.* (2010) Structural conversion of neurotoxic amyloid-beta(1-42) oligomers to fibrils. *Nat Struct Mol Biol* 17(5):561-567.

145. Sandberg A, *et al.* (2010) Stabilization of neurotoxic Alzheimer amyloid-beta oligomers by protein engineering. *Proc Natl Acad Sci U S A* 107(35):15595-15600.

146. Glabe CG (2008) Structural classification of toxic amyloid oligomers. *The Journal of biological chemistry* 283(44):29639-29643.

147. Sakono MZ, T. (2010) Amyloid oligomers: formation and toxicity of Aβ oligomers. *FEBS J* 277:1348-1358.

148. Orte A, *et al.* (2008) Direct characterization of amyloidogenic oligomers by single-molecule fluorescence. *Proc Natl Acad Sci U S A* 105(38):14424-14429.

149. Compton LAaJ, W.C., Jr. (1986) Analysis of protein circular dichroism spectra for secondary structure using a simple matrix multiplication. *Anal Biochem* 155(2):155-167.

150. Woody RW (1995) Circular dichroism. *Methods in enzymology* 246:34-71.

151. Kelly SM, Jess TJ, & Price NC (2005) How to study proteins by circular dichroism. *Biochim Biophys Acta* 1751(2):119-139.

152. Carulla N, *et al.* (2009) Experimental characterization of disordered and ordered aggregates populated during the process of amyloid fibril formation. *Proc Natl Acad Sci U S A* 106(19):7828-7833.

153. Makin OS, Atkins E, Sikorski P, Johansson J, & Serpell LC (2005) Molecular basis for amyloid fibril formation and stability. *Proc Natl Acad Sci U S A* 102(2):315-320.

154. Larson JL KE, Miranker AD (2000) Direct measurement of islet amyloid polypeptide fibrillogenesis by mass spectrometry. *Protein Sci* 9:427-431.

155. Smith AM JT, Ashcroft AE, Radford SE (2006) 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.

156. Fabian H GK, Laue M, Misselwitz R, Uchanska-Ziegler B, Ziegler A, Naumann D (2008) Early stages of misfolding and association of β2microglobulin: insights

from infrared spectroscopy and dynamic light scattering. *Biochemistry* 47:6895-6906.

157. Pecora R (1985) Dynamic Light Scattering: Applications of Photon Correlation Spectroscopy. *Plenum Press*.

158. Lebowitz J, Lewis MS, & Schuck P (2002) Modern analytical ultracentrifugation in protein science: a tutorial review. *Protein Sci* 11(9):2067-2079.

159. Laue TM (2001) Analytical Ultracentrifugation. *Current Protocols in Protein Science*, (John Wiley & Sons, Inc.).

160. Arosio P, Vendruscolo M, Dobson CM, & Knowles TP (2014) Chemical kinetics for drug discovery to combat protein aggregation diseases. *Trends Pharmacol Sci* 35(3):127-135.

161. Oosawa FA, S. (1975) Thermodynamics of the Polymerization of Protein. *Academic Press.*

162. Oosawa F & Kasai M (1962) A theory of linear and helical aggregations of macromolecules. *J Mol Biol* 4:10-21.

163. Ferrone FA, Hofrichter, J., Sunshine, H. R. & Eaton, W. A. (1980) Kinetic studies on photolysis-induced gelation of sickle cell hemoglobin suggest a new mechanism. . *Biophys J* 32:361–380.

164. Ferrone FA, Hofrichter J, & Eaton WA (1985) Kinetics of sickle hemoglobin polymerization. II. A double nucleation mechanism. *J Mol Biol* 183(4):611-631.

165. Bishop MFF, F. A. (1984) Kinetics of nucleation-controlled polymerization. A perturbation treatment for use with a secondary pathway. *Biophys J* 46:631-644.

166. Ferrone F (1999) Analysis of protein aggregation kinetics. *Methods in enzymology* 309:256-274.

167. Knowles TP*, et al.* (2009) An analytical solution to the kinetics of breakable filament assembly. *Science* 326(5959):1533-1537.

168. Cohen SI*, et al.* (2011) Nucleated polymerization with secondary pathways. I. Time evolution of the principal moments. *The Journal of chemical physics* 135(6):065105.

169. Cohen SIA, Vendruscolo M, Dobson CM, & Knowles TPJ (2011) Nucleated polymerization with secondary pathways. III. Equilibrium behavior and oligomer populations. *The Journal of chemical physics* 135(6):065107.

170. Cohen SI, Vendruscolo M, Dobson CM, & Knowles TP (2012) From macroscopic measurements to microscopic mechanisms of protein aggregation. *J Mol Biol* 421(2–3):160-171.

171. Arosio P, Knowles TP, & Linse S (2015) On the lag phase in amyloid fibril formation. *Phys Chem Chem Phys* 17(12):7606-7618.

172. Uversky VN (2002) Natively unfolded proteins: a point where biology waits for physics. *Protein Sci* 11(4):739-756.

173. Uversky VNaF, A.L. (2004) Conformational constraints for amyloid fibrillation: the importance of being unfolded. *Biochim Biophys Acta* 1698:131-153.

174. Chiti F, Stefani M, Taddei N, Ramponi G, & Dobson CM (2003) Rationalization of mutational effects on protein aggregation rates. *Nature* 424:805-808.

175. Adamcik J*, et al.* (2010) Understanding amyloid aggregation by statistical analysis of atomic force microscopy images. *Nature nanotechnology* 5(6):423-428.

176. Ferrone FA HJ, Eaton A (1985) Kinetics of sickle hemoglobin polymerization. 1. Studies using temperature-jump and laser photolysis techniques. *J Mol Biol* 183:591-610.

177. Frokjaer SaO, D.E. (2005) Protein drug stability: a formulation challenge. *Nat. Rev. Drug Discov.* 4:298-306.

178. Oosawa F, Asakura S, Hotta K, Imai N, & Ooi T (1959) G-F Transformation of Actin as a Fibrous Condensation. *J Polym Sci* 37(132):323-336.

179. Rochet JC & Lansbury PT, Jr. (2000) Amyloid fibrillogenesis: themes and variations. *Current opinion in structural biology* 10(1):60-68.

180. Murphy RM & Pallitto MM (2000) Probing the kinetics of beta-amyloid self-association. *J Struct Biol* 130(2-3):109-122.

181. Jarrett JT LP (1992) Amyloid fibril formation requires a chemically discriminating nucleation event: studies of an amyloidogenic sequence from the bacterial protein OsmB. *Biochemistry* 31:12345-12352.

182. Ponder JW CD (2003) Force fields for protein simulation. *Advances in protein chemistry* 66:27-85.

183. Yoda T SY, Okamotoa Y (2004) Comparisons of force fields for proteins by generalized-ensemble simulations. *Chem Phys Lett* 386:460-467.

184. Best RB, Buchete NV, & Hummer G (2008) Are current molecular dynamics force fields too helical? *Biophys J* 95(1):L07-09.

185. Lomakin A. CDS, Benedek G.B., Kirschner D.A. and Teplow D.B. (1996) On the nucleation and growth of amyloid beta-protein fibrils: detection of nuclei and quantitation of rate constants. *Proc. Natl Acad. Sci. USA* 93:1125-1129.

186. Saric A, Chebaro YC, Knowles TP, & Frenkel D (2014) Crucial role of nonspecific interactions in amyloid nucleation. *Proc Natl Acad Sci U S A* 111(50):17869-17874.

187. Lomakin A, Teplow DB, Kirschner DA, & Benedek GB (1997) Kinetic theory of fibrillogenesis of amyloid beta-protein. *Proc Natl Acad Sci U S A* 94(15):7942-7947.

188. Morris AM, Watzky MA, & Finke RG (2009) Protein aggregation kinetics, mechanism, and curve-fitting: a review of the literature. *Biochim Biophys Acta* 1794(3):375-397.

189. Di Fede G*, et al.* (2009) A Recessive Mutation in the APP Gene with Dominant-Negative Effect on Amyloidogenesis. *Science (New York, N.Y.)* 323(5920):1473-1477.

190. Jonsson T*, et al.* (2012) A mutation in APP protects against Alzheimer's disease and age-related cognitive decline. *Nature* 488(7409):96-99.

191. Betts V*, et al.* (2008) Aggregation and catabolism of disease-associated intra-Aβ mutations: reduced proteolysis of AβA21G by neprilysin. *Neurobiology of disease* 31(3):442-450.

192. Hoyer W, Cherny D, Subramaniam V, & Jovin TM (2004) Impact of the acidic C-terminal region comprising amino acids 109-140 on alpha-synuclein aggregation in vitro. *Biochemistry* 43(51):16233-16242.

193. Pike CJ, Overman MJ, & Cotman CW (1995) Amino-terminal deletions enhance aggregation of beta-amyloid peptides in vitro. *The Journal of biological chemistry* 270(41):23895-23898.

194. Ecroyd H & Carver JA (2009) Crystallin proteins and amyloid fibrils. *Cellular and molecular life sciences : CMLS* 66(1):62-81.

195. Hoyer W, Gronwall C, Jonsson A, Stahl S, & Hard T (2008) Stabilization of a beta-hairpin in monomeric Alzheimer's amyloid-beta peptide inhibits amyloid formation. *Proc Natl Acad Sci U S A* 105(13):5099-5104.

196. Kokkoni N, Stott K, Amijee H, Mason JM, & Doig AJ (2006) N-Methylated peptide inhibitors of beta-amyloid aggregation and toxicity. Optimization of the inhibitor structure. *Biochemistry* 45(32):9906-9918.

197. Shammas Sarah L*, et al.* (2011) Binding of the Molecular Chaperone αB-Crystallin to Aβ Amyloid Fibrils Inhibits Fibril Elongation. *Biophysical Journal* 101(7):1681-1689.

198. Grey M, Linse S, Nilsson H, Brundin P, & Sparr E (2011) Membrane interaction of alpha-synuclein in different aggregation states. *Journal of Parkinson's disease* 1(4):359-371.

199. Campioni S*, et al.* (2012) Salt anions promote the conversion of HypF-N into amyloid-like oligomers and modulate the structure of the oligomers and the monomeric precursor state. *J Mol Biol* 424(3-4):132-149.

200. Klement K*, et al.* (2007) Effect of different salt ions on the propensity of aggregation and on the structure of Alzheimer's abeta(1-40) amyloid fibrils. *J Mol Biol* 373(5):1321-1333.

201. Morel B, Varela L, Azuaga AI, & Conejero-Lara F (2010) Environmental Conditions Affect the Kinetics of Nucleation of Amyloid Fibrils and Determine Their Morphology. *Biophys. J.* 99(11):3801-3810.

202. Marek PJ, Patsalo V, Green DF, & Raleigh DP (2012) Ionic Strength Effects on Amyloid Formation by Amylin Are a Complicated Interplay between Debye Screening, Ion Selectivity, and Hofmeister Effects. *Biochemistry* 51(43):8478-8490.

203. Buell AK*, et al.* (2014) Solution conditions determine the relative importance of nucleation and growth processes in α-synuclein aggregation. *P Natl Acad Sci USA* 111(21):7671-7676.

204. Giehm L & Otzen DE (2010) Strategies to increase the reproducibility of protein fibrillization in plate reader assays. *Anal Biochem* 400(2):270-281.

205. Nielsen L*, et al.* (2001) Effect of environmental factors on the kinetics of insulin fibril formation: elucidation of the molecular mechanism. *Biochemistry* 40(20):6036-6046.

206. Hong LQ, X.; Zhang, Y. (2011) Dissecting the kinetic process of amyloid fiber formation through asymptotic analysis. *J. Phys. Chem. B* 116:6611-6617.

207. CJ R (2007) Non-native protein aggregation kinetics. *Biotechnology and bioengineering* 98:927-938.

208. Knowles TP, Vendruscolo M, & Dobson CM (2014) The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell Biol.* 15:384-396.

209. Xue W.F. HSW, and Radford S.E. (2008) Systematic analysis of nucleation-dependent polymerization reveals new insights into the mechanism of amyloid self-assembly. *Proc. Natl Acad. Sci. USA* 105:8926–8931.

210. Jarrett JT, Berger EP, & Lansbury PT (1993) The carboxy terminus of the [beta]-amyloid protein is critical for the seeding of amyloid formation: implications for the pathogenesis of Alzheimer's disease. *Biochemistry* 32:4693-4697.

211. Jarrett JT LP, Jr (1993) Seeding "one-dimensional crystallization" of amyloid: A pathogenic mechanism in Alzheimer's disease and scrapie? *Cell* 73(6).

212. Hofrichter J RP, Eaton WA (1974) Kinetics and mechanism of deoxyhemoglobin S gelation: A new approach to understanding sickle cell disease. *Proc Natl Acad Sci U S A* 71.

213. Auer S, Ricchiuto P, & Kashchiev D (2012) Two-step nucleation of amyloid fibrils: omnipresent or not? *J Mol Biol* 422(5):723-730.

214. Serio TR*, et al.* (2000) Nucleated conformational conversion and the replication of conformational information by a prion determinant. *Science* 289(5483):1317-1321.

215. Lee J CE, Powers ET, Kelly JW (2011) Amyloid-beta forms fibrils by nucleated conformational conversion of oligomers. *Nat Chem Biol* 7:602-609.

216. Hellstrand E BB, Walsh DM, Linse S (2010) Amyloid β-protein aggregation produces highly reproducible kinetic data and occurs by a two-phase process. *ACS Chem Neurosci* 1(1):13-18.

217. Anonymous (!!! INVALID CITATION !!!).

218. Bitan G*, et al.* (2003) Amyloid beta -protein (Abeta) assembly: Abeta 40 and Abeta 42 oligomerize through distinct pathways. *Proc Natl Acad Sci U S A* 100(1):330-335.

219. Shankar GM*, et al.* (2008) Amyloid-beta protein dimers isolated directly from Alzheimer's brains impair synaptic plasticity and memory. *Nature medicine* 14(8):837-842.

220. Liang Y, Lynn DG, & Berland KM (2010) Direct observation of nucleation and growth in amyloid self-assembly. *J Am Chem Soc* 132(18):6306-6308.

221. Bernstein SL*, et al.* (2005) Amyloid beta-protein: monomer structure and early aggregation states of Abeta42 and its Pro19 alloform. *J Am Chem Soc* 127(7):2075-2084.

222. Bernstein SL*, et al.* (2009) Amyloid-beta protein oligomerization and the importance of tetramers and dodecamers in the aetiology of Alzheimer's disease. *Nature chemistry* 1(4):326-331.

223. Bleiholder C, Dupuis NF, Wyttenbach T, & Bowers MT (2011) Ion mobility–mass spectrometry reveals a conformational conversion from random assembly to β-sheet in amyloid fibril formation. *Nature chemistry* 3(2):172-177.

224. Sabate R & Estelrich J (2005) Evidence of the existence of micelles in the fibrillogenesis of beta-amyloid peptide. *J Phys Chem B* 109(21):11027-11032.

225. Yong W*, et al.* (2002) Structure determination of micelle-like intermediates in amyloid β-protein fibril assembly by using small angle neutron scattering. *Proceedings of the National Academy of Sciences* 99(1):150-154.

226. Dobson CM (2004) Experimental investigation of protein folding and misfolding. *Methods* 34(1):4-14.

227. Banerjee PR & Deniz AA (2014) Shedding light on protein folding landscapes by single-molecule fluorescence. *Chem Soc Rev* 43(4):1172-1188.

228. Cornish PV & Ha T (2007) A survey of single-molecule techniques in chemical biology. *ACS Chem Biol* 2(1):53-61.

229. Hatzakis NS (2014) Single molecule insights on conformational selection and induced fit mechanism. *Biophys Chem* 186:46-54.

230. Joo C, Balci H, Ishitsuka Y, Buranachai C, & Ha T (2008) Advances in single-molecule fluorescence methods for molecular biology. *Annu Rev Biochem* 77(1):51-76.

231. JØRgensen SK & Hatzakis NS (2013) Insights in Enzyme Functional Dynamics and Activity Regulation by Single Molecule Studies. *Biophysical Reviews and Letters* 08(03n04):137-160.

232. Michalet X, Weiss S, & Jager M (2006) Single-molecule fluorescence studies of protein folding and conformational dynamics. *Chem Rev* 106(5):1785-1813.

233. Moerner WE (2007) New directions in single-molecule imaging and analysis. *Proc Natl Acad Sci U S A* 104(31):12596-12602.

234. Schuler B & Eaton WA (2008) Protein folding studied by single-molecule FRET. *Current opinion in structural biology* 18(1):16-26.

235. Xie XS, Choi PJ, Li GW, Lee NK, & Lia G (2008) Single-molecule approach to molecular biology in living bacterial cells. *Annu Rev Biophys* 37(1):417-444.

236. Zhuang X (2005) Single-molecule RNA science. *Annu Rev Biophys Biomol Struct* 34:399-414.

237. Walter NG, Huang C-Y, Manzo AJ, & Sobhy MA (2008) Do-it-yourself guide: how to use the modern single-molecule toolkit. *Nat Meth* 5(6):475-489.

238. Deniz AA, Mukhopadhyay S, & Lemke EA (2008) Single-molecule biophysics: at the interface of biology, physics and chemistry. *J R Soc Interface* 5(18):15-45.

239. Claessen VI*, et al.* (2010) Single-biomolecule kinetics: the art of studying a single enzyme. *Annu Rev Anal Chem (Palo Alto Calif)* 3(1):319-340.

240. Hohng S, Lee S, Lee J, & Jo MH (2014) Maximizing information content of single-molecule FRET experiments: multi-color FRET and FRET combined with force or torque. *Chem Soc Rev* 43(4):1007-1013.

241. Neuman KC & Nagy A (2008) Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nat Methods* 5(6):491-505.

242. Greenleaf WJ, Woodside MT, & Block SM (2007) High-resolution, single-molecule measurements of biomolecular motion. *Annu Rev Biophys Biomol Struct* 36(1):171-190.

243. Oddershede LB (2012) Force probing of individual molecules inside the living cell is now a reality. *Nat. Chem. Biol.* 8:879-886.

244. Fazal FM & Block SM (2011) Optical tweezers study life under tension. *Nature photonics* 5:318-321.

245. Engel A & Muller DJ (2000) Observing single biomolecules at work with the atomic force microscope. *Nat Struct Mol Biol* 7(9):715-718.

246. Fisher TE, Marszalek PE, & Fernandez JM (2000) Stretching single molecules into novel conformations using the atomic force microscope. *Nat Struct Mol Biol* 7(9):719-724.

247. Kedrov A, Janovjak H, Sapra KT, & Muller DJ (2007) Deciphering molecular interactions of native membrane proteins by single-molecule force spectroscopy. *Annu Rev Biophys Biomol Struct* 36:233-260.

248. Muller DJD, Y.F. (2008) Atomic force microscopy as a multifunctional molecular toolbox in nanobiotechnology. *Nat. Nanotech.* 3:261-269.

249. Sheiko SS & Möller M (2001) Visualization of MacromoleculesA First Step to Manipulation and Controlled Response. *Chemical Reviews* 101(12):4099-4124.

250. Veigel C & Schmidt CF (2011) Moving into the cell: single-molecule studies of molecular motors in complex environments. *Nat Rev Mol Cell Biol* 12(3):163-176.

251. Pinaud FC, S.; Sittner, A.; Dahan, M. (2010) Probing cellular events, one quantum dot at a time. *Nat. Methods* 7:275-285.

252. Yang H (2010) Progress in single-molecule spectroscopy in cells. *Curr Opin Chem Biol* 14(1):3-9.

253. Ha T, Kozlov AG, & Lohman TM (2012) Single-molecule views of protein movement on single-stranded DNA. *Annu Rev Biophys* 41:295-319.

254. Smiley RD & Hammes GG (2006) Single molecule studies of enzyme mechanisms. *Chem Rev* 106(8):3080-3094.

255. Min W*, et al.* (2005) Fluctuating enzymes: lessons from single-molecule studies. *Acc Chem Res* 38(12):923-931.

256. Tinnefeld P & Sauer M (2005) Branching out of single-molecule fluorescence spectroscopy: challenges for chemistry and influence on biology. *Angew. Chem. Int. Edn. Engl.* 44:2642-2671.

257. Roeffaers MB*, et al.* (2007) Single-molecule fluorescence spectroscopy in (bio)catalysis. *Proc Natl Acad Sci U S A* 104(31):12603-12609.

258. Kapanidis AN & Strick T (2009) Biology, one molecule at a time. *Trends Biochem. Sci.* 34(5):234-243.

259. Bhattacharya M & Mukhopadhyay S (2016) Studying Protein Misfolding and Aggregation by Fluorescence Spectroscopy. *Reviews in Fluorescence 2015*, ed Geddes DC (Springer International Publishing, Cham), pp 1-27.

260. Schmolze DB, Standley C, Fogarty KE, & Fischer AH (2011) Advances in microscopy techniques. *Archives of pathology & laboratory medicine* 135(2):255-263.

261. Weisenburger S & Sandoghdar V (2015) Light microscopy: an ongoing contemporary revolution. *Contemporary Physics* 56(2):123-143.

262. Gust A*, et al.* (2014) A starting point for fluorescence-based single-molecule measurements in biomolecular research. *Molecules* 19(10):15824-15865.

263. Magde D, Webb WW, & E E (1972) Thermodynamic fluctuations in a reacting systemmeasurement by fluorescence correlation spectroscopy. *Phys Rev Lett* 29:705–708.

264. Gerard M*, et al.* (2006) The aggregation of alpha-synuclein is stimulated by FK506 binding proteins as shown by fluorescence correlation spectroscopy. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 20(3):524-526.

265. Nath S, Meuvis J, Hendrix J, Carl SA, & Engelborghs Y (2010) Early Aggregation Steps in a-Synuclein as Measured by FCS and FRET: Evidence for a Contagious Conformational Change. *Biophys. J.* 98(7):1302-1311.

266. Kapusta P, Wahl M, Benda A, Hof M, & Enderlein J (2007) Fluorescence lifetime correlation spectroscopy. *Journal of fluorescence* 17(1):43-48.

267. Pitschke M, Prior R, Haupt M, & Riesner D (1998) Detection of single amyloid beta-protein aggregates in the cerebrospinal fluid of Alzheimer's patients by fluorescence correlation spectroscopy. *Nature medicine* 4(7):832-834.

268. Bieschke J, *et al.* (2000) Ultrasensitive detection of pathological prion protein aggregates by dual-color scanning for intensely fluorescent targets. *Proc Natl Acad Sci U S A* 97(10):5468-5473.

269. Birkmann E, *et al.* (2006) Detection of prion particles in samples of BSE and scrapie by fluorescence correlation spectroscopy without proteinase K digestion. *Biological chemistry* 387(1):95-102.

270. Funke SA, *et al.* (2007) Single particle detection of Abeta aggregates associated with Alzheimer's disease. *Biochemical and biophysical research communications* 364(4):902-907.

271. Lakowicz JR (2006) Principles of Fluorescence Spectroscopy. (Springer US).

272. Roy R, Hohng S, & Ha T (2008) A practical guide to single-molecule FRET. *Nat Methods* 5(6):507-516.

273. Förster T (1948) Zwischenmolekulare Energiewanderung und Fluoreszenz. *Annalen der Physik* 437(1-2):55-75.

274. Ha T, *et al.* (1996) Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. *Proc Natl Acad Sci U S A* 93(13):6264-6268.

275. Deniz AA, *et al.* (1999) Single-pair fluorescence resonance energy transfer on freely diffusing molecules: observation of Forster distance dependence and subpopulations. *Proc Natl Acad Sci U S A* 96(7):3670-3675.

276. Jia Y, *et al.* (1999) Folding dynamics of single GCN-4 peptides by fluorescence resonant energy transfer confocal microscopy. *Chemical Physics* 247(1):69-83.

277. Talaga DS, *et al.* (2000) Dynamics and folding of single two-stranded coiled-coil peptides studied by fluorescent energy transfer confocal microscopy. *Proc Natl Acad Sci U S A* 97(24):13021-13026.

278. Deniz AA, *et al.* (2000) Single-molecule protein folding: diffusion fluorescence resonance energy transfer studies of the denaturation of chymotrypsin inhibitor 2. *Proc Natl Acad Sci U S A* 97(10):5179-5184.

279. Hoffmann A, Neupane K, & Woodside MT (2013) Single-molecule assays for investigating protein misfolding and aggregation. *Phys Chem Chem Phys* 15(21):7934-7948.

280. Itzhaki LS, Otzen DE, & Fersht AR (1995) The structure of the transition state for folding of chymotrypsin inhibitor 2 analysed by protein engineering methods: evidence for a nucleation-condensation mechanism for protein folding. *J Mol Biol* 254(2):260-288.

281. Schuler B, Lipman EA, & Eaton WA (2002) Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy. *Nature* 419(6908):743-747.

282. Rhoades E, Gussakovsky E, & Haran G (2003) Watching proteins fold one molecule at a time. *Proc Natl Acad Sci U S A* 100(6):3197-3202.

283. Rhoades E, Cohen M, Schuler B, & Haran G (2004) Two-state folding observed in individual protein molecules. *J Am Chem Soc* 126(45):14686-14687.

284. Kuzmenkina EV, Heyes CD, & Nienhaus GU (2005) Single-molecule Forster resonance energy transfer study of protein dynamics under denaturing conditions. *Proc Natl Acad Sci U S A* 102(43):15471-15476.

285. Ruedas-Rama MJ, Alvarez-Pez JM, & Orte A (2013) Solving Single Biomolecules by Advanced Fret-Based Single-Molecule Fluorescence Techniques. *Biophysical Reviews and Letters* 08(03n04):161-190.

286. Li H, Ying L, Green JJ, Balasubramanian S, & Klenerman D (2003) Ultrasensitive coincidence fluorescence detection of single DNA molecules. *Analytical chemistry* 75(7):1664-1670.

287. Kapanidis AN*, et al.* (2005) Alternating-laser excitation of single molecules. *Acc Chem Res* 38(7):523-533.

288. Kapanidis AN*, et al.* (2004) Fluorescence-aided molecule sorting: analysis of structure and interactions by alternating-laser excitation of single molecules. *Proc Natl Acad Sci U S A* 101(24):8936-8941.

289. Edel JB, Eid JS, & Meller A (2007) Accurate single molecule FRET efficiency determination for surface immobilized DNA using maximum likelihood calculated lifetimes. *J Phys Chem B* 111(11):2986-2990.

290. Kudryavtsev V*, et al.* (2012) Combining MFD and PIE for accurate single-pair Forster resonance energy transfer measurements. *Chemphyschem* 13(4):1060-1078.

291. Widengren J*, et al.* (2006) Single-molecule detection and identification of multiple species by multiparameter fluorescence detection. *Analytical chemistry* 78(6):2039-2050.

292. Muller BK, Zaychikov E, Brauchle C, & Lamb DC (2005) Pulsed interleaved excitation. *Biophys J* 89(5):3508-3522.

293. Hillger F, Nettels D, Dorsch S, & Schuler B (2007) Detection and analysis of protein aggregation with confocal single molecule fluorescence spectroscopy. *Journal of fluorescence* 17(6):759-765.

294. Grohmann DW, F.; Tinnefeld, P. (2013) Making connections—Strategies for single molecule fluorescence biophysics. *Curr. Opin. Chem. Biol.* 17:691–698.

295. RP. F (1992) There's plenty of room at the bottom. *J Microelectromech Syst.* 1:60-66.

296. Bustamante C, Bryant Z, & Smith SB (2003) Ten years of tension: single-molecule DNA mechanics. *Nature* 421(6921):423-427.

297. Kapanidis AN & Weiss S (2002) Fluorescent probes and bioconjugation chemistries for single-molecule fluorescence analysis of biomolecules. *J. Chem. Phys.* 117:10953-10964.

298. Higuchi R, Krummel B, & Saiki RK (1988) A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. *Nucleic Acids Res* 16(15):7351-7367.

299. Braman J (2001) *In Vitro Mutagenesis Protocols. 2nd.* (Totowa, NJ).

300. MW P (1994) Site-specific chemical modification procedures. *Methods Mol Biol*), Vol 35, pp 171–185.

301. Pennington MW (1994) Site-specific chemical modification procedures. *Methods Mol Biol* 35:171-185.

302. Duzdevich D & Greene EC (2013) Towards physiological complexity with in vitro single-molecule biophysics. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 368(1611):20120271.

303. Nettels D, Hoffmann A, & Schuler B (2008) Unfolded protein and peptide dynamics investigated with single-molecule FRET and correlation spectroscopy from picoseconds to seconds. *J Phys Chem B* 112(19):6137-6146.

304. Ha T, Laurence TA, Chemla DS, & Weiss S (1999) Polarization Spectroscopy of Single Fluorescent Molecules. *The Journal of Physical Chemistry B* 103(33):6839-6850.

305. Giese A*, et al.* (2005) Single particle detection and characterization of synuclein co-aggregation. *Biochemical and biophysical research communications* 333(4):1202-1210.

306. Levin J, Bertsch U, Kretzschmar H, & Giese A (2005) Single particle analysis of manganese-induced prion protein aggregates. *Biochemical and biophysical research communications* 329(4):1200-1207.

307. Puchalla J, Krantz K, Austin R, & Rye H (2008) Burst analysis spectroscopy: a versatile single-particle approach for studying distributions of protein aggregates and fluorescent assemblies. *Proc Natl Acad Sci U S A* 105(38):14400-14405.

308. Kostka M*, et al.* (2008) Single particle characterization of iron-induced pore-forming alpha-synuclein oligomers. *The Journal of biological chemistry* 283(16):10992-11003.

309. Hillmer AS*, et al.* (2010) Converse modulation of toxic alpha-synuclein oligomers in living cells by N'-benzylidene-benzohydrazide derivates and ferric iron. *Biochemical and biophysical research communications* 391(1):461-466.

310. Hogen T*, et al.* (2012) Two different binding modes of alpha-synuclein to lipid vesicles depending on its aggregation state. *Biophys J* 102(7):1646-1655.

311. Bader B*, et al.* (2011) Single particle analysis of tau oligomer formation induced by metal ions and organic solvents. *Biochemical and biophysical research communications* 411(1):190-196.

312. Duering M*, et al.* (2011) Co-aggregate formation of CADASIL-mutant NOTCH3: a single-particle analysis. *Human molecular genetics* 20(16):3256-3265.

313. Orte A, Clarke R, Balasubramanian S, & Klenerman D (2006) Determination of the fraction and stoichiometry of femtomolar levels of biomolecular complexes in an excess of monomer using single-molecule, two-color coincidence detection. *Analytical chemistry* 78(22):7707-7715.

314. Chiou A*, et al.* (2009) Probing neuroserpin polymerization and interaction with amyloid-beta peptides using single molecule fluorescence. *Biophys J* 97(8):2306-2315.

315. Narayan P, *et al.* (2012) The extracellular chaperone clusterin sequesters oligomeric forms of the amyloid-β1−40 peptide. *Nature Struct. Mol. Biol.* 19:79-83.

316. Cremades N, *et al.* (2012) Direct Observation of the Interconversion of Normal and Toxic Forms of a-Synuclein. *Cell* 149(5):1048-1059.

317. Horrocks MH, *et al.* (2012) Single Molecule Fluorescence under Conditions of Fast Flow. *Anal. Chem.* 84(1):179-185.

318. Horrocks MH, *et al.* (2015) Fast flow microfluidics and single-molecule fluorescence for the rapid characterization of alpha-synuclein oligomers. *Analytical chemistry* 87(17):8818-8826.

319. Shammas SL, *et al.* (2015) A mechanistic model of tau amyloid aggregation based on direct observation of oligomers. *Nature communications* 6:7025.

320. Tosatto L, *et al.* (2015) Single-molecule FRET studies on alpha-synuclein oligomerization of Parkinson's disease genetically related mutants. *Sci Rep* 5:16696.

321. Iljina M, *et al.* (2016) Kinetic model of the aggregation of alpha-synuclein provides insights into prion-like spreading. *Proc Natl Acad Sci U S A* 113(9):E1206-1215.

322. Vendruscolo M, Knowles TPJ, & Dobson CM (2011) Protein Solubility and Protein Homeostasis: A Generic View of Protein Misfolding Disorders. *Cold Spring Harbor Perspectives in Biology* 3(12):a010454.

323. Eichner T & Radford Sheena E (A Diversity of Assembly Mechanisms of a Generic Amyloid Fold. *Molecular cell* 43(1):8-18.

324. Stefani M (2007) Generic cell dysfunction in neurodegenerative disorders: role of surfaces in early protein misfolding, aggregation, and aggregate cytotoxicity. *The Neuroscientist : a review journal bringing neurobiology, neurology and psychiatry* 13(5):519-531.

325. Guijarro JI, Sunde M, Jones JA, Campbell ID, & Dobson CM (1998) Amyloid fibril formation by an SH3 domain. *Proc Natl Acad Sci U S A* 95(8):4224-4228.

326. Jimenez JL, *et al.* (1999) Cryo-electron microscopy structure of an SH3 amyloid fibril and model of the molecular packing. *The EMBO journal* 18(4):815-821.

327. Zurdo J, Guijarro JI, Jimenez JL, Saibil HR, & Dobson CM (2001) Dependence on solution conditions of aggregation and amyloid formation by an SH3 domain. *J Mol Biol* 311(2):325-340.

328. Morel B, Ruzafa D, & Conejero-Lara F (2015) SH3 Domains as Suitable Models to Study Amyloid Aggregation. *SH Domains: Structure, Mechanisms and Applications*, ed Kurochkina N (Springer International Publishing, Cham), pp 1-15.

329. Kuriyan J & Cowburn D (1993) Structures of SH2 and SH3 domains. *Current opinion in structural biology* 3:828-837.

330. Mayer BJ (2001) SH3 domains: complexity in moderation. *J Cell Sci* 114:1253-1263.

331. Kay BK, Williamson MP, & Sudol M (2000) The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 14:231-241.

332. Zhuang X & Jiang L (2015) SH Domain Proteins in Plants: Roles in Signaling Transduction and Membrane Trafficking. *SH Domains: Structure, Mechanisms and Applications*, ed Kurochkina N (Springer International Publishing, Cham), pp 17-33.

333. Musacchio A, Noble M, Pauptit R, Wierenga R, & Saraste M (1992) Crystal structure of a Src-homology 3 (SH3) domain. *Nature* 359(6398):851-855.

334. Varela L, Morel B, Azuaga AI, & Conejero-Lara F (2009) A single mutation in an SH3 domain increases amyloid aggregation by accelerating nucleation, but not by destabilizing thermodynamically the native state. *FEBS Lett.* 583(4):801-806.

335. Ruzafa D, Varela L, Azuaga AI, Conejero-Lara F, & Morel B (2014) Mapping the structure of amyloid nucleation precursors by protein engineering kinetic analysis. *Phys Chem Chem Phys* 16(7):2989-3000.

336. Ruzafa D, Conejero-Lara F, & Morel B (2013) Modulation of the stability of amyloidogenic precursors by anion binding strongly influences the rate of amyloid nucleation. *Phys Chem Chem Phys* 15(37):15508-15517.

337. Ruzafa D, Morel B, Varela L, Azuaga AI, & Conejero-Lara F (2014) Correction: Characterization of Oligomers of Heterogeneous Size as Precursors of Amyloid Fibril Nucleation of an SH3 Domain: An Experimental Kinetics Study. *PLoS*

ONE 9(1):10.1371/annotation/dbb84118-84119ada-84143e84114-88734-84118f88322be84111a84159.

338. Morel B, Casares S, & Conejero-Lara F (2006) A single mutation induces amyloid aggregation in the alpha-spectrin SH3 domain: analysis of the early stages of fibril formation. *J Mol Biol* 356(2):453-468.

339. Ruzafa D, Morel B, Varela L, Azuaga AI, & Conejero-Lara F (2012) Characterization of oligomers of heterogeneous size as precursors of amyloid fibril nucleation of an SH3 domain: an experimental kinetics study. *PLoS One* 7(11):e49690.

340. Paredes JM*, et al.* (2012) Early Amyloidogenic Oligomerization Studied through Fluorescence Lifetime Correlation Spectroscopy. *Int. J. Mol. Sci.* 13:9400-9418.

341. Rüttinger S, Kapusta P, Patting M, Wahl M, & Macdonald R (2010) On the Resolution Capabilities and Limits of Fluorescence Lifetime Correlation Spectroscopy (FLCS) Measurements. *J. Fluoresc.* 20(1):105-114.

342. Chiti F & Dobson CM (2006) Protein Misfolding, Functional Amyloid, and Human Disease. *Annu Rev Biochem* 75(1):333-366.

343. Orte A, Clarke R, Balasubramanian S, & Klenerman D (2006) Determination of the Fraction and Stoichiometry of Femtomolar Levels of Biomolecular Complexes in an Excess of Monomer Using Single-Molecule, Two-Color Coincidence Detection. *Anal. Chem.* 78(22):7707-7715.

344. Chiou A*, et al.* (2009) Probing Neuroserpin Polymerization and Interaction with Amyloid-β Peptides Using Single Molecule Fluorescence. *Biophysical Journal* 97(8):2306-2315.

345. Narayan P*, et al.* (2012) The extracellular chaperone clusterin sequesters oligomeric forms of the amyloid-β1−40 peptide. *Nat. Struct. Mol. Biol.* 19(1):79-83.

346. Orte A*, et al.* (2008) Direct characterization of amyloidogenic oligomers by single-molecule fluorescence. *Proc. Natl. Acad. Sci. U.S.A.* 105(38):14424-14429.

347. Cremades N*, et al.* (2011) Direct observation of the interconversion of normal and pathogenic forms of alpha-synuclein. *Eur. Biophys. J. Biophys. Lett.* 40:215-216.

348. Casares S, Sadqi M, Lopez-Mayorga O, Conejero-Lara F, & van Nuland NA (2004) Detection and characterization of partially unfolded oligomers of the SH3 domain of alpha-spectrin. *Biophys J* 86(4):2403-2413.

349. Morel B, Casares S, & Conejero-Lara F (2006) A single mutation induces amyloid aggregation in the alpha-spectrin SH3 domain: Analysis of the early stages of fibril formation. *J. Mol. Biol.* 356(2):453-468.

350. Paredes JM, *et al.* (2012) Early Amyloidogenic Oligomerization Studied through Fluorescence Lifetime Correlation Spectroscopy. *Int. J. Mol. Sci.* 13(8):9400-9418.

351. ATTO-Tec (Förster-Radius of ATTO-Dye-Pairs for FRET https://www.atto-tec.com/index.php?id=65&L=1 (accessed Oct. 2014).

352. Eggeling C. ea (2001) Data registration and selective single-molecule analysis using multi-parameter fluorescence detection. *Journal of Biotechnology* 86(3):163-180.

353. M. Wahl RE, K. Lauritsen, H. J. Rahn (1998) Hardware solution for continuous time-resolved burst detection of single molecules in flow. *Proc. SPIE* 3259:173-178.

354. Hoffmann A, *et al.* (2007) Mapping protein collapse with single-molecule fluorescence and kinetic synchrotron radiation circular dichroism spectroscopy. *Proc Natl Acad Sci U S A* 104(1):105-110.

355. Edel JB, Eid JS, & Meller A (2007) Accurate Single Molecule FRET Efficiency Determination for Surface Immobilized DNA Using Maximum Likelihood Calculated Lifetimes. *J. Phys. Chem. B* 111(11):2986-2990.

356. Widengren J, *et al.* (2006) Single-Molecule Detection and Identification of Multiple Species by Multiparameter Fluorescence Detection. *Anal. Chem.* 78(6):2039-2050.

357. Schuler B, Lipman EA, Steinbach PJ, Kumke M, & Eaton WA (2005) Polyproline and the "spectroscopic ruler" revisited with single-molecule fluorescence. *Proc Natl Acad Sci U S A* 102(8):2754-2759.

358. Orte A, *et al.* (2008) Direct characterization of amyloidogenic oligomers by single-molecule fluorescence. *Proc. Natl. Acad. Sci. U.S.A.* 105:14424-14429.

359. Lakowicz JRe & (2006) *Principles of Fluorescence Spectroscopy* (Springer) 3rd Ed.

360. Narayan P*, et al.* (2012) The extracellular chaperone clusterin sequesters oligomeric forms of the amyloid-beta(1-40) peptide. *Nat Struct Mol Biol* 19(1):79-83.

361. Dempsey GT*, et al.* (2009) Photoswitching mechanism of cyanine dyes. *J Am Chem Soc* 131(51):18192-18193.

362. Kapanidis AN*, et al.* (2008) Alternating-Laser Excitation of Single Molecules. *Single Molecule Techniques. A Laboratory Manual*, eds Selvin PR & Ha T (Cold Spring Harbor Laboratory Press, New York).

363. Schindelin J*, et al.* (2012) Fiji: an open-source platform for biological-image analysis. *Nat Methods* 9(7):676-682.

364. Mark HF (1984) Polymer chemistry—the basic concepts, by Paul C. Hiemenz, Marcel Dekker, New York, 1984, 738 pp. No price given. *Journal of Polymer Science: Polymer Letters Edition* 22(9):509-510.

365. Johnson WC, Jr. (1990) Protein secondary structure and circular dichroism: a pratical guide. *Proteins* 7:205-214.

366. G.D. GNF (1969) computed circular dichroism spectra for the evaluation of protein conformation. *Biochemistry* 8:4108-4116.

367. N. KSaP (2000) The use of circular dichroism in the investigation of protein structure and function. *Current protein and peptide science* 1:349-384.

368. Sgarbossa A, Buselli D, & Lenci F (2008) In vitro perturbation of aggregation processes in β-amyloid peptides: A spectroscopic study. *FEBS letters* 582(23–24):3288-3292.

369. Wallace LWaBA (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. in *Nucleic Acids Research*.

370. Sreerama N & Woody RW (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(2):252-260.

371. Bohm G, Muhr R, & Jaenicke R (1992) Quantitative analysis of protein far UV circular dichroism spectra by neural networks. *Protein engineering* 5(3):191-195.

372. Ventura S & Serrano L (2003) Spectroscopic Analysis of Amyloid Fibril Formation in SH3-Domains. *Spectroscopy* 17(4).

373. Plotnikov VV, Brandts JM, Lin LN, & Brandts JF (1997) A new ultrasensitive scanning calorimeter. *Anal Biochem* 250(2):237-244.

374. Chiu MH & Prenner EJ (2011) Differential scanning calorimetry: An invaluable tool for a detailed thermodynamic characterization of macromolecules and their interactions. *J Pharm Bioallied Sci* 3(1):39-59.

375. Makhatadze GI & Privalov PL (1990) Heat capacity of proteins. I. Partial molar heat capacity of individual amino acid residues in aqueous solution: hydration effect. *J Mol Biol* 213(2):375-384.

376. privalov P.L. and potekhin SA (1986) Scanning microcalorimetry in studying temperature-induced changes in proteins. *Methods in enzymology* 131:4-51.

377. Johnson CM (2013) Differential scanning calorimetry as a tool for protein folding and stability. *Archives of Biochemistry and Biophysics* 531:100–109.

378. Privalov PL & Potekhin SA (1986) Scanning microcalorimetry in studying temperature-induced changes in proteins. *Methods in enzymology* 131:4-51.

379. Viguera ARM, J.C.; Filimonov, V.V.; Mateo, P.L. and Serrano, L. (1994) Thermodynamic and kynetic analysis of the SH3 domain of spectrin shows a two-state folding transition. *Biochemistry* 33(8):2142-2150.

380. Kapoor M (1996) How to cross-link proteins. *Cellular, Molecular and Microbial Biology Division, University of Calgary, Calgary, Canada, T2N 1N4*.

381. Migneault I, Dartiguenave C, Bertrand MJ, & Waldron KC (2004) Glutaraldehyde: behavior in aqueous solution, reaction with proteins, and application to enzyme crosslinking. *BioTechniques* 37(5):790-796, 798-802.

382. Fadouloglou VE, Kokkinidis M, & Glykos NM (2008) Determination of protein oligomerization state: two approaches based on glutaraldehyde crosslinking. *Anal Biochem* 373(2):404-406.

383. Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227(5259):680-685.

384. Böhm G, Muhr R, & Jaenicke R (1992) Quantitative analysis of protein far UV circular dichroism spectra by neural networks. *Protein Eng.* 5:191-195.

385. Doran TM, Kamens AJ, Byrnes NK, & Nilsson BL (2012) Role of amino acid hydrophobicity, aromaticity, and molecular volume on IAPP(20-29) amyloid self-assembly. *Proteins* 80(4):1053-1065.

386. Castello F, Casares S, Ruedas-Rama MJ, & Orte A (2015) The First Step of Amyloidogenic Aggregation. *J Phys Chem B* 119(26):8260-8267.

387. Narayan P*, et al.* (2012) The extracellular chaperone clusterin sequesters oligomeric forms of the amyloid-beta(1-40) peptide. *Nat Struct Mol Biol* 19:79-83.

388. Cremades N*, et al.* (2012) Direct Observation of the Interconversion of Normal and Toxic Forms of a-Synuclein. *Cell* 149:1048-1059.

389. Zampagni M*, et al.* (2011) A comparison of the biochemical modifications caused by toxic and non-toxic protein oligomers in cells. *J Cell Mol Med* 15(10):2106-2116.

390. Vega MC, Martinez JC, & Serrano L (2000) Thermodynamic and structural characterization of Asn and Ala residues in the disallowed II' region of the Ramachandran plot. *Protein Sci* 9(12):2322-2328.

391. Orte A, Clarke RW, & Klenerman D (2008) Fluorescence Coincidence Spectroscopy for Single-Molecule Fluorescence Resonance Energy-Transfer Measurements. *Anal. Chem.* 80(22):8389-8397.

392. Vega MC, Martinez JC, & Serrano L (2000) Thermodynamic and structural characterization of Asn and Ala residues in the disallowed II' region of the Ramachandran plot. *Protein Sci.* 9(12):2322-2328.

393. Wilkins DKG, S. B.; Receveur, V.; Dobson, C. M.; Jones, J. A.; Smith, L. J. (1999) Hydrodynamic Radii of Native and Denatured Proteins Measured by Pulse Field Gradient NMR Techniques *Biochemistry* (38):16424-16431.

394. Wilkins DK*, et al.* (1999) Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques. *Biochemistry* 38(50):16424-16431.

395. Casares S, Sadqi M, Lopez-Mayorga O, Martinez JC, & Conejero-Lara F (2003) Structural cooperativity in the SH3 domain studied by site-directed mutagenesis and amide hydrogen exchange. *FEBS letters* 539(1-3):125-130.

396. David Ruzafa BM, Lorena Varela, Ana I. Azuaga, Francisco Conejero-Lara (2012) Characterization of Oligomers of Heterogeneous Size as Precursors of Amyloid Fibril Nucleation of an SH3 Domain: An Experimental Kinetics Study. *PLOS ONE* 7(11).

397. Cohen SI*, et al.* (2015) A molecular chaperone breaks the catalytic cycle that generates toxic Abeta oligomers. *Nat Struct Mol Biol* 22(3):207-213.

398. Marshall KE*, et al.* (2011) Hydrophobic, aromatic, and electrostatic interactions play a central role in amyloid fibril formation and stability. *Biochemistry* 50(12):2061-2071.

399. Cohen SI, Vendruscolo M, Dobson CM, & Knowles TP (2011) Nucleated polymerization with secondary pathways. II. Determination of self-consistent solutions to growth processes described by non-linear master equations. *The Journal of chemical physics* 135(6):065106.

400. Bertrand Morel DRaFC-L (2015) *SH Domains Structure, Mechanisms and Applications*.

401. Campioni S*, et al.* (2010) A causative link between the structure of aberrant protein oligomers and their toxicity. *Nat Chem Biol* 6(2):140-147.