Mechanism of Disorder-to-Order Amyloid Transition: The Role of Conformational Plasticity and Water Mobility of Disordered Proteins

SHRUTI ARYA

Thesis submitted for the degree of

DOCTOR OF PHILOSOPHY

Department of Chemical Sciences

Indian Institute of Science Education and Research (IISER) Mohali

October 2016

CERTIFICATE

The work presented in this thesis has been carried out by me under the supervision of Dr. Samrat Mukhopadhyay at the Department of Chemical Sciences, Indian Institute of Science Education and Research (IISER) Mohali.

This work has not been submitted in part or full for a degree, a diploma, or a fellowship to any other university or institute.

Whenever contributions of others are involved, every effort is made to indicate this clearly, with due acknowledgement of collaborative research and discussions. This thesis is a bonafide record of original work done by me and all sources listed within have been detailed in the bibliography.

Date **Place** **Shruti Arya**

In my capacity as supervisor of the candidate's thesis work, I certify that above statements made by the candidate are true to the best of my knowledge.

Dr. Samrat Mukhopadhyay

(Supervisor)

****

Dedicated to

My family

****

Acknowledgements

This thesis would not have been possible without the help and support of a large number of individuals who in some way or the other have made a huge difference during my PhD journey. I take this opportunity to thank all of them for being there for me.

First of all, I would like to express my sincere gratitude to my PhD advisor Dr. Samrat Mukhopadhyay for his constant encouragement and support throughout my PhD. He has always been a great source of inspiration and has given me tremendous freedom to pursue my research, which aided in my growth as an independent scientist. He always gave me his valuable suggestions so that I could improve my research and writing skills. His passion for science is extremely contagious and has always motivated me to pursue my research with immense joy and enthusiasm. Besides the extraordinary scientific temperament, he also has all the other ingredients of a great mentor and I consider myself extremely fortunate to have worked with him for my graduate studies. I really appreciate all his support towards making my PhD experience productive and stimulating, and thank him for believing in me.

I gratefully acknowledge Dr. Mily Bhattacharya for her valuable advices, supervision and crucial contribution to this thesis. I want to thank her for the generous time and promptness she has always offered in giving me feedback on my research. She has always been a source of encouragement during difficult times.

I am thankful to Professor N. Sathyamurthy, the Director of IISER Mohali for providing me the excellent research facilities.

I would like to thank the members of my doctoral committee, Professor K. S. Viswanathan and Professor Purnananda Guptasarma, for their encouragement and valuable suggestions on my work.

I thank all the current and former members of the Mukhopadhyay lab. This includes Anupa, Sourav, Hema, Karishma, Priyanka, Priyanka Madhu, Aishwarya, Debapriya, Sayanta, Mily, Pushpender, Arpana, Neha, Dominic, Vijit, Anubhuti and Vinodh. All of them have contributed immensely to my personal and professional

time at IISER. The group has been a source of friendships as well as great advices and collaborations. I owe special thanks to Mily, Arpana, Neha, Karishma and Priyanka for their contributions into my thesis work.

I would like to thank Professor Anindya Datta at the Indian Institute of Technology (IIT), Bombay and the members of his lab, Avinash and Tuhin for their help in the femtosecond experiments.

I gratefully acknowledge Dr. Payel Das at IBM Thomas J. Watson Research Center, Yorktown Heights, NY USA for collaborating with us on α-synuclein project.

I would like to thank Professor Vinod Subramaniam (University of Twente, Netherlands) for providing with α-synuclein plasmid.

I gratefully acknowledge the help and support from all the faculty members, staff, library, and instrumentation facility at IISER.

I would like to thank CSIR and CPSDE for the financial support. The financial support from DST, IISER and the Biophysical Society for attending the international conference is also duly acknowledged.

I thank all my friends and cousins for their immense care and support. I particularly thank Chinmoy, Satyam, Sumyra, Shilpa, Maheshwararao, Moumita, Gaurav, Abhishek, Vinca, Anindya, Shilpi, Gagandeep, Mansi, Siddharth, Smriti, Aamya, Divyadeep, Sanchi, Vinni, Ritika, Viddur, Vaishali, Himanshu, Ruchi, Mukesh and Sandhya. I owe a special thanks to Chinmoy for his constant support, care and brainstorming discussions.

Finally, I would like to express my deepest gratitude to my grandmother and my parents for their immeasurable love, care and encouragement throughout my life. They always stood by me like a pillar of strength.

Shruti Arya

Synopsis

Mechanism of Disorder-to-Order Amyloid Transition: The Role of Conformational Plasticity and Water Mobility of Disordered Proteins

Chapter 1. Introduction

According to the traditional sequence-structure-function paradigm, a protein must adopt a well-defined 3-dimensional structure in order to perform its function. However, an emerging class of proteins, known as intrinsically disordered proteins (IDPs), confronts the traditional structure-function paradigm. It is interesting to note that nearly 33% of genome codes for IDPs or IDRs (intrinsically disordered regions). Under certain unfortunate circumstances inside the cell, both folded proteins and IDPs can go to an alternate thermodynamically stable state known as amyloid state. These amyloids are implicated in a large number of debilitating neurodegenerative diseases, such as Parkinson's and Alzheimer's diseases. The intricate balance between chain-chain and chain-solvent (water) interactions underlies the major physical driving force for protein aggregation leading to amyloid formation. During the amyloid transition, various intermediates have been observed and there is an emerging consensus that these intermediate species might perhaps be much more cytotoxic than the amyloid fibrils. Thus, the morphologies of these species might play an important role in the pathogenesis of amyloid diseases. An enlarging body of data have provided significant insights into the mechanism of coupled conformational switch and aggregation from the protein perspective. However, the precise mechanism of amyloid formation, the behavior of water molecules in and around IDPs, and the role of water in amyloid formation remains elusive. In this thesis, efforts were directed towards understanding the role of both chain-chain and chain-solvent interactions in protein aggregation, using a variety of biophysical tools. The conformational changes (chain-chain interactions) during the amyloid formation were mainly studied using circular dichroism (CD), Raman spectroscopy and atomic force microscopy (AFM), and the nature of water molecules in IDPs and amyloids (chain-solvent)

was monitored using femtosecond- and picosecond-resolved fluorescence spectroscopy (Figure 1).

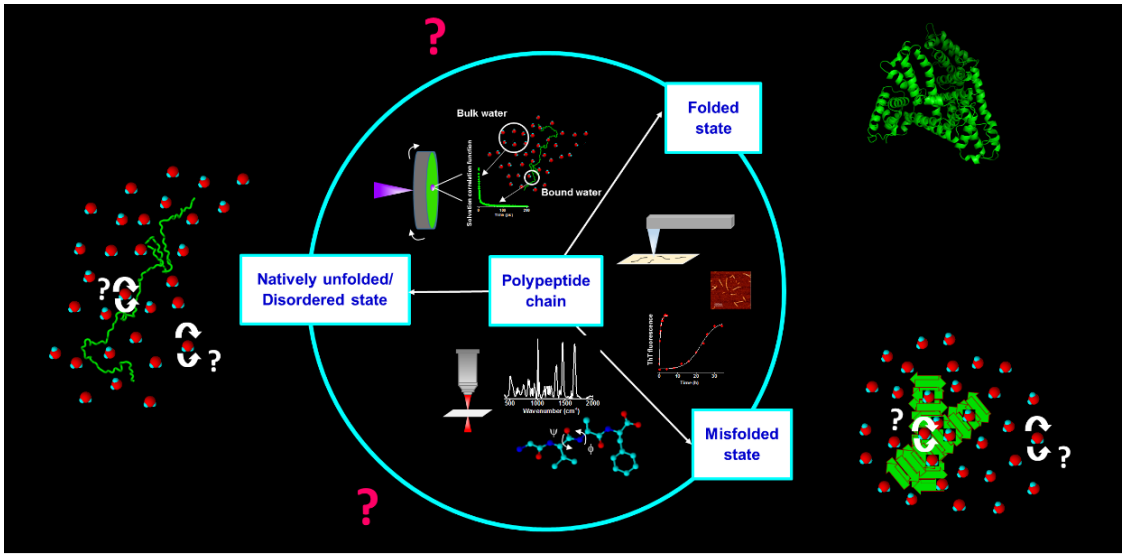


Figure 1. Schematic representation showing the questions addressed in this thesis along with the spectroscopic/imaging approaches that were utilized.

Chapter 2. Ordered Water within the Collapsed Globules of an Amyloidogenic Intrinsically Disordered Protein

Water molecules residing in the vicinity of the polypeptide chain play potentially important roles in directing the course of binding-induced folding and amyloid aggregation of IDP. In this work, we have characterized the nature of water molecules entrapped within the collapsed globules of an amyloidogenic IDP, namely, κ-casein. These globules undergo further compaction in the presence of an anionic detergent that is capable of diminishing the intra-chain repulsion from the positively charged glutamine/asparagine-rich amyloidogenic N-terminal domain comprising 100 residues. Using time-resolved fluorescence spectroscopy, we estimated the longer component of the solvation time to be ~1.4 ns, which is 3 orders of magnitude slower than that in bulk water and more than an order of magnitude slower than the "biological water" present at the protein surface. Profoundly restrained water within the collapsed IDP globules resembles the ordered water cluster found under nanoconfinement (Figure 2). We suggest that the association of these globules would result in the release of ordered water molecules into the bulk milieu causing an entropic gain that would eventually drive the formation of the key (obligatory) oligomeric intermediates on the pathway to

amyloids via nucleation-dependent polymerization. The change in the behavior of water molecules upon disorder-to-amyloid transition of κ-casein has been described in Chapter 3.

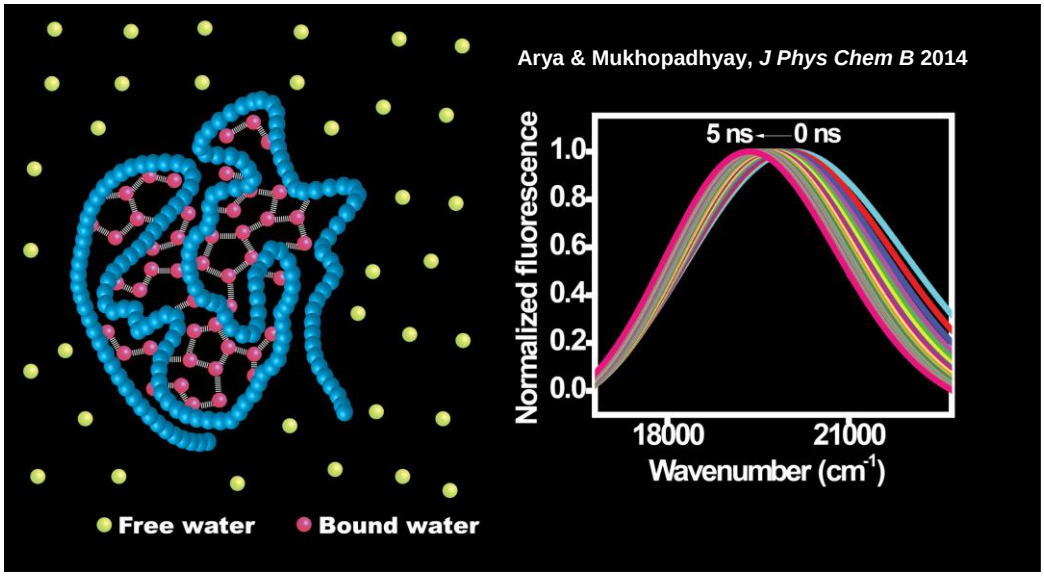


Figure 2. A schematic showing the ordered water molecules within the collapsed globule of κ-casein along with the time-resolved emission spectra associated with the solvation dynamics of these water molecules.

Chapter 3. Water Rearrangements upon Disorder-to-Order Amyloid Transition

Water plays a critical role in governing the intricate balance between chain-chain and chain-solvent interactions during protein folding, misfolding and aggregation. Protein aggregation leading to amyloid formation is implicated in a large variety of debilitating neurodegenerative diseases. Although a wealth of information is available on the mechanism of amyloid formation from the protein perspective, the role of water in protein aggregation remains elusive. In this work, using femtosecond and picosecond time-resolved fluorescence measurements, we have probed the solvation dynamics from ultrafast- to ultraslow timescale both in the monomeric- and in the amyloid-state of an intrinsically disordered protein, namely bovine κ-casein, the amyloid formation of which is linked with mammary cancer in cows. Our femtosecond results indicate that the mobility of "biological water" (type-I) gets restrained as a result of conformational sequestration during amyloid formation. Additionally, a significant decrease in the bulk water component with a concomitant increase in the ultraslow relaxation component revealed the ordering of interstitial water (type-II) upon disorder-to-order amyloid transition (Figure 3). Our results provide an experimental underpinning of profound water rearrangements associated with both chain desolvation and

water confinement upon amyloid formation. The detailed mechanism of amyloid formation from κ-casein is discussed in Chapter 5.

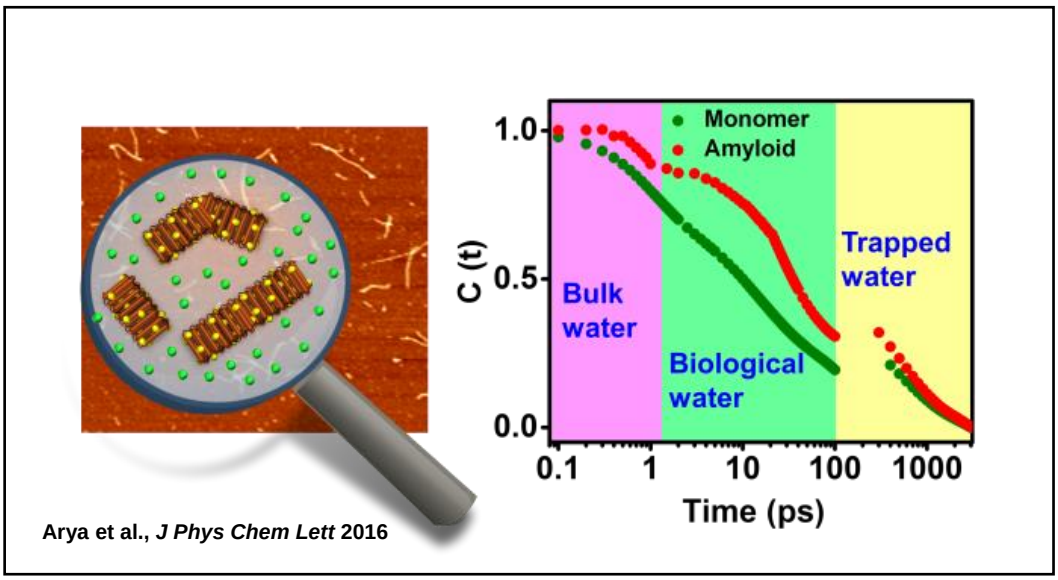


Figure 3. AFM (atomic force microscopy) image of κ-casein amyloid fibrils with the magnifying glass showing the presence of different types of water, that are further illustrated in the complete hydration correlation plot ranging over four orders of magnitude time scale constructed by combining the femtosecond up-conversion data and the picosecond TCSPC data for IDP (olive) and amyloid (red).

Chapter 4. Site-specific hydration dynamics of α-synuclein at the femtosecond time resolution

Amyloid aggregation of a presynaptic protein α-synuclein, which is an intrinsically disordered protein (IDP), underlies the pathogenesis of Parkinson's disease. The water present in the vicinity of the protein surface is believed to play a central role during this process and is known to strongly influence protein dynamics and function. Though the intimate relationship between proteins and water is very well appreciated, the behavior of water molecules in and around IDPs remains poorly understood. Here, using ultrafast laser spectroscopy, we have probed the hydration dynamics within natively unfolded α-synuclein with a femtosecond time resolution. By creating six cysteine mutations in N-terminus, NAC domain and C-terminus, respectively, we have been able to delineate the differences in the water arrangement within largely disordered α-synuclein. The timescales of surface bound water dynamics were found to be markedly different for different mutants. These differences in the water structure are anticipated to have a significant influence on the secondary

structure propensities within different regions of α-synuclein and might be tailored to the distinct roles of different domains in α-synuclein function and disease. Most intriguingly, the water molecules around the amyloidogenic NAC domain demonstrated retarded mobility compared to N- and C-terminus (Figure 4). Therefore, the presence of ordered water around the NAC domain underscores the importance of the NAC domain in the context of α-synuclein aggregation. In order to observe the release of water molecules from NAC domain upon amyloid formation, hydration dynamics studies were carried out with acrylodan labeled (Cys-78) α-synuclein. A drastic drop in the contribution of bulk water suggested the entropic liberation of ordered water molecules from the NAC domain upon fibrillation.

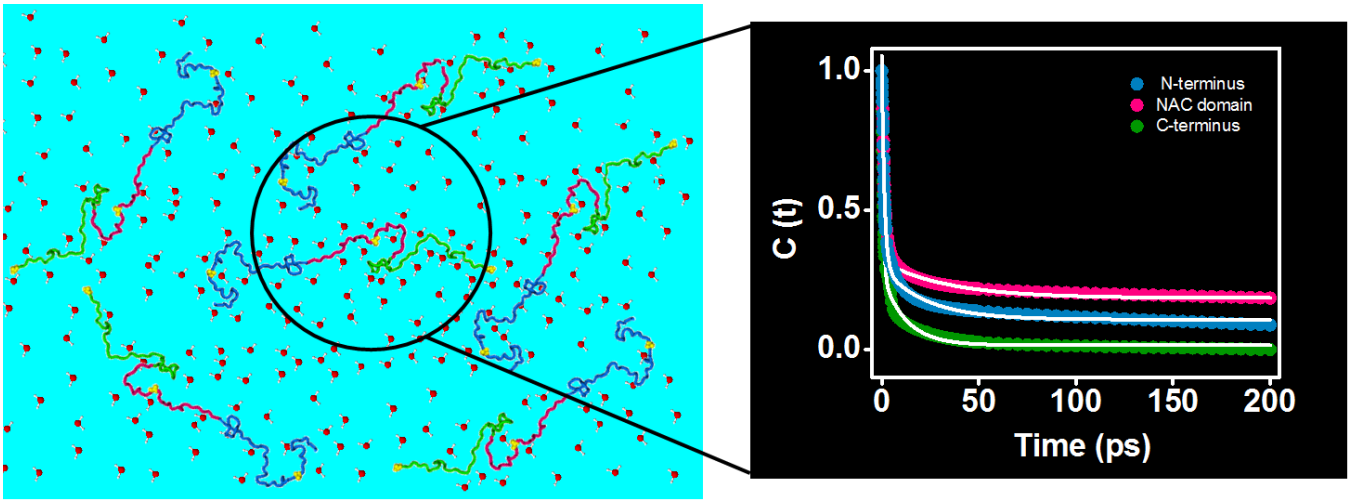


Figure 4. A schematic showing the conformational ensemble of α-synuclein, with water molecules present in the surrounding along with the hydration correlation function plots with different regions being shown in different colors (N- terminus in blue, NAC domain in red and C-terminus in olive).

Chapter 5. Mechanism of aggregation of a model amyloidogenic intrinsically disordered protein

The fact that many pathological amyloids are associated with intrinsically disordered proteins (IDPs) underscores the importance of IDPs in amyloid biology. The huge conformational plasticity associated with IDPs confers them the flexibility to adopt a large number of conformations depending upon their immediate biochemical environment. For example, many IDPs undergo a profound conformational change in the presence of lipids. In certain cases, lipid bilayers can also act as conformational catalysts and favor protein misfolding. Here, we have shown the effect of a well-known lipid mimetic, sodium dodecyl sulfate (SDS)

on the mechanism of aggregation of an amyloidogenic IDP, namely, bovine κ-casein using a variety of biophysical tools. The deposition of amyloids formed by κ-casein is linked with mammary cancer. The amyloid formation from κ-casein followed a nucleation dependent polymerization in the absence of SDS, as shown in Chapter 3. In the presence of SDS (sub-micellar concentration), a drastic change in kinetics and a switch in the mechanism was observed (Figure 5). Additionally, the morphologies of the aggregates formed in the absence and presence of SDS were found to be different (Figure 5).

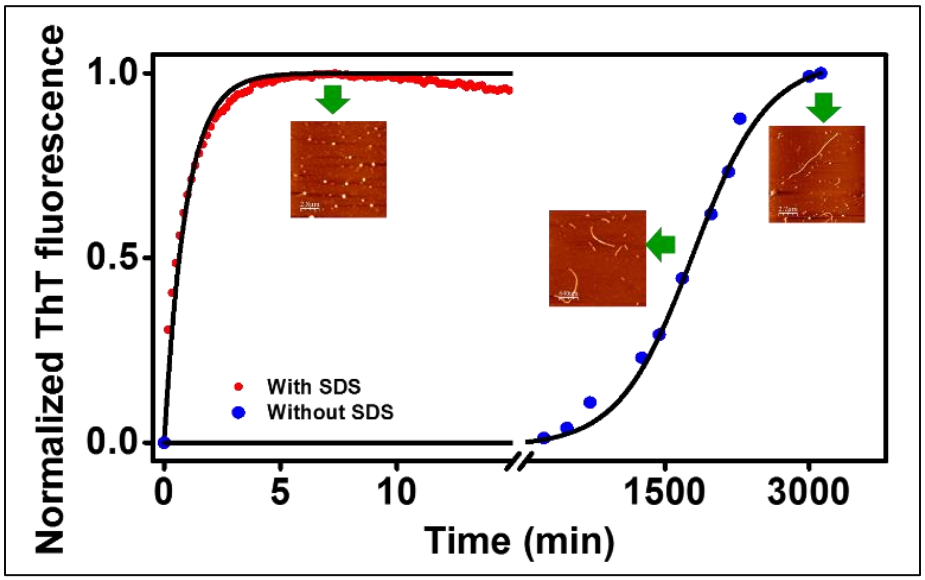


Figure 5. ThT aggregation kinetics along with AFM (atomic force microscopy) images showing two distinct pathways of aggregation followed by κ-casein. (A) Nucleation dependent polymerization occurring at high concentration leads to fibril formation. (B) Isodesmic polymerization occurring at low protein concentration, in the presence of SDS, leads to the formation of large spherical oligomeric species.

Chapter 6. Appearance of annular ring-like intermediates during amyloid fibril formation from human serum albumin

The self-assembly of proteins triggered by a conformational switch into highly ordered β-sheet rich amyloid fibrils has captivated burgeoning interest in recent years due to the involvement of amyloids in a variety of human diseases and a diverse range of biological functions. Here, we have investigated the mechanism of fibrillogenesis of human serum albumin (HSA), an all-α-helical protein, using an array of biophysical tools that include steady-state as well as time-resolved fluorescence, circular dichroism and Raman spectroscopy in conjunction with atomic force microscopy (AFM). Investigations into the

temporal evolution of nanoscale morphology using AFM revealed the presence of ring-like intermediates that subsequently transformed into worm-like fibrils presumably by a ring-opening mechanism (Figure 6). Additionally, a multitude of morphologically-diverse oligomers were observed on the pathway to amyloid formation. Kinetic analysis using multiple structural probes in-tandem indicated that HSA amyloid assembly is a concerted process encompassing a major structural change that is primarily mediated by hydrophobic interactions between thermally-induced disordered segments originating in various domains. A slower growth kinetics of aggregates suggested that the protein structural reorganization is a prerequisite for fibril formation. Moreover, time-dependent Raman spectroscopic studies of HSA aggregation provided key molecular insights into the conformational transitions occurring within the protein amide backbone and at the residue-specific level. Our data revealed the emergence of conformationally-diverse disulfides as a consequence of structural reorganization and sequestration of tyrosines into the hydrophobic amyloid core comprising antiparallel cross β-sheets.

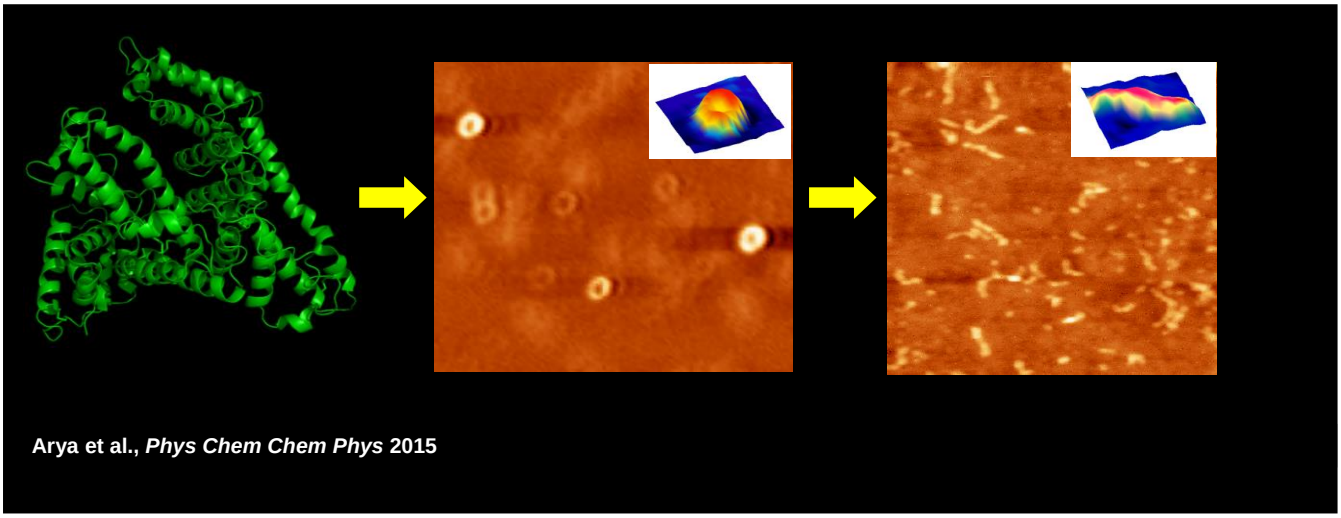


Figure 6. The amyloid formation from human serum albumin (HSA) involves the formation of ring-like intermediates that subsequently transformed into worm-like fibrils presumably by a ring-opening mechanism.

Chapter 7. Conclusions and future direction

The work described in this thesis provides insights into (i) the nature of water molecules in IDPs and amyloids, and (ii) the mechanism of amyloid formation from a model amyloidogenic IDP and from a partially disordered state of a model globular protein. This chapter provides a bigger picture obtained from the results and conclusions of the individual

chapters. The salient features of the thesis includes: (i) how the water arrangement in the native state of a collapsed IDP (Chapter 2 and 3) differs from the water structure in an expanded IDP (Chapter 4), (ii) the water rearrangement upon disorder-to-amyloid transition in a collapsed (Chapter 3) as well as an expanded IDP (Chapter 4) and (iii) the mechanism of amyloid formation using a model amyloidogenic IDP (Chapter 5) and a partially disordered form of a globular protein (Chapter 6). In the last part of this chapter, I discuss how the knowledge gained from this thesis can be utilized to understand the behavior of another interesting protein, namely, p53, which is a tumor suppressor protein and has a significant fraction of intrinsically disordered regions (IDRs) (37% of its structure intrinsically disordered). The importance of p53 in biology stems from the fact that it is a master regulatory protein that participates in cellular processes such as apoptosis, DNA repair, and cell cycle control. It functions as a homotetrameric nuclear phosphoprotein, and its function is lost in more than 50% of human cancers. The amyloid aggregates of p53 have been detected in tumor tissues. Therefore, aggregation might be a crucial aspect of cancer development, as p53 would lose its functions in an aggregated state. The strategies and some preliminary results for addressing the key aspects of p53 aggregation are discussed briefly in this chapter.

List of Publications

1. **S. Arya**, A. Singh, T. Khan, M. Bhattacharya, A. Datta & S. Mukhopadhyay, “Water Rearrangements upon Disorder-to-Order Amyloid Transition” ***J. Phys. Chem. Lett.*** **2016,** *7,* 4105-4110.
2. **S. Arya**, A. Kumari, V. Dalal, M. Bhattacharya & S. Mukhopadhyay, "Appearance of Annular Ring-like Intermediates during Amyloid Fibril Formation from Human Serum Albumin“ ***Phys. Chem. Chem. Phys. 2015***, *17*, 22862-22871.
3. **S. Arya** & S. Mukhopadhyay, "Ordered Water within the Collapsed Globules of an Amyloidogenic Intrinsically Disordered Protein" ***J. Phys. Chem. B 2014***, *118*, 9191–9198.
4. N. Jain, D. Narang, K. Bhasne, V. Dalal, **S. Arya**, M. Bhattacharya & S. Mukhopadhyay, “Direct Observation of the Intrinsic Backbone Torsional Mobility of Disordered Proteins” ***Biophys. J. 2016****, 4,* 768-774.
5. V. Dalal, **S. Arya** & S. Mukhopadhyay, "Confined Water in Amyloid-Competent Oligomers of the Prion Protein" ***ChemPhysChem 2016***, *17*, 2804-2807.
6. V. Dalal, **S. Arya**, M. Bhattacharya & S. Mukhopadhyay, "Conformational Switching and Nanoscale Assembly of Human Prion Protein into Polymorphic Amyloids via Structurally-Labile Oligomers" ***Biochemistry 2015***, *54*, 7505-7513.
7. S. Mukhopadhyay, V. Dalal & **S. Arya**, "Nanoscale Optical Imaging of Amyloid Fibrils in “Bionanoimaging: Insights into Protein Misfolding and Aggregation" in **Bio-Nanoimaging**: **Protein Misfolding** & **Aggregation** Elsevier 2013.
8. **S. Arya**, K. Bhasne, A. Singh, P. Dogra, A. Datta & S. Mukhopadhyay, “Site-Specific Hydration Dynamics of α-synuclein at the Femtosecond Time Resolution” (in preparation).
9. **S. Arya**, K. Bhasne & S. Mukhopadhyay, “Ordering of Water Molecules in the Amyloidogenic NAC Domain of α-synuclein upon Disorder-to-Order Amyloid Transition” (in preparation).
10. **S. Arya**, P. Dogra, N. Jain & S. Mukhopadhyay, “Mechanistic Insights into the Surfactant-induced Aggregation of a Model Amyloidogenic Intrinsically Disordered Protein” (in preparation).

Presentations and abstracts

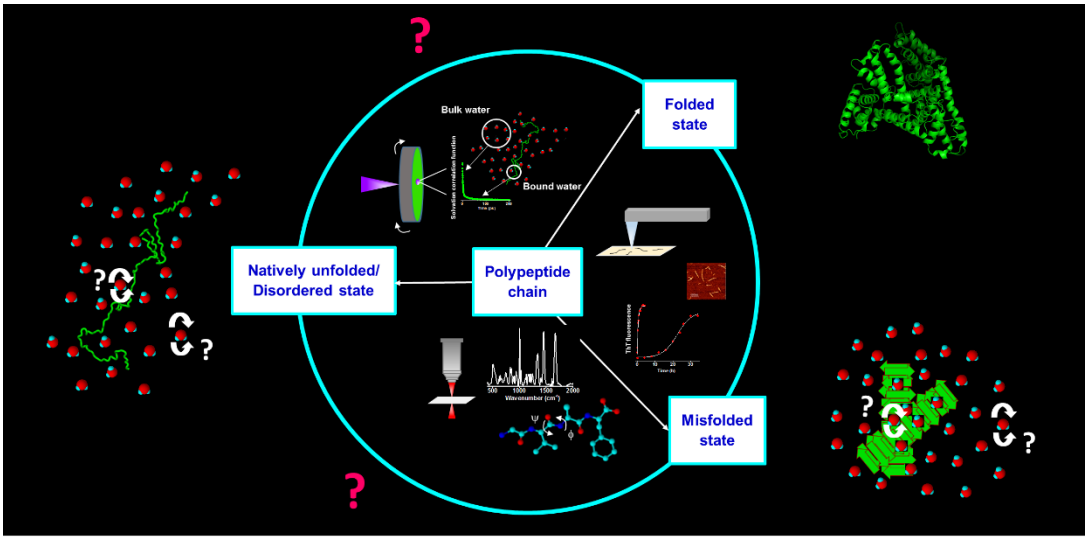
- **S. Arya** & S. Mukhopadhyay: Water in Amyloidogenic Intrinsically Disordered Proteins: Interplay of Conformational Preference and Amyloid Aggregation, 60th *Annual Meeting of the Biophysical Society (2016)*, Los Angeles, California, USA.

- **S. Arya**, M. Bhattacharya & S. Mukhopadhyay: Water in an Intrinsically Disordered Protein: From Chain Collapse to Amyloid Formation, *International Symposium on Advances in Spectroscopy and Ultrafast Dynamics (2014)*, Indian Association for the Cultivation of Science, Kolkata, India.

- **S. Arya**, N. Jain, M. Bhattacharya & S. Mukhopadhyay: Disorder to Order Transition in an Amyloidogenic Intrinsically Disordered Proteins: From Chain Collapse to Binding-induced Folding, *International Symposium on Protein Folding and Dynamics (2012)*, National Centre for Biological Sciences, Bangalore, India.

Table of Contents

Introduction



1.1 Protein folding, misfolding and aggregation

Proteins are the most ubiquitous and functionally versatile biological macromolecules that make up the fundamental molecular machinery of the cell and play a crucial role in the maintenance of life. Traditionally, in order to function, newly synthesized polypeptide chains generally fold into their 'native state', an ensemble of a few closely related and well-defined three-dimensional structures.[1-2] However, there is an emerging class of proteins known as natively unfolded or intrinsically disordered proteins (IDPs) that confront the traditional sequence-structure-function paradigm and remain unfolded in their native state.[3] These proteins have a special role in biology and are discussed later in great detail (See section 1.3). How the task of protein folding is accomplished remains a central problem in biology.[4] Nevertheless, considerable progress has been made in recent years through experiments and theoretical analyses, in understanding the protein folding problem.[1, 5-6] The total number of possible conformations that any protein chain can adopt are immense and a meticulous search for the native state structure would require an astronomical length of time. However, it is now quite clear that the process of protein folding involves a stochastic search of the diverse conformations accessible to a protein chain.[1, 5-6] In the current model of protein folding, the search of the unfolded polypeptide chain is described by a free energy landscape which views molecules as dynamic ensembles of rapidly interconverting conformational states.[7] The polypeptide chains are thought to navigate through the funnel-like energy surfaces, along various downhill routes, towards the native structure (Figure 1.1). The surface of this folding funnel is unique to a specific protein and is believed to be dictated by amino-acid sequence.[4] The folding landscape of small proteins (<100 amino acid residues) is relatively smooth and these proteins convert from the unfolded state to the native state (that are separated by a single transition state barrier), without the involvement of highly populated intermediates.[4, 8] However, for larger proteins (>100 amino acid residues), the free energy surface is often rugged in which partially folded intermediates may become transiently populated as kinetically trapped species during the folding process.[8-9] This seems to occur due to the higher tendency of larger proteins to undergo hydrophobic collapse, resulting in the formation of compact globular conformations that may have substantial native-like structural elements. These compact states may reorganize either into: (i) transient population of partially folded states lacking specific contacts and retaining large configurational entropy or (ii) intermediates that may be stabilized by non-native interactions. The search for the critical native contacts within the partially folded intermediates might be the rate limiting step for protein folding in the former case, whereas in

the latter case, the substantial reorganization event leading to the breakage of non-native interactions may be rate-limiting.[1] The presence of these intermediates *en route* to the native state can often be problematic as they tend to aggregate in a concentration-dependent manner.

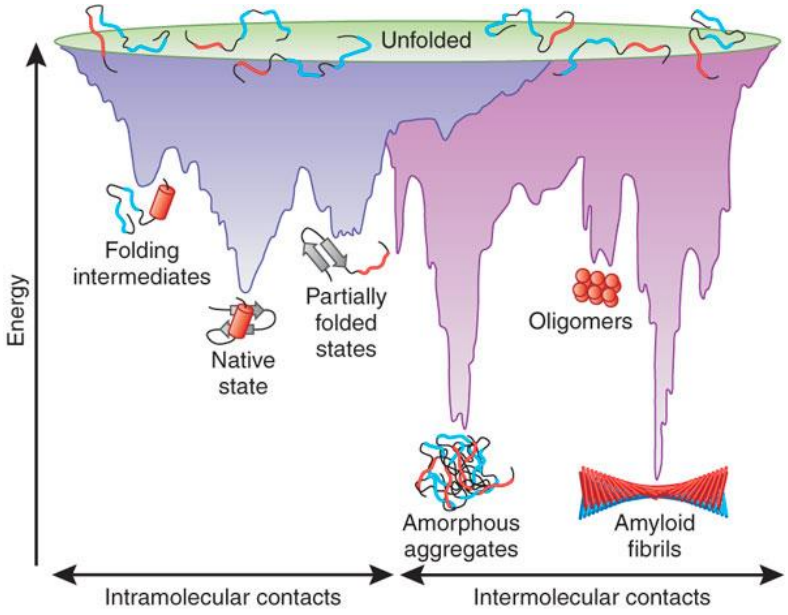


Figure 1.1. A schematic energy landscape of protein folding and aggregation. The purple surface corresponds to the diverse conformations 'funneling' towards the native state via intramolecular contacts and the pink surface corresponds to the conformations that lead to the formation of amorphous aggregates, toxic oligomers or ordered amyloid fibrils via intermolecular contacts. Protein aggregation may initiate from the intermediates that get populated during de novo folding or by the destabilization of the native state into partially folded states. The formation of aggregates is normally prevented by molecular chaperones. Reproduced with permission from [8].

The tendency of these intermediates to aggregate stems from the fact that these states have unstructured regions and hydrophobic amino-acid residues exposed to the solvent. The hydrophobic force that is known to drive the folding process, also is the major driving force in protein aggregation.[8] The protein folding process *in vivo* is very tightly regulated by the highly sophisticated chaperone machinery.[10] Molecular chaperones increase the efficiency of the folding process by restricting the formation of protein aggregates.[11] The protein aggregates become more prevalent under stress or when protein 'quality control' fails.[8] The protein aggregation can either lead to the formation of amorphous aggregates or a special kind of protein aggregates known as amyloids (Figure 1.1).

1.2 Amyloid state of proteins

Amyloid can be defined as a highly ordered, unbranched, fibrous quaternary structure which is composed of long filaments with a cross-β sheet arrangement in which the β-strands are oriented perpendicular to the fibril axis, and the inter-strand hydrogen bonds – as well as the amino acid side chains – are parallel to the fibril axis (Figure 1.2).[12] There is a growing body of evidence that suggests that the amyloid state is perhaps thermodynamically more stable state than the functional native state of many protein molecules, even under physiological conditions.[9, 11, 13-15]

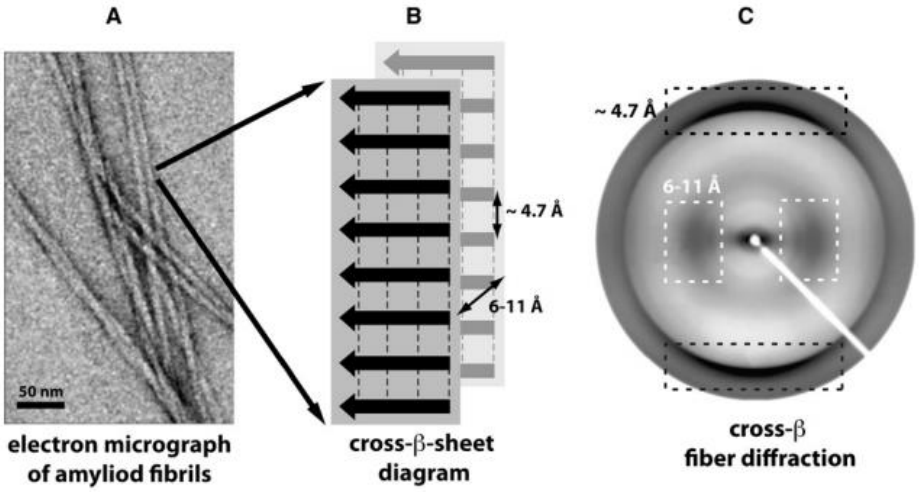


Figure 1.2. (A) Transmission electron micrograph of amyloid fibrils shows that they are composed of long filaments. (B) The schematic representation of the cross-β sheet arrangement within fibrils, with the backbone hydrogen bonds shown by dashed lines, and (C) the characteristic cross-β diffraction pattern with a meridional reflection at ~4.7 Å and an equatorial reflection at ~6–11 Å, that correspond to inter-strand and inter-sheet spacing, respectively. Reproduced with permission from [46].

The hallmark of amyloid fibrils is the typical X-ray diffraction pattern characteristic of the cross-β motif (Figure 1.2C).[16] They also exhibit apple-green birefringence under cross-polarized light when they bind the amyloidophilic dye Congo Red (see the next section).[17] Although the amyloids were initially identified in the context of disease, it became clear in the late 1990s that under certain environmental conditions, irrespective of their sequence, size and secondary structure, any protein can convert into aggregates with all the characteristics of the amyloid fibrils that are linked with disease.[4, 18] Thus, amyloid formation is now considered a generic property of proteins.[4, 18]

1.2.1 A historical perspective

The term 'amyloid' was first introduced by Schleiden in 1838 and then by Virchow in 1854 to describe the macroscopic abnormality of cerebral corpora amylacea that exhibited a positive iodine stain.[19] The macroscopic appearance of the brain tissue was similar to the "lardaceous" or "waxy" substances observed in the spleen, liver, and kidneys of deceased patients, perhaps as early as 1639 by Fontanus.[20] Virchow was convinced that the cerebral corpora amylacea was identical to starch. Therefore, the amyloid deposits were thought to be carbohydrate in nature until their high nitrogen content was established later in 1859 by Friedreich and Kekule.[19] The first report on primary amyloidosis came in 1856 by Wilks.[21] Further advancements in understanding of amyloid structure were made with light microscopy and histopathological dyes such as thioflavin (ThT) and Congo red. ThT was found to bind avidly to the amyloid deposits in vivo.[20] However, due to its interaction with other tissue components, it was considered less useful for diagnostic histology. Congo red is an aniline dye that became quite popular initially for its use in the textile industry (in 1883) and later on (in 1886), for tissue staining.[22] In 1922, Congo red was found to bind amyloids quite strongly.[22] The specific binding of amyloids to Congo red produced an apple green birefringence when examined using polarization light microscope, under cross polarized light.[20, 22] Thus, Congo red binding with apple green birefringence was the very first criterion of amyloid to be adopted. The observed birefringence was suggestive of an ordered submicroscopic structure. However, until the 1950s, the nature of amyloid structure remained elusive. In 1959, using electron microscopy, Cohen and Calkins were the first one to demonstrate that all types of amyloid have fibrillar structure.[23] Thus, fibrillar morphology, observed using negatively stained tissue sections became the second criterion for identifying amyloid. In 1968, Eanes & Glenner reported the X-ray fiber diffraction pattern of amyloid fibrils.[24-25] The amyloid fibrils exhibited a characteristic cross-β fiber diffraction pattern. In fact, X-ray diffraction pattern for amyloid fibrils formed in vitro was found to be similar to amyloids extracted from the tissues.[25] The cross- β diffraction pattern was first observed in 1935 by a well-known biophysicist Astbury[26] and is now considered a hallmark of amyloid structure. Further advancement in biochemical and biophysical techniques facilitated the isolation of amyloid fibrils from tissues and a significant progress has been made in understanding amyloid fibril structure since the early 1970's. A detailed analysis of the amino acid sequence composition of the proteins that comprise a range of ex vivo amyloid fibrils indicated that each amyloid disorder was associated with a particular protein or peptide.[27] There are atleast 20 plasma proteins that are known to

form amyloid, until now. It is interesting to note that despite the differences in the amino acid sequences and the native structure of the proteins, the amyloid fibrils formed by them appear to have similar morphological features with characteristic cross-β core.[16, 24] The common cross-β core can account for the common biochemical characteristics that amyloids share, for example, protease resistance nature of most amyloids. The amyloid from human immunoglobulin light chain isolated from the tissue of a patient suffering from "primary amyloidosis" was the first amyloid to be identified and characterized chemically.[28] Further, it was found that the lysosomal extracts were able to digest immunoglobulin light chain precursor protein into amyloidogenic fragments.[28] This observation was followed by several other biochemical investigations that led to the identification of amyloidogenic fragments in about 20 soluble proteins, which in turn led to a general realization that many of the amyloid forming peptides are often formed by the proteolytic processing of the respective precursor proteins.[29] In 1982, the 'prion' hypothesis describing an infectious protein particle capable of causing scrapie (which was first recognized in western Europe in the 1700s), a fatal neurodegenerative disease in sheep, was introduced by Prusiner.[30] The protein responsible was amyloidogenic in nature but unlike other proteins, it was found to be infectious and was termed as prion protein (PrP). The proteinaceous agent involved in the transmission of disease was termed prion and it was found that these prions are the altered forms of cellular proteins that can transform the unaltered (non-infectious) form into the altered (infectious) form. This positive-feedback feature underlies the basis for the self-propagation and infectivity of prions. Two years later, Aβ peptide fragment from the amyloid precursor protein (APP), which is the major component of amyloid plaques found in the brain of Alzheimer's patients, was identified and characterized.[31] However, in 1986, prions again captivated enormous attention upon the outbreak of Bovine spongiform encephalopathy (BSE), a prion disease similar to scrapie in sheep. Ten years later, there was an onset of the human prion disease variant Creutzfeldt-Jakob disease (vCJD). The transmissibility of prions from one species to other could not be denied in case of vCJD which was strongly linked with the exposure to the BSE agent. Around this time, prion diseases and amyloid related disorders became subject of intense interest. In 1994, the yeast prions were discovered and were found to form amyloids.[32-33] The mild effects of some of the yeast prion variants and the occurence of the clearly functional [Het-s] prion of *Podospora anserina* led to a belief that yeast prions may actually benefit their hosts. Like yeast prions, many other proteins capable of forming amyloids with functional importance for the host were identified in various lower and higher organisms, including humans. Such amyloids

are known as "functional amyloids".[34-35] Subsequent years witnessed a remarkable progress in the understanding of amyloids and prions, especially the structural aspects.

1.2.2 Characteristic features of amyloid structures

Amyloid fibrils are often insoluble and heterogeneous in nature, and hence are not amenable to the commonly used methods for structure determination such as X-ray crystallography and solution nuclear magnetic resonance (NMR).[29, 36-37] However, many important structural characteristics of amyloid fibrils have progressively emerged from cryo-electron microscopy[38-39], electron paramagnetic resonance (EPR)[40], solid-state NMR (ss-NMR)[41], and single-crystal X-ray diffraction[42]. The structural insights obtained from all these techniques suggest that all amyloid fibrils share a common core cross-β structure in which β-sheets run parallel to the fibril axis, with their hydrogen-bonded β-strands arranged orthogonally with respect to the fibril axis. Recent X-ray crystallographic studies on amyloids formed by short amyloidogenic peptides have revealed the presence of side-chain interactions which stabilize the association of β-sheets that can be further categorized as polar (hydrogen bonding or electrostatic) and non-polar (π-π stacking and hydrophobic) interactions.[43-45] The atomic-level structural information obtained for microcrystals of short amyloidogenic peptides or segments of amyloid-forming proteins indicated the presence of two distinct types of β-sheet stacking interfaces that were termed as dry interface and wet interface.[46] The dry interface, also known as 'dry steric zipper', represents an interface that is mostly devoid of water and hence, mainly involves complementary side-chain interactions that result in dense packing.[46] In contrast, the wet interface is mostly composed of inter-strand hydrogen bonds.[46] Although complementarity is expected for steric-zipper formation, the sequences involved may not be self-complementary as is evident from solid-state NMR studies on Aβ and Het-s, a fungal prion that form cross-complementary sheets. Such steric-zippers are known as 'heterosteric zippers'.[47] The structural information obtained for many of the amyloidogenic polypeptides and proteins suggests that amyloids can have different structural arrangements. Due to differences in the packing of the protofilaments, numerous morphologies of mature fibrils can arise, leading to a phenomenon called 'amyloid polymorphism'.[48] The phenomenon of polymorphism describes the multiple structural forms and emphasizes the diverse and dynamic nature of amyloid fibrils. Polymorphism in amyloids has been differentiated by electron microscopy (EM) and NMR studies.[48] Although the details of amyloid polymorphism need to be investigated, some initial studies have indicated that the morphological diversity can result from the changes in conditions such as physicochemical environment, seeding and template effects.[49] The variation

in structure at the molecular level due to differences in the packing of either the amino acid side chains or a few peptide stretches (segments) gives rise to packing and segmental polymorphism, respectively.[48] In packing polymorphism, the nature and position of the side chains present on the outer side of the fibrils differ and results in different morphologies for the two polymorphs. Therefore, in packing polymorphs, the sequence involved in the formation of the cross-β core remains the same, only the packing is different. In case of segmental polymorphism, two or more segments from the same amyloid forming protein can give rise to different morphologies, and hence the sequence involved in the formation of the cross-β spine is different for the two segmental polymorphs, as has been observed for Pmel 17 fibrils.[50] The phenomenon of amyloid polymorphism is believed to play a crucial role in the biology of prion propagation, the strain phenomenon and the species barrier.[51-53] The generation of distinct prion strains that give rise to structurally and morphologically different amyloid fibrils can be explained by different models that explain polymorphism in amyloids. Based on the atomic structure information available on amyloids, various models have been proposed for amyloid polymorphs.[54] Various structural arrangements possible for amyloids are summarized below.

(i) Parallel arrangement of β-strands: In this case, the β-strands are aligned in such a manner that the side chains of the respective amino acid residues are positioned on top of one another along the fiber axis. The backbone hydrogen bonds are arranged parallel to the fiber axis, and the interstrand distance, also termed as 'main chain spacing'[55], is 4.7–4.8 Å (Figure 1.3A). This type of packing is known as parallel in-register packing and is the most commonly observed architecture among pathological as well as functional amyloids. This kind of morphology has been observed for amyloid fibrils formed by α-synuclein[56] and β_2-microglobulin[57].

(ii) Antiparallel arrangement of β-strands: The antiparallel arrangement of β-strands is the commonly observed β-sheet arrangement in globular proteins[58]. In the antiparallel arrangement, the β-strands run antiparallel to each other (Figure 1.3B). Antiparallel β-sheet structures have been observed for fibrils formed by short peptides, such as fragments of Aβ[59].

(iii) β-helical or nanotube structure: When one or more β-sheets, arranged in a parallel manner, wrap around a hollow core forming a helical structure; the structure so formed is known as a β-helical or nanotube structure (Figure 1.3C). Intramolecular backbone hydrogen bonding is observed between the parallel strands present in the β-helix. For fibrils formed by polyglutamine peptide, a water-filled nanotube β-helical model has

been proposed based on X-ray diffraction patterns[60]. β-helical models has also been proposed for the NM-domain of Sup35 yeast prion protein using pyrene excimer fluorescence studies[61].

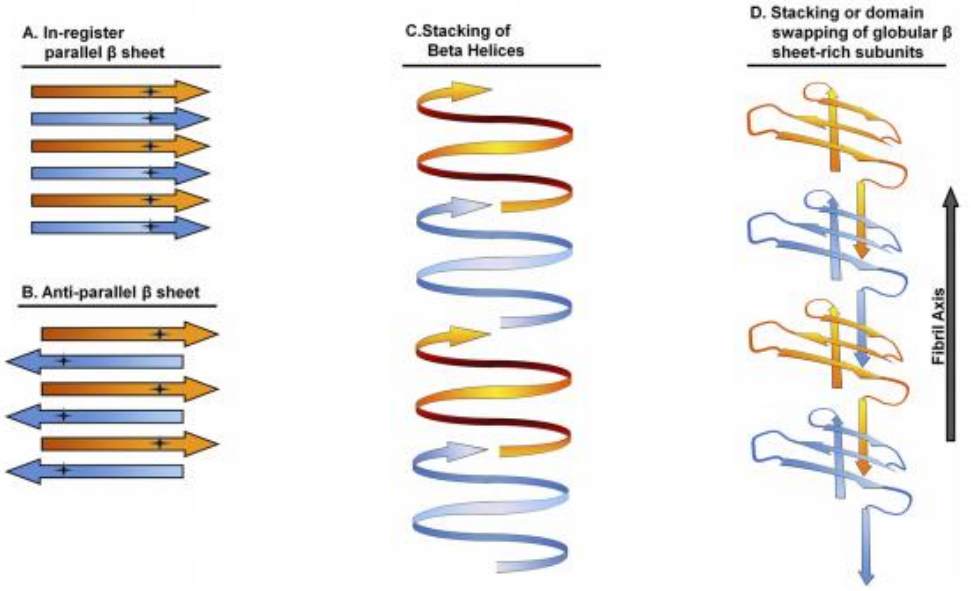


Figure 1.3. Various structural arrangements possible for amyloids (A) parallel in-register -sheet structures are composed of individual polypeptides stacking in-register every 4.7 Å along the fibril axis (B) antiparallel β-sheet structures are also composed of polypeptides stacking every ~4.7 Å, but β-strands alternately run in opposite directions. (C) β-helices are composed of a single polypeptide wrapping around an axis, forming intramolecular parallel β-sheets. (D) some β-sheet-rich proteins can linearly assemble into filamentous structures by domain swapping mechanism. Reproduced with permission from [58].

(iv) Domain swapping: In certain proteins such as serine protease inhibitors, aggregation involves linear domain swapping.[62] X-ray crystallographic studies on a stable serpin dimer suggested that the fibrillar structure is formed via domain-swapping mechanism, in which a β-hairpin of one promoter gets inserted into the β-sheet of the next one (Figure 1.3D). Studies on trimeric serpin suggested an entirely different mechanism of domain swap (C-terminal swap).[63] Many other proteins, such as immunoglobulin-binding domain B1 of streptococcal protein G, and cystatin, have also been proposed to aggregate or misfold through domain-swapping mechanism.[64]

1.2.3 Mechanism of amyloid formation

The conversion of a protein from its functional soluble state to the amyloid state is a fairly complex process and involves a multiple precursor species, which are often much smaller in length and width than mature amyloid fibrils.[65] The rate of fibril formation, the nature of intermediates getting populated during fibrillation and the lifetime of the intermediate species are believed to dictate the onset and the progression of amyloid diseases.[66] Therefore, in vitro studies of amyloid formation are critical for understanding the role of intermediates in the disease and for developing drugs that can intervene the fibrillation.[67] Major advances in our understanding of the molecular mechanism that underlie protein aggregation have been made in past few years. However, the understanding of the mechanism of amyloid formation still remains elusive. The process of amyloid formation can be described by both protein polymerization and protein aggregation.[68] Polymerization is more specific while aggregation implies a non-specific process.[68] However, since amyloid formation involves both specific and non-specific interactions, the distinction between polymerization and aggregation becomes obsolete and therefore, can be interchangeably used to describe amyloid formation. Although a variety of models to describe the mechanism of amyloid formation have been proposed, the two of the polymerization mechanism that have existed for a long time and reflect two extreme scenarios are: - (i) Nucleation-dependent polymerization and (ii) Isodesmic polymerization (Figure 1.4).

(i) Nucleation-dependent (or nucleation elongation) polymerization

Typically, amyloid formation follows nucleation dependent polymerization (NDP) pathway (Figure 1.4A).[69-70] In an NDP reaction, the initial steps are slower compared to the later steps. A generic nucleation elongation process involves: (i) formation of nucleus (nucleation phase), and (ii) addition of monomers to the critical nucleus formed (elongation phase). The forward and reverse reaction rate constants are required for each step in order to provide a comprehensive mathematical depiction of an NDP reaction. The initial steps in an NDP reaction (nucleation) consist of a number of unfavorable equilibria or in other words, the free energy barrier associated with nucleation is high. Once the barrier is overcome or the nucleus is formed, the polymerization (elongation) becomes thermodynamically favorable. Therefore, in the nucleation phase, the dissociation is much more favored than the association. Once the nucleus is formed, the elongation phase is reached wherein the association rate constants are much higher compared to the dissociation rate constants.

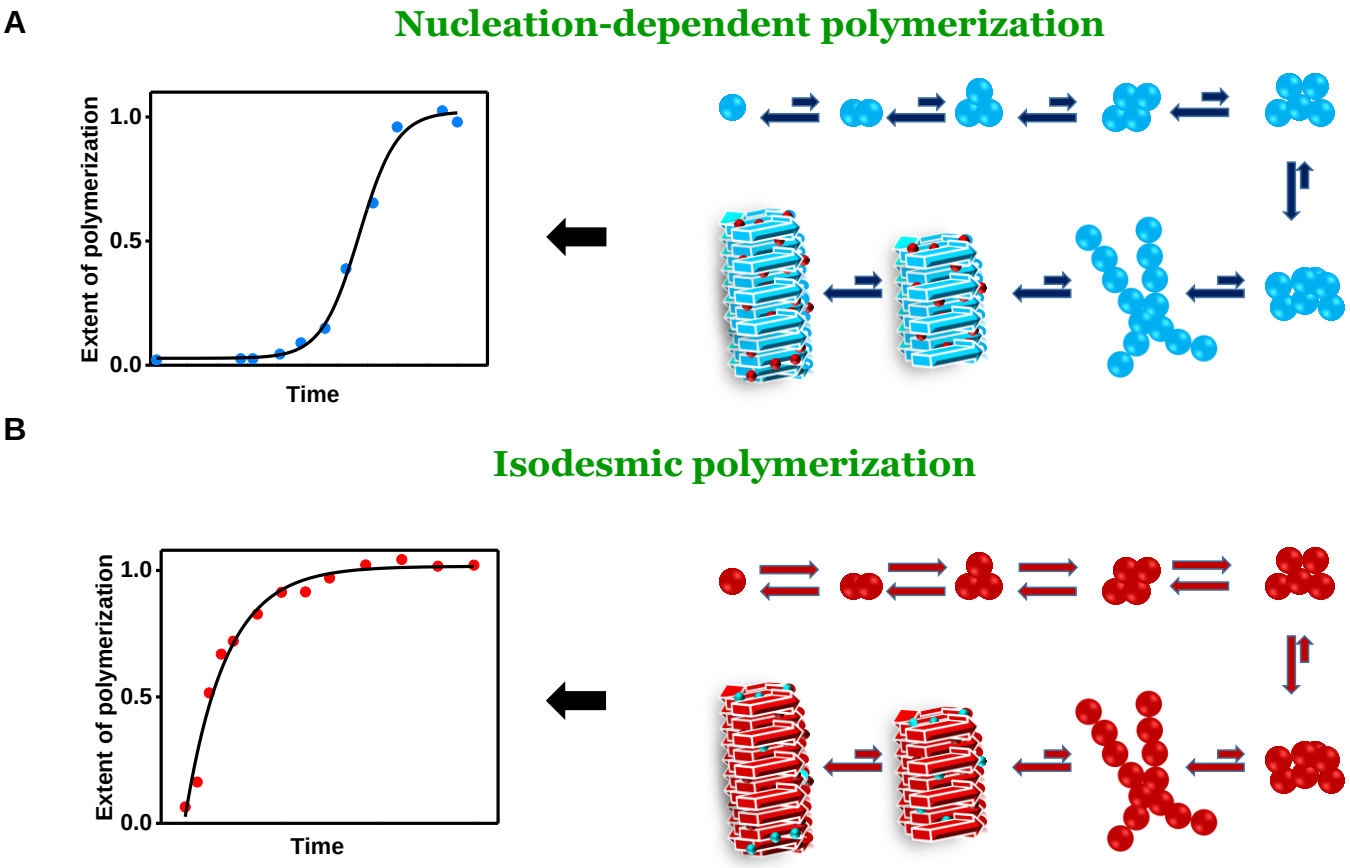


Figure 1.4. Protein aggregation mechanism. Schematic representation of (A) nucleation-dependent polymerization (B) isodesmic polymerization.

An NDP reaction has following characteristics that make them different from isodesmic polymerization reaction[71-72]: (i) A typical lag phase is seen during an NDP reaction. The duration of lag time is proportional to the sharpness of the energy curve in the initial phase and depends on protein concentration. The size of nucleus also plays a critical role, (ii) The addition of a preformed nucleus at the beginning of an NDP reaction leads to the disappearance of the lag phase, as a result of seeding, a process analogous to phenomenon of crystallization (The preformed nucleus added is referred to as seed and the reaction taking place in the presence of seed is called seeding reaction)[73], and (3) There is a critical concentration associated with the formation of polymer. The critical concentration is that concentration of monomeric protein,

below which the polymerization would not take place. It is typically determined from a plot of the rate of polymer formation (or amount of polymer) versus protein concentration. This concentration would be different for different proteins. The nucleation–elongation polymerization must fulfill all three criteria because even in the isodesmic case, at least two of the three features described above can be observed (see below).

(ii) Isodesmic (linear) polymerization

Unlike NDP, there are no discrete nucleation and elongation phases (Figure 1.4B).[68, 74] No lag phase is seen in case of isodesmic polymerization and the reaction is fastest at the beginning of the reaction where the concentration of monomers is highest and gets slowed down as the equilibrium is reached. There is no critical concentration barrier. In general, an isodesmic polymerization reaction does not demonstrate any of the three characteristic features of an NDP reaction that are mentioned above. However, it is not always easy to differentiate between the two polymerization pathways, because the differences between them are rather subtle, and rely solely on the size of the nucleus and the rates for dissociation and association reactions. Therefore, under certain circumstances, isodesmic polymerization reactions can readily switch to the NDP mechanism.[68]

In this thesis, the mechanism of amyloid aggregation from a model amyloidogenic IDP κ-casein and disordered state of a model globular protein human serum albumin (HSA) has been discussed in Chapter 5 and Chapter 6, respectively.

Table 1.1 **A list of human diseases associated with amyloid formation along with the details of the proteins or peptides and the tissues involved.**

Disease	Aggregating protein or peptide	Structure of protein or peptide	Tissue distribution of protein deposits
Neurodegenerative diseases			
Alzheimer's disease	Amyloid-β peptide and Tau	Intrinsically disordered	Brain
Spongiform encephalopathies	Prion protein or its fragments	Intrinsically disordered and α-helical	Brain
Parkinson's disease	α-synuclein	Intrinsically disordered	Brain
Amyotrophic lateral sclerosis	Superoxide dismutase 1	β-sheet and Ig-like	Brain and spinal cord
Huntington's disease	Huntingtin fragments	Mostly intrinsically disordered	Brain
Familial amyloidotic polyneuropathy	Transthyretin mutants	β-sheet	Brain
Non-neuropathic systemic amyloidosis			
Amyloid light chain (AL) amyloidosis	Immunoglobulin (Ig) light chains or its fragments	β-sheet and Ig-like	Almost all organs and tissues
Amyloid A (AA) amyloidosis	Serum amyloid A1 protein fragments	α-helical and unknown fold	Bladder, stomach, thyroid, kidney, liver, spleen, GI tract
Senile systemic amyloidosis	Wild-type transthyretin	β-sheet	Almost all organs and tissues
Haemodialysis-related amyloidosis	β_2-microglobulin	β-sheet and Ig-like	Musculoskeletal tissues, peripheral nervous system, GI tract, tongue, heart, urogenital tract
Lysozyme amyloidosis	Lysozyme mutants	α-helical and β-sheet	Several visceral organs and tissues
Non-neuropathic localized amyloidosis			
Apolipoprotein A1 (Apo A1) amyloidosis	Apo A-1 fragments	Intrinsically disordered	Eyes
Type II diabetes	Amylin	Intrinsically disordered	Pancreas
Injection-localized amyloidosis	Insulin	α-helical and insulin-like	Skin, muscles

1.2.4 Diseases associated with amyloid formation

There are now approximately 50 disorders which are associated with the aberrant aggregation of normally soluble, functional proteins or peptides, into intractable aggregates, often amyloid

fibrils.[65] A list of some of these amyloid-related human diseases is given in Table 1.1, along with the specific proteins associated with them. The diseases have been broadly classified on the basis of the region(s) that is (are) involved: (i) Neurodegenerative diseases: aggregation occurs in the brain, (ii) Non-neuropathic systemic amyloidosis: aggregation occurs in multiple tissues, and (ii) Non-neuropathic localized amyloidosis: aggregation occurs in a single type of tissue other than the brain. It is interesting to note that many pathological amyloids are associated with proteins that are either IDPs or have intrinsically disordered regions (IDRs) (Table 1.1).

1.3 Intrinsically disordered or natively unfolded proteins

Traditionally, for most proteins, the function is believed to be dictated by its unique, well-defined three-dimensional structure. This classical structure-function paradigm has mainly emerged from the concepts explaining the specificity of enzymes (the famous lock-and-key hypothesis).[75] Though most of the proteins adopt a well-defined three-dimensional structure in order to perform their respective functions, a significant fraction of the proteome of any

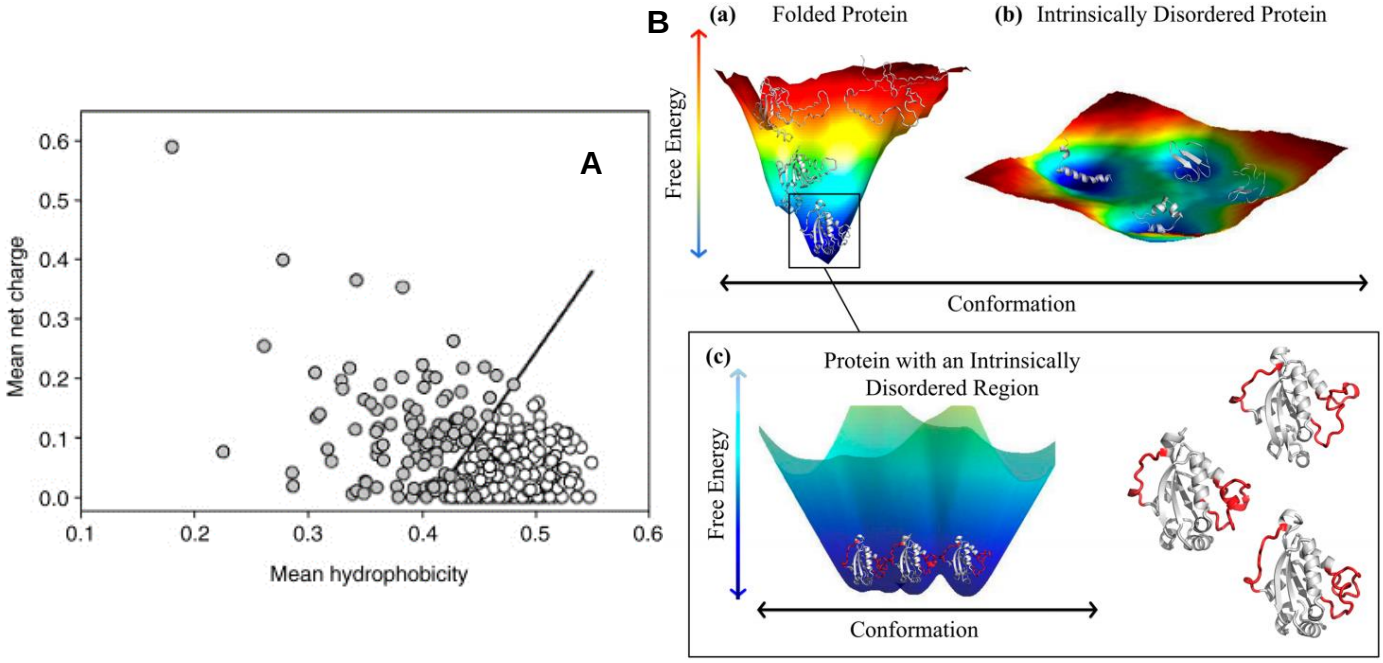
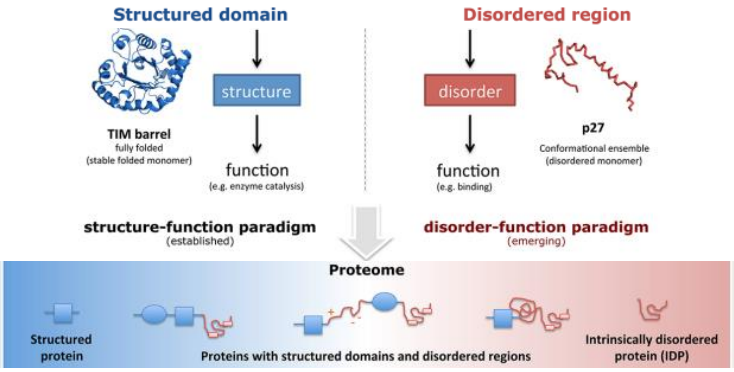


Figure 1.5. (A) The charge-hydropathy plot comparing ordered/folded proteins (275, open circles) and disordered proteins (91, grey circles) based on their net charges and hydropathies. The solid line represents the border between folded and disordered proteins. Reproduced with permission from [89]. (B) A schematic representation of energy landscapes for (**a**) a folded protein (human nucleoside diphosphate kinase (NDPK), PDB ID: 1nsk) and (**b**) an intrinsically disordered peptide (CcdA C-terminal, PDB ID: 3tcj); (**c**) close-up of the minimal free energy well in (**a**), where IDRs are shown in red and ordered regions are shown in white. Reproduced with permission from [93].

organism contains proteins or protein segments that lack the ability to fold to form a well-defined three-dimensional structure, but are still functional.[3, 76-91] These proteins and protein segments are referred to as intrinsically disordered proteins (IDPs) and intrinsically disordered regions (IDRs), respectively.[92] One of the frequently used prediction methods for IDPs is based on a very simple model which proposes that the repulsion from like charges favors unfolding while increased hydrophobicity favors folding. Using this approach, a charge-hydropathy (C-H) plot is constructed by plotting normalized net charge against normalized hydropathy (calculated from the Kyte-Doolittle hydropathy scale).[79] This simple C-H plot beautifully separates IDPs from structured/folded proteins (Figure 1.5A). The difference between folded and IDPs can further be realized by their potential energy surfaces. As discussed in section 1.1, the proteins typically explore the funnel-like energy surface during their search for the native structure. This holds true for folded proteins because IDPs have relatively flat energy surface (Figure 1.5B).[93]

In case of folded proteins, there is a global energy minimum which is the lowest energy state and corresponds to the native structure. The width of the global energy minimum is proportional to the conformational entropy associated with the native state. On the contrary, IDPs have multiple local energy minima that are separated by small barriers. The interconversion between different local energy minima can take place quickly, leading to an ensemble of a large number of conformational states having approximately equal energies.[93] Therefore, IDPs exist as dynamic ensembles of rapidly fluctuating, interconvertible structures under physiological conditions. Additionally, the energy minimum of a folded protein with an IDR is not smooth due to many smaller minima corresponding to different states sampled by the IDR within the folded protein in its native state (Figure 1.5B).[93] The conformational plasticity of IDPs and the absence of rigid structure confers on them a large number of functional advantages over folded proteins.[89] This explains why intrinsic disorder is a common feature of proteins involved in signaling, regulation, and recognition. The dynamic behavior of IDPs can also account for their diverse roles in modulating the functions of their binding partners and in promoting the association of supramolecular complexes. The biological functions of IDPs are often controlled by post translation modifications (PTMs) such as phosphorylation, ubiquitination, acetylation and alternative splicing.[94] Many IDPs are known to have multiple functional elements as a result of which they can interact with numerous structurally unrelated binding partners.[94] Therefore, in the bound state, the same protein can now have very different structures depending upon the binding partners. However, many IDPs

remain disordered upon binding their partners and still retain their function.[93] In case of IDPs that undergo folding upon binding their partners, the question that whether folding occurs prior to, during, or after contacting the partner is of significant interest due to its important implication in the designing of molecules that can potentially inhibit or stabilize the IDP state of protein. In order to understand this, various hypotheses have been proposed. According to the conformational selection hypothesis, unbound IDP fluctuates through their bound conformations and the binding partner selectively binds the IDP, the moment it attains the appropriate binding conformation.[95] On the other hand, the induced fit hypothesis proposes that IDP first make low-affinity and non-specific contacts with its binding partner and then gets folded upon binding.[96-97] Another explanation of binding-induced folding can be given by fly-casting which states that extended IDP conformations are associated with a relatively large capture radius which leads to fast association rates.[98-99] Though the high capture radius has not been shown to substantially increase the binding rate of IDPs over folded proteins, the possibility of the weak complexes transiently formed with an extended IDP to evolve into the native complex was found to be higher compared to the folded form of IDP.[100] In general, IDPs can switch between bound and unbound state and therefore return back to their disordered state after the completion of a particular function. These unique and dynamical properties that IDPs possess make them particularly different from folded proteins and allow them to perform a diverse array of functions that folded proteins are incapable of. Various disorder predictors



applied on different

Figure 1.6. Structured domains and intrinsically disordered regions (IDRs) are two fundamental classes of functional building blocks of proteins. The synergy between disordered regions and structured domains increases the functional versatility of proteins. Reproduced with permission from [102].

proteomes have revealed that IDPs are highly abundant in nature and the general extent of disorder in proteins increases on going from bacteria to eukaryotes.[80, 85, 101] The high occurrence of IDRs in any genome in combination with their unique properties calls for the amendment in

the classical view of the structure−function paradigm and hence that of protein function.[102] Therefore, now, structured domains and intrinsically disordered proteins/regions (IDP/Rs) are considered two fundamental classes of functional building blocks of proteins (Figure 1.6).[103]

Table 1.2 **Functional classification of IDPs**

Protein	Partner	Function
Entropic chains		
Nup2p FG repeat region	-	Gating in nucleoporin complex (NPC)
MAP2 projection domain	-	Spacing in cytoskeleton
Titin PEVK domain	-	Elasticity of muscle
Display sites		
CREB KID	PKA	Site of phosphorylation
Cyclin B N-terminal domain	E3 ubiquitin ligase	Site of ubiquitination
Chaperones		
β-synuclein	α-synuclein	Prevention of aggregation
ERD 10/14	luciferase	Prevention of aggregation
hnRNP A1	DNA	Strand re-annealing
Effectors		
4E-BP1	eIF4E	Inhibition of translation initiation
$p27^{Kip1}$	Cyclin A-Cdk2	Inhibition of cell-cycle
Securin	Separase	Inhibition of anaphase
Assemblers		
$p21^{Cip1}$	Cyclin A-Cdk2	Assembly of Cyclin-Cdk complex
CREB	p300/CBP	Initiation of transcription
Scavengers		
Caseins	Calcium phosphate	Solubilization of calcium phosphate clusters in milk
ERD 10/14	Water	Retention of water in dehydration stress

1.3.1 Functional classification of IDPs/IDRs

Various classifications of IDPs and IDRs are available in the literature since the start of meticulous studies into their functions about 17 years ago. They can be classified based on function, structure, sequence, functional elements, protein interactions, evolution, regulation etc. Based on their function, these IDPs/IDRs can be broadly divided into six categories[104]:

(i) Entropic chains: IDPs that fall in this category carry out functions that take advantage of their high degree of conformational disorder.

(ii) Display sites: The conformational flexibility of IDPs as display sites facilitates the exposure of motifs and easy access for proteins that introduce and recognize PTMs.

(iii) Chaperones: The versatility of IDPs that fall in this category, and their unique binding properties (many different partners, rapid association/disassociation, and folding upon binding) are well suited for chaperone functions.

(iv) Effectors: The IDPs that fall in this category modify the activity of their partner proteins. Upon binding their partners, they often undergo a disorder-to-order transition.

(v) Assemblers: The IDPs that belong to this class have large binding interfaces that scaffold multiple binding partners and promote the formation of supramolecular protein complexes.

(vi) Scavengers: These IDPs are known to store and neutralize small ligands.

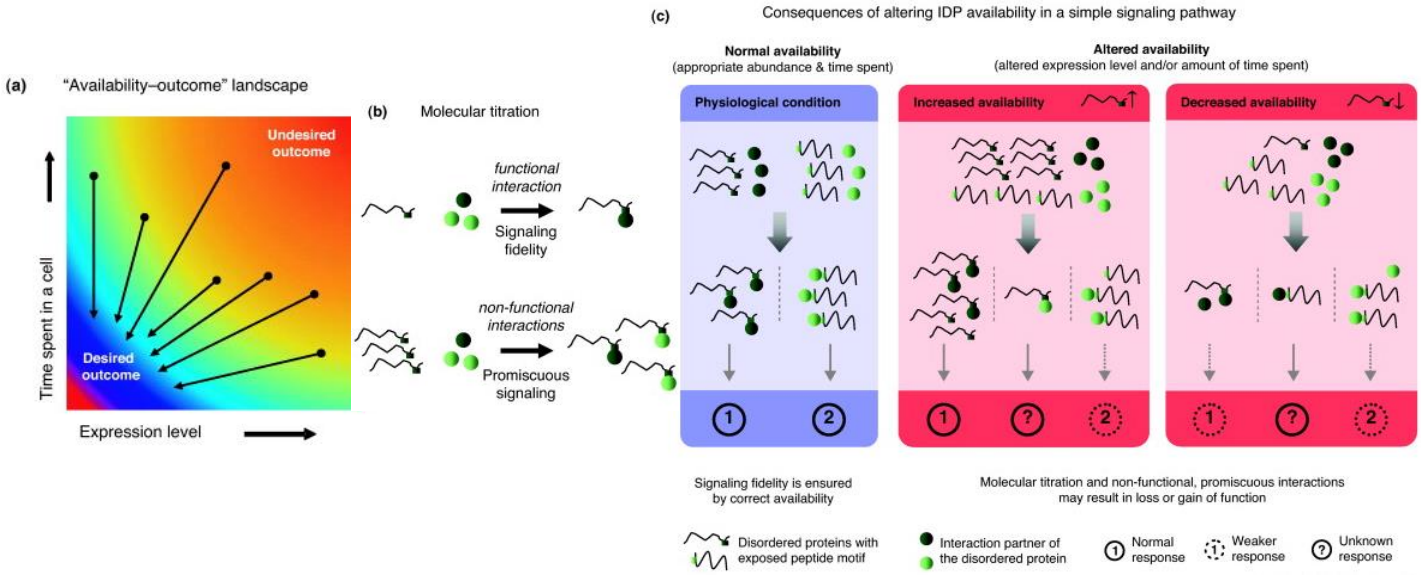
The examples of the IDPs that fall in each of these categories along with their functions, are shown in Table 1.2.

1.3.2 Regulation of IDPs and diseases

Due to their unusual structural properties and important functional properties, the existence of IDPs inside the cell needs to be carefully scrutinized. Various studies have investigated how IDPs are regulated inside a cell. Mainly, it has been investigated how cellular systems have evolved to control the availability of IDPs.[105] A comprehensive study with multiple large-scale datasets that describe the aspects of gene regulation at the transcriptional and translational level (i.e., abundance, rate of synthesis and half-life of mRNA and proteins) revealed that rather than a single regulatory mechanism, multiple mechanisms during transcription and translation control the availability of IDPs, many of which are present in low levels and for short periods. of time (Figure 1.7).[106]

In particular, it was shown that the short residence time is mainly a consequence of post-transcriptional regulation that involves increased transcript clearance and proteolytic degradation. In fact, studies on *S. pombe* and human cells indicated that the tight regulation of IDPs is possibly an evolutionary conserved strategy.[105] Further, these observations suggest that the tight control of the availability of IDPs might provide fidelity in signaling, regulation and recognition by minimizing the prospect of non-functional interactions and unwanted sequestration of proteins into non-physiological complexes.[106] Disordered regions have higher

tendency towards promiscuous interactions when their concentration is high. This is likely to be the origin of the toxicity when genes are overexpressed. For example, over-expression and



mutations in certain IDPs such as tau, ataxin-1 and α-synuclein are known to increase the possibility of aggregation and have been linked to various pathological conditions in humans.[11]

Figure 1.7. (a) The hypothetical 'availability–outcome' landscape and the regulation of IDPs. Inside a cell, a well-coordinated, tight regulation (arrows) of IDPs (black dots) at several stages during transcription and translation ensures that they are present not only at appropriate levels (x-axis) but also for the right amount of time (y-axis). (b) The tight regulation of IDPs may provide fidelity in signaling and regulation by minimizing non-functional protein-peptide interactions. Wavy lines denote disordered proteins exposing a peptide motif (small dark green box). At the right expression level and time period (top part), functional interaction is mediated with the relevant partner (dark green circle). However, at high concentrations of the IDP (bottom part), structurally similar proteins may interact and be sequestered because of non-functional interaction with the IDP, known as molecular titration. This may result in a perturbed cellular state and may cause cellular toxicity. (c) Potential consequences of altered availability of an IDP that is employed in a simple signaling pathway are shown. Reproduced with permission from [106].

From these examples, it is clear that for normal cellular function, it is important to maintain the right endogenous levels of IDPs and the integrity of the intricate regulatory mechanisms associated with IDPs. However, all IDPs may not be harmful when they are overexpressed.[107] In fact, there are many IDPs that may be required in high levels or for longer period of time in response to certain specific stimuli such as under conditions of stress or during specific phases of the cell-cycle.[105] Studies have suggested that the IDP availability, in such cases, can be fine-tuned by PTMs and by other regulatory factors that would result in the increased abundance

and longer half-life, via change in cellular localization or by protection from the degradation machinery.[108] Therefore, in order to function properly, cells must make sure that IDPs are employed in the correct pathways at the right time to ensure an appropriate response (Figure 1.7). Thus, it is clear that the availability of IDPs is tightly regulated at multiple levels.

1.3.3 Sequence-conformational ensemble relationships

The amino acid sequence composition of IDPs and IDRs is quite unique. They are rich in polar and charged amino acid residues but deficient in bulky hydrophobic residues.[79] Due to absence of canonical hydrophobic residues, these proteins lack the ability to undergo hydrophobic collapse that is believed to drive the process of protein folding. As mentioned above, IDPs exist as a dynamic ensemble of fluctuating conformations. A large number of experimental and computational studies have addressed the relationship between the information encoded in amino acid sequences and the conformational ensemble accessible to IDPs under different environments. These studies have mainly focused on three distinct compositional classes that are based on the fraction of charged versus polar residues: (i) Polar tracts, (ii) Polyelectrolytes and (iii) Polyampholytes.[109] Polar tracts are enriched in polar residues like glutamine, asparagine, serine, glycine, threonine and histidine, and are deficient in charged and hydrophobic residues. Polyampholytes and polyelectrolytes can either be weak or strong depending on the fraction of charged residues (FCR), which is sum of the fraction of positively charged residues (f_+) and fraction of negatively charged residues (f_-).[110] Polyelectrolytes are enriched in charged residues and have one type of charge present in excess, that is, either $f_+ > f_-$ or $f_- > f_+$. Polyampholytes have roughly equal number of positive and negative charges, that is, $f_+ \approx f_-$. In the case of polyelectrolyte and polyampholyte, it is finally the FCR value that determines whether a polyelectrolyte/polyampholyte is weak or strong.[110] The FCR value is high in case of strong polyampholytes/polyelectrolytes and it encodes an inherent tendency for populating expanded coil-like conformations because charged residues prefer to be solvated in aqueous milieus.[110] Both FCR and net charge per residue (NCPR: $f_+ - f_-$) are important for predicting the conformation ensemble for an IDP, especially in case of polyampholytes.[110]

It has been shown by both experiments and simulations that increasing the NCPR beyond a certain threshold can reverse the tendency of polypeptide backbones to form ensembles of collapsed conformations. The switch between globules and expanded coils is sharp which suggests that even slight changes in the NCPR of an IDP via PTMs could lead to the reversible coil-to-globule transition. The influence of the NCPR on the conformational

propensities of IDPs has been explained by a diagram-of-states[111], which generalizes the original charge-hydropathy plot (discussed above).[79] The original diagram-of-states that was proposed had three axes: f_+, f_- and hydropathy (Figure 1.8A). However, considering the fact that the hydropathy for IDPs is intrinsically low, the diagram-of-states was modified to focus only on the other two axes, that is, f_+ and f_- (Figure 1.8B).[112] This modified diagram-of-states

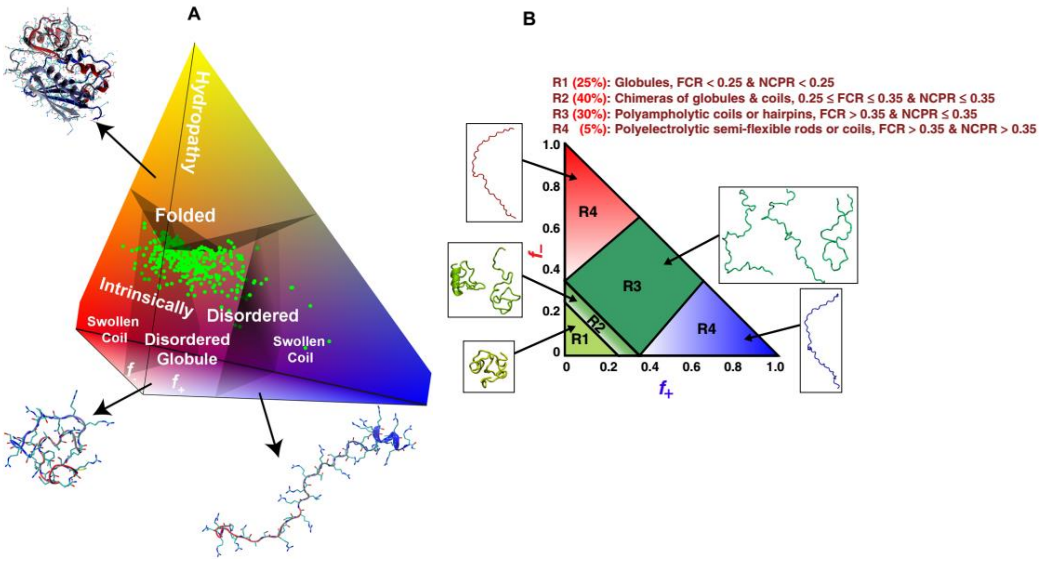


Figure 1.8. Original and modified diagram-of-states to classify predicted conformational properties of IDPs. (A) The original diagram-of-states has three axes that denote the fraction of positively charged residues, f_+, the fraction of negatively charged residues, f_-, and the hydropathy. These three parameters are calculated from the amino acid composition. Green dots in the diagram correspond to 364 curated disordered sequences extracted from the DisProt database. The hydropathy values of these sequences designate them as being disordered; that is, they lie in the bottom portion of the pyramid by definition. (B) Modified diagram-of-states from panel (A) with a focus only on the bottom portion of the pyramid (since the hydropathy is low enough to be ignored). Statistics for different regions (percentages) are from analysis of bona fide IDPs in DISPROT. Reproduced with permission from [102] and [110].

classifies IDPs into four distinct conformational classes referred to as R1, R2, R3 and R4, respectively.[110] Polar tracts and weak polyampholytes/polyelectrolytes fall in region R1 and exist as disordered globules.[110] Strong polyampholytes fall in region R3 and exist either as hairpins or coils, depending upon the combination of charge pattern and FCR values.[110] Since region R2 lies between R1 and R3, the sequence composition that falls in this region is expected to have conformational characteristics of both R1 and R3.[110] Therefore, IDPs that fall in R2 region are likely to be chimeras of globules and coils. Mostly, IDPs that undergo folding upon binding to their partners, belong to this region. Strong polyelectrolytes fall in R4 region, that

occupies two portions of the diagrams and corresponds to acid rich ($f_- > f_+$) and base rich ($f_+ > f_-$) polyelectrolytes, respectively. In this case, the electrostatic repulsions between charged residues and the favorable solvation free energies of these residues together result in semi-flexible coil-like structural ensemble. It is important to emphasize that the diagram-of-states holds valid only for IDP sequences that have at least thirty residues, low overall hydropathy, and low proline content.[110] Analysis of the disordered sequences from the DisProt database[113] suggests that at least 70% of known IDP sequences belong to regions R2 and R3. Regions R1, R2, and R3 together comprise at least 95% of the known sequences of IDPs which indicates that a vast majority of IDPs have amino acid compositions that predispose them to be globule formers. Thus, the diagram-of-states classification clearly describes the distinct conformational classes for IDP sequences and the underlying principles of sequence-conformational ensemble relationship enable the use of de novo sequence design as a tool for modulating these properties and assessing their impact on functions associated with IDPs and IDRs, that have been discussed in the above section. In this thesis, we have studied two IDPs, namely, κ-casein (Chapter 2, 3 and 5) and α-synuclein (Chapter 4) that fall in R1 and R2 region, respectively in the diagram-of-states.

1.4 Polymer physics description of IDPs and protein aggregation

Since IDPs are unfolded in their native state, they can be treated like polymers and can be described by the scaling laws of polymer physics. Additionally, the proteins that are prone to aggregation tend to either be intrinsically disordered (see Table 1.1) or go through the partially disordered state during the process of aggregation and this disorder facilitates non-specific interactions between protein molecules.[114] Therefore, the concepts of polymer physics can provide the necessary framework for understanding the driving forces and mechanism of protein aggregation. The intricate balance between chain-chain and chain-solvent interactions, and the magnitude and nature of spontaneous conformational fluctuations that promote intermolecular interactions underlies the physical driving forces for any homopolymer or heteropolymer aggregation.[115] The conformational and phase equilibria for synthetic polymers as well as IDPs are dictated by the balance of chain-solvent, chain-chain and solvent-solvent interactions.[116] Under dilute conditions, a single, linear, flexible polymer can either exist as swollen coil, favored in good solvent or as compact globule, favored in poor solvent.

According to Flory's polymer theory, the radius of gyration, given by R_g, has a strong dependence on the length of the polymer (N) and follows the polymer law given by $R_g = R_0 N^{\nu}$

(Figure 1.9). For polymer chains in a good solvent, $R_g \sim N^{2/3}$. The effective pairwise inter-residue interactions are described by a second virial coefficient B_2.[116-117] For good solvents, $B_2 > 0$ which means that the effective pairwise inter-residue interactions are net repulsive as chain-solvent interactions dominate chain-chain interactions. In case of poor solvents, $B_2 < 0$ which

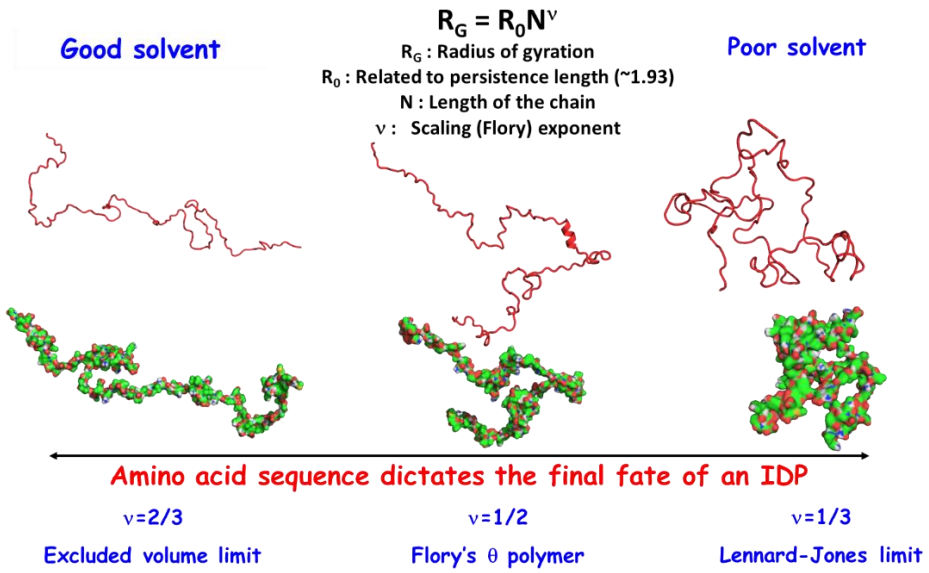


Figure 1.9. In general, for polymer chains, ν is an indicator of the solvent quantity. In case of IDPs, the conformational ensemble most accessible to an IDP is dictated by the solvent quantity as well as amino acid sequence composition.

means that the effective pairwise inter-residue interactions are net attractive as chain-chain interactions are preferred over chain-solvent interactions and $R_g \sim N^{1/3}$.[116-117] Therefore, in a good solvent, swollen coils are preferred and in a poor solvent, collapsed globules are preferred. While going from good solvent to poor solvent or vice-versa, polymers go through the theta state. At theta state, polymer is an ideal Flory's random coil and is characterized by maximum conformational entropy. For theta solvent, $R_g \sim N^{1/2}$ and $B_2 = 0$ because at theta point, the chain-chain and chain-solvent interactions cancel each other.[116-117]

While discussing the concepts underlying protein aggregation from IDPs or short peptides, it is important to realize that there are distinct differences between them and therefore, their aggregation behaviors are also very different. They are well separated in terms of length scale. This separation is considered an important hallmark in polymer studies and leads to the idea of "blobs".[115] A blob is that length scale of a polymer, beyond which the balance between

chain-chain, chain-solvent and solvent-solvent interaction is at least of the order of k_BT.[115] In a good solvent, the balance of interactions between blobs is net repulsive and in case of poor solvent, it is net attractive. In contrast, within a blob, the balance of the interactions is considerably smaller than k_BT.[115] Therefore, for a blob that has *x* residues, the radius of gyration of the blob scales as $x^{1/2}$, irrespective of the solvent quality.[116] Proteins having multiple segments of length *x* can undergo dramatic changes in both size and shape with change in the solvent quality and their conformational equilibria would influence their phase behavior as a function of protein concentration.[115] However, the conformational equilibria would only have a little influence on the phase behavior of peptides that are approximately *x* residue long.[115] Therefore, the driving forces in the aggregation of blob-sized peptides would be similar to the small molecules and very different from the aggregation of proteins with multiple blobs. The aggregation of proteins with multiple blobs is discussed below.

The quality of solvent for a protein solution can be varied by changing several thermodynamic parameters such as temperature, hydrostatic pressure, concentration of cosolutes such as osmolytes and denaturants, and pH, either one at a time or a combination of them. A parameter that is used to quantify the balance of interactions between chain and solvent as a function of multiple thermodynamic parameters mentioned above is given by χ, which refers to the Flory interaction parameter. If we consider temperature to depict solvent quality, then increase in temperature would lead towards good solvent and vice-versa.[115] However, it is important to note that phase separation and aggregation are not feasible at good and theta solvents, and require the proteins to be in poor solvent. The poor solvent condition is only attained upon decreasing the temperature ($T<T_\theta$). With decrease in temperature, under dilute conditions, the polymer collapses to minimize the interface with the surrounding solvent and forms homogeneously mixed solutions of isolated globules. As the polymer concentration is increased in a poor solvent, we enter the two-phase regime and the system can now separate into distinct solvent-rich and polymer-rich phases.[115] The driving force for phase separation/aggregation in a poor solvent can be increased further by increasing the length of polymer.[115] In case of homopolymers, a single χ value characterizes the balance of interactions between chain and solvent, and χ varies with T.[115] However, for a heteropolymer, there will be a distribution of χ values.[115] Thus, it is clear that the sequence characteristics of blobs play a critical role in governing the strengths of attractive interactions between blobs and the energy penalty associated with the unfavorable interface between a blob and the solvent. This, in turn,

suggests that some regions within the sequence would be more prone to promoting aggregation and appears to be consistent with the previous findings.[118-119]

In case of IDPs or partially unfolded proteins or the partially disordered states of the folded proteins, the conformational ensemble in a poor solvent will be heterogeneous and this conformational heterogeneity leads to the increased exposure of solvophobic blobs. As a result, the probability of making either specific or non-specific intermolecular contacts in order to reduce the exposure of solvophobic blobs is also higher. The interaction between disordered globules results in the formation of clusters. As the size of the cluster increases, pairs of blobs between chains have access to attractive intermolecular and inter-blob interactions.[115] Due to these inter-blob interactions, the driving force for restricting the chain in globular geometries is overcome and therefore, in the polymer-rich phase, each chain now unravels to sample a variety of conformations. The direct consequence of this is reflected in both theory and experiment, which showed that the scaling of radius of gyration for the chain (with chain length N) in the aggregate/polymer-rich phase takes the form $R_g \sim N^{1/2}$ as opposed to $N^{1/3}$ scaling observed for individual chains in the soluble phase.[116] Furthermore, the polymer theories, particularly the theory given by Raos and Allegra, have suggested alternative explanations for the origin of longer lag phase associated with protein aggregation.[120] According to these theories, it is the slow chain dynamics that limits the process of phase separation/aggregation and not the well-defined energy barrier associated with the formation of a structured nucleus (discussed in section 1.2.3).[121]

1.5 Water: the matrix of life

In section 1.3.3, we have discussed the role of solvent in protein aggregation. The solvent in biology is water which is very unique in itself and is considered the "matrix of life".[122] Leonardo Da Vinci described water as "the vehicle of nature" and mentioned that "Water is the extension and humor (nourishing fluid) of all living forms. Without it, nothing can maintain its form." Our body is about 70% water by weight, out of which about 46% (~2/3) is inside cells, and about 23% (~1/3) is present outside cells in blood plasma and other body fluids. Despite acknowledging the importance of water in life very long ago, biologists have ignored the central role of water in life's machinery and have treated it essentially as a mere bystander in biology, for a long time. However, it has become increasingly clear over the past two decades that water is a highly versatile and an active component of the cell that is involved in a wide range of biomolecular interactions. Thus, water performs a variety of important tasks in various

biological processes, as thoroughly reviewed by Philip Ball.[123] Let us now see how the water inside the cells is different from the bulk water (the water that we drink).

In bulk liquid water, each water molecule forms up to four directed hydrogen bonds with the neighboring water molecules in a tetrahedral manner.[124] Each of these water molecule can also establish dipole and induced dipole interactions with other molecules. This clustering of water molecules gives liquid water a heterogeneous character that, in turn, is sensitive to the physicochemical environment. Inside the cell, the behavior of water is very different from the bulk.[124-125] This is because the inside of the cell is extremely crowded due to very high concentration of macromolecules of ~400 g/L, which approximates to about 40% of the total volume of the cell.[126] Due to this crowding, macromolecules are typically separated by only 1-2 nm. Such narrow confinement is expected to alter the structure of liquid water via the disruption of its hydrogen-bond network, as a result of which, the average coordination number is reduced from about 3.6 to 2.2.[123] Molecular crowding also leads to slower diffusion rates in the mitochondrion and endoplasmic reticulum, typically 3-8 fold, compared to the bulk.[127] Further, proton NMR studies have revealed that the relaxation times of water in cells differ from that in the bulk and additionally, the relaxation has been found to be become more "bulk-like" in cancerous cells.[128-129] Despite many studies indicating the differences between bulk water and water inside the cells, determining how different the water inside the cells is compared to the bulk remains poorly understood. In fact, discerning how differently water molecules are around specific biomolecules such as DNA, RNA and proteins itself remains a challenging task. My major interest is in understanding the specific role of water in the structure and behavior of IDPs, and protein aggregation. Below, I briefly discuss the behavior of water in proteins.

1.6 Water in proteins

As mentioned above, water present close to the biomolecules is different from the bulk water as in the presence of a biomolecule, the hydrogen-bonded network of bulk water gets disrupted. These water molecules that are present in the close vicinity of biomolecules such as proteins are known as bound water or biological water.[124, 130-131] The dielectric and dynamical properties of biological water are significantly different from the bulk.[132-134] The water molecules in the bulk only make transient contact with their nearest neighbors. These transient contacts get broken very fast and neighbors get changed. In other words, the water molecules in the bulk can exchange their positions with their neighbors very fast. In contrast, the water molecules

that are present in the vicinity of a protein move in a coordinated manner and retain contact with the closest partner for a longer time. The presence of different types of water molecules leads to the heterogeneity in the water network, which is considered essential for water's role in biological functions.[124] This is why water molecules present in the proximity of hydrophobic and hydrophilic surfaces behave quite differently, and perform different functions. The layer of water molecules that surrounds a protein is known as hydration layer. In the older view, the hydration layer surrounding the protein was thought to have a rigid ice-like structure (Figure 1.10).[135] However, this "iceberg model" was rejected when many experimental and theoretical studies proposed that water indeed slows down to a certain extent at the protein surface but, it never slows down to the extent expected for an iceberg.[124, 131, 136-139] The hydration layer provides stability to the native state of a protein and is believed to influence the biological functions associated with it. Therefore, structure and dynamics of water in the hydration shell are intimately related to the functioning of proteins, and are dependent on the width of the shell. Now, the important question is how thick this hydration layer is? The thickness of a protein hydration shell has remained a subject of interest, and also a topic of debate. Depending upon the experimental probe being used, the estimated thickness of the hydration layer can vary anywhere from 1-2 monolayers to 10 monolayers.[124] The thickness of the hydration layer relates to the size of the protein itself, and is expected to be smaller for a small protein and more for a large protein. Historically, dielectric measurements provided the first ever estimate of the width of the protein hydration layer. The thickness estimated was 3–4 Å, which was believed to be quite accurate for a long time, until more advanced time-dependent studies and computer simulations became available. From computer simulations, it was found that the hydration layer of smaller proteins can extend merely upto the first layer of solvent molecules. However, in case of larger proteins, the hydration layer around the protein was found to have a thickness of more than 10Å, which accounts for more than three-monolayer thickness.[124] Recent terahertz studies estimated that the hydration layer can extend up to a distance of 15-20 Å.[124, 140] Since the biological function is intimately linked to the structure and dynamics of water in the hydration shell, it is important to understand how the structure of water is in the hydration shell and how dynamic it is.

Structure of water: The molecular arrangement of bulk water gets substantially altered in the proximity of the protein surface that contains both hydrophilic and hydrophobic residues.[123-124] The hydrophilic residues mostly consist of charged and polar atoms and hydrophobic residues contain non-polar atoms. Water can form strong hydrogen bonds with the polar or charged

surface atoms, especially with residues like glutamine and asparagine that have amide in their side chains. On the contrary, hydrophobic surface atoms lack the ability to form such hydrogen-bonded network. In addition to the nature of the amino acid residues, the heterogeneity of the charge distribution and the surface topology are anticipated to have a crucial influence on the local structure of water molecules present at the protein surface.

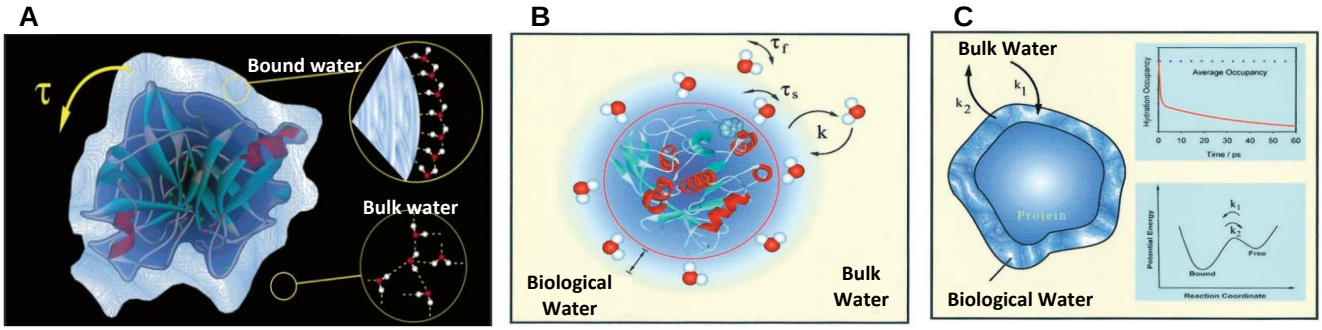


Figure 1.10. (A) Iceberg model and (B) Dynamic exchange model of hydration showing the presence of two different types of water in proteins. (C) The exchange of biological and bulk water is displayed with two rate constants, *k1* or *k2*. (Lower) The solvation behavior is correlated to a hydration occupancy, but the time scale shown is only for illustration. The free energy change is also shown. Reproduced with permission from [136] and [143].

Dynamics of water: The water dynamics in the hydration layer exhibits slower dynamics compared to the bulk, however, quantitative characterization of the dynamics has proven to be fairly difficult. The dynamic properties of the protein hydration layer have been studied by many different techniques (reviewed thoroughly in Bagchi's book)[124] that have probed properties of water on different length and time scales as a result of which, different hydration timescales have been recovered. For example, dielectric relaxation and solvation dynamics are mostly sensitive to the rotational motion of the water molecules and the protein side-chain motions,[137] while NMR based nuclear Overhauser effect (NOE) is sensitive to the relative translational motion between the protein and the water molecules.[141]. NOE is sensitive to very short length scales but is limited by the time resolution, whereas, dielectric relaxation has no sensitivity to length scales at all. Two experimental approaches that are sensitive to both length and time scales, and therefore, qualify as good probes for studying the dynamics of hydration layer are solvation dynamics and inelastic neutron scattering. I have chosen solvation dynamics approach to monitor the water reorientation dynamics in IDPs and amyloids. The techniques that have been used for the carrying out solvation dynamics studies are briefly discussed below.

1.7 Biophysical techniques used to carry out solvation dynamics studies

In order to carry out solvation dynamics studies in protein, one can either make use of an intrinsic fluorophore such as tryptophan or covalently label the protein using extrinsic fluorophores.[131, 136, 142-146] Mostly thiol-reactive fluorophores are used to label the cysteines in protein, using thiol labeling chemistry. The fluorophores that are highly sensitive to the solvent and exhibit a large Stokes' shift are preferred for solvation dynamics studies. Solvation dynamics probe the dielectric response of proteins by monitoring the time-dependent fluorescence Stokes shift (TDFSS) of the fluorophore which is either part of the protein or is covalently linked to it.[124, 131, 147] Immediately upon excitation, at t = 0, there is a creation of a new dipole around the fluorophore which gives rise to an instantaneous electric field on the solvent (water) molecules. Due to the interaction of the permanent dipoles of water molecules with the electric field, the free energy minimum of water gets shifted to a non-zero polarization and the water molecules at t = 0 find themselves in a high energy configuration (Figure 1.11).[147] This non-equilibrated hydration water network around the fluorophore now has to reorganize themselves *via* hydrogen bond rearrangement to reach their new equilibrium positions (Figure 1.11).[147] The nuclear motion that water undergoes can be broadly classified into rotational and translational motions.

In case of bulk water, rotation would also include libration (hindered rotation) and translational motion would include intermolecular vibrations due to hydrogen bonded network.[131] The librational and translational motion are particularly high in frequency and therefore, are thought to play a critical role in the initial part of solvent relaxation.[131] The quantity measured in solvation dynamics studies is solvation energy of the fluorescent probe due to interactions with the dipole moment of water molecules. Since solvation energy originates from the surrounding dipolar solvent molecules, it is considered as a collective quantity. However, this energy contribution from a particular solvent molecule is also dependent on the spatial separation *r* between the fluorescent probe and the chosen solvent molecule. This suggests that the time-dependent contribution depends on the length scale, which makes solvation dynamics a very useful technique.

In the linear response approximation, the solvation time correlation function given by S (t) is directly related to the auto time correlation function of solvation energy fluctuation given by C (t).[124] Thus,

$$S(t) = C(t) = \frac{\delta E(0)\delta E(t)}{\delta E^2} = \frac{E(t) - E(\infty)}{E(0) - E(\infty)} \qquad (1)$$

where δE(t) is the fluctuation of solvation energy from the average, equilibrium value at time t. It is important to note that the equality in above equation indicates a direct relation between the average of the fluctuations over the equilibrium distribution (left) and the non-equilibrium function (right). The initial part of the solvent relaxation is dominated by the available high-frequency modes of the solvent, that is, libration and the intermolecular vibration.[131] The long-amplitude rotation or molecular translation are not significant in the initial part. The last part comprises larger amplitude rotational as well as translational motions of solvent molecules in the first solvation shell. In the intermediate part, moderately damped rotational motions of water molecules mostly contribute.[131]

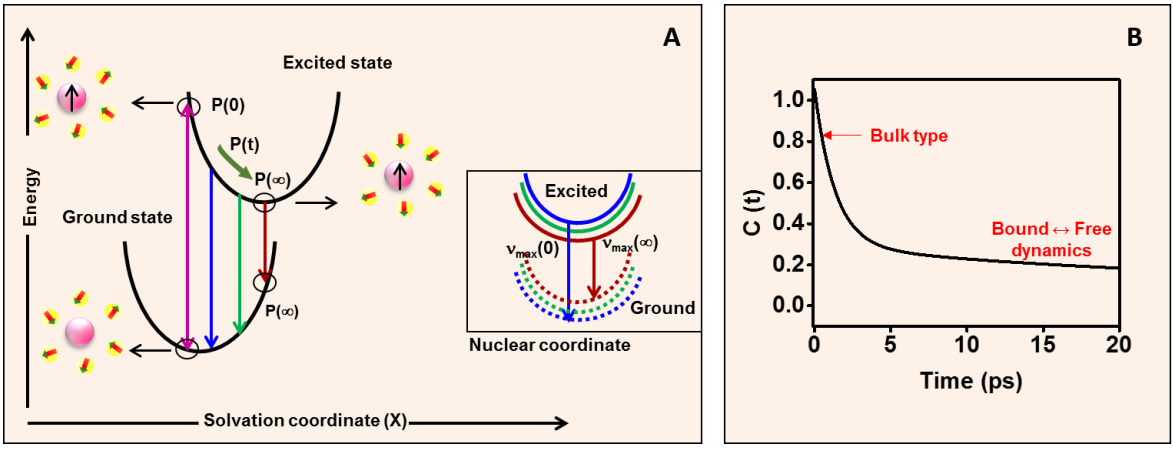


Figure 1.11. (A) Schematic illustration of the potential energy surfaces involved in solvation dynamics, showing the water orientational motions along the solvation coordinate together with instantaneous polarization P. Inset: The change in the potential energy along the intramolecular nuclear coordinate. As solvation proceeds, the energy of the solute decreases, giving rise to a red shift in the fluorescence spectrum. Note the instantaneous P, e.g., P(∞), on the two connected potentials. (B) A typical solvation time correlation function for water is shown.

The solvation dynamics studies in case of proteins is rather complex because it derives contributions not only from the surface and the bulk water molecules but also from various sources, such as amino acid side-chains and ions. When considering protein hydration, we mainly consider two types of water, those bound to the surface and those that are free. The equilibrium between these two types of water is governed by the timescale. In the hydration layer around the protein surface, the interaction with water involves hydrogen bonding to the polar and charged groups of the surface, and the strength of this bonding is dependent on the amino acid side chain composition. When the water molecules are strongly bound to the

protein, they can neither rotate nor translate and thus, cannot contribute to solvation dynamics. However, the hydrogen bonding is transient, and there is a dynamic equilibrium between free and bound water (Figure 1.10 and Figure 1.11).[136] The potential of interaction can be represented by a double-well structure to represent the processes of bond breaking and bond forming. Most of the experimental data on proteins indicated the biphasic behavior of hydration, thereby, suggesting the presence of bound and free water in the surface layer of proteins. In a simplest dynamic exchange model, the two rate constants, k_{bf} and k_{fb}, were introduced to describe the transition from bound (to the surface) to free (from the surface) and the reverse, respectively:

bound ↔ free (rotating and translating)

$(\tau_{s\ bulk}(k))^{-1}$ is close to 10^{12} s^{-1} and the transition rates, k_{bf} and k_{fb}, are expected to be smaller.[136] The residence time of the bound water molecules is essentially given by k_{bf}^{-1}, which is in the range of tens of picoseconds.

The model explained above was further modified to take into consideration the effect of re-entry of the bulk water into the layer as well as the role of orientational order and surface inhomogeneity in the observed decay of solvation correlation function.

Experimentally, the process of solvation dynamics is followed by constructing the time-resolved emission spectra (TRES). Typically, for every TRES measurement, at least 10-12 fluorescence transients are measured as a function of detection wavelength across the emission spectrum of the fluorescent probe. In femtosecond transients, a decay in the fluorescence at the blue edge of the emission spectrum and rise at the red edge is observed, a feature typical of solvent relaxation. The fluorescence transients obtained are then fitted to a function with a sum of exponentials, taking the instrument response function into account. The parameters recovered are then used to simulate the decay profiles which are then normalized using the steady-state spectrum and average lifetime of the fluorophore. The TRES so obtained is then fitted with a Lognormal or Lorentzian or Gaussian shape function to estimate the spectrum maximum, ν(t). The time-dependent Stokes' shift in the fluorophore's emission frequency (peak) within a particular time window is then a measure of the dynamics of solvation.

Finally, solvation correlation function given by C (t) that follows from equation is plotted as:

$$C(t) = [\nu(t) - \nu(\infty)] / [\nu(0) - \nu(\infty)] \qquad (2)$$

A large number of studies on the folded proteins have indicated that the hydration dynamics in proteins is extremely different from the bulk water dynamics and mainly two types (biphasic) of water have been identified at the interface: (i) water molecules that are labile, or bulk-like (~1 ps), and (ii) water molecules that are bound to the surface (~20-40 ps).[136] Broadly, the water molecules present far away from the protein surface are referred to as bulk water and the water molecules present close to the protein surface are termed as bound water or biological water.[136] The bulk water relaxation time is ~1-2 ps, whereas, the biological water relaxation time spans from 10-100 ps.[136] Some of the earlier studies have shown the presence of a long solvation component of the order of 10 ns. This solvation component can result from the highly constrained water molecules that are most likely present inside the hydrophobic core of the protein. However, such a slow timescale is likely to have a contribution from the ultraslow motion of the fluorescent probe or due to the slow conformational fluctuation of proteins.[124] In fact, in certain cases, the ps solvation can also have contribution from the local motion of the fluorophore. In order to rule out this possibility of conformational fluctuations contributing to the solvation, anisotropy measurements are often carried out.

In order to observe the ultrafast solvation (1-100 ps) in IDPs and amyloids, we have carried out solvation dynamics studies using femtosecond upconversion (see below) and to look at slower solvation occurring on the timescale of few ns, time-correlation single photon counting (TCSPC) (see below) measurements were carried out. We also carried out femtosecond and nanosecond anisotropy measurements using femtosecond upconversion and TCSPC, respectively.

1.7.1 Femtosecond upconversion

In past few years, the emergence and commercial availability of a range of advanced and stable ultrafast lasers (predominantly Ti:sapphire-based lasers with ~ 100-fs or shorter pulse widths) and the associated optoelectronic instruments have enabled the spectroscopic studies of biological and chemical systems on the subpicosecond timescale. The best way to achieve fluorescence time resolution, comparable to laser pulse width (~100 fs), is by the use of non-linear optical techniques. One of the popular non-linear techniques that involves frequency mixing is femtosecond fluorescence upconversion.[148-149] It is based on the phenomenon of sum frequency generation of light (second-order nonlinear effect) in a nonlinear crystal such as potassium dihydrogen phosphate (KDP) or β-barium borate (BBO). The illustration describing the process of frequency mixing that dictates the time resolution in fluorescence upconversion

and the upconversion setup used for studies described in this thesis is shown in Figure 1.12. The upconversion is basically a cross-correlation between the fluorescence and the probe laser pulse.[150-152] At time t=0, the sample is electronically excited by second or third harmonic of an ultrafast laser with frequency ω_p. The incoherent fluorescence from the sample (ω_f) and the probe laser pulse (ω_p) arriving at time t = τ are co-focused into a non-linear optical crystal oriented at an appropriate angle with respect to the fluorescence and laser beams, such that the

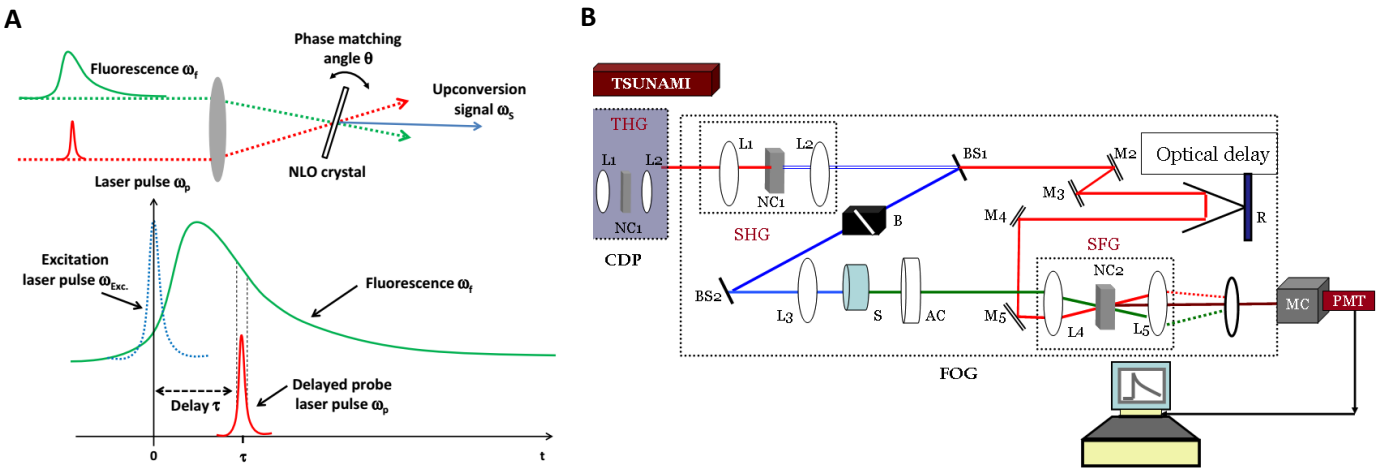


Figure 1.12. (A) Schematic diagram showing the basic principles involved in the fluorescence upconversion technique. Reproduced with permission from [152]. (B) Schematic diagram of our fluorescence upconversion setup.

phase matching condition is fulfilled.[152] The probe laser pulse is called as 'gate' pulse because the sum frequency photons are generated only during the time when this pulse is present in the crystal, thereby, acting as a 'gate'. Thus, the time resolution is within the laser pulse width. By varying the delay τ, the time evolution of fluorescence can be recovered. For solvation dynamics studies, the fluorescence upconversion transients are collected by changing the emission wavelength and TRES is plotted following the procedure discussed above. The details of the measurement parameters are discussed later in the experimental section of the individual chapters.

1.7.2 Time-correlated single photon counting (TCSPC)

Today, most of the time-domain fluorescence measurements are performed using TCSPC. These measurements can be efficiently carried out with inexpensive systems using pulsed-laser diodes (LDs), light-emitting diodes (LEDs) as excitation source, and small, fast PMTs for detection of fluorescence signal.[147] The sample is excited with a pulse of light and the

conditions are adjusted in such a manner that less than one photon is detected per laser pulse. In fact, the rate of detection is typically 1 photon per 100 excitation pulses. The time is measured between the excitation pulse and the observed photon and stored in the form of a histogram.

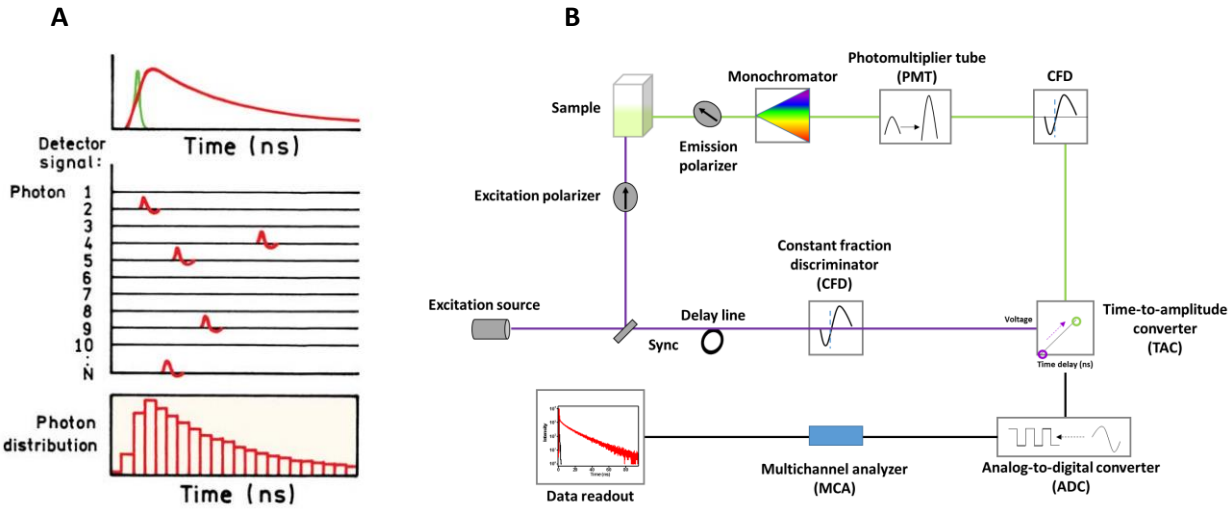


Figure 1.13. (A) Principle of TCSPC. The pulses in the middle panel represent the output from a constant fraction discriminator. Reproduced with permission from [147]. (B) Schematic diagram of our TCSPC setup.

The time difference between the excitation pulse and the arrival photon is taken as x-axis and the number of photons detected for this time difference as the y-axis. If much less than 1 photon is detected per excitation pulse, the histogram signifies the waveform of the decay. However, when the count rate is high, the histogram is biased to shorter times because with TCSPC, only the first photon can be observed (Figure 1.13). The TCSPC measurement starts when the excitation pulse hits the samples and a signal is sent to the electronics. This signal goes through a constant function discriminator (CFD), which accurately measures the arrival time of the pulse. This signal is further sent to a time-to-amplitude converter (TAC), which generates a voltage ramp, that is, a voltage that increases linearly with time on the nanosecond timescale. In other channel, the pulse from the single detected photon is received, the accurate arrival time of which is determined using a CFD that sends a signal to stop the voltage ramp. At this point, the TAC contains a voltage proportional to the time delay (Δt) between the excitation and emission signals. As per the need, the voltage is amplified by a programmable

gain amplifier (PGA) and converted to a numerical value by the analog-to-digital converter (ADC), which is stored as a single event with the measured time delay. A histogram of the decay is built by repeating this process multiple times. Typically, there are three curves associated with a fluorescence decay curve: the measured data, the instrument response function and the calculated (fitted) decay. These functions are in the form of discrete times as the photons are collected into channels, each with a known time and width. The instrument response function (IRF) is the response of the instrument to a sample with zero lifetime and is usually collected using a dilute scattering solution such as colloidal silica (Ludox). The shortest time that can be measured by the instrument can be estimated by IRF. The time resolution of a TCSPC setup is limited by the width of the IRF which is a resultant of the characteristics of the detector and the timing of the electronics. In this thesis, TCSPC measurements have been carried out to monitor the nanosecond solvation dynamics using the setup shown in Figure 1.13. The details of the measurement parameters are discussed later in the experimental section of the individual chapters.

1.8 Thesis motivation and perspective

Under certain unfortunate circumstances inside the cell, both folded proteins and IDPs can go to an alternate thermodynamically stable state known as amyloid state, which is implicated in a large number of debilitating neurodegenerative diseases, such as Parkinson's and Alzheimer's diseases. The rate of amyloid formation, the nature of the intermediate species and the lifetime of these species are believed to contribute to the pathogenesis of amyloid disease. Therefore, in vitro studies of amyloid formation are crucial for understanding the role of intermediates in the disease and for developing drugs that can intervene the process of amyloid formation. The intricate balance between chain-chain and chain-solvent (water) interactions underlies the major physical driving force for amyloid formation. An enlarging body of data have provided significant insights into the mechanism of coupled conformational switch and aggregation from the protein perspective. However, the precise mechanism of amyloid formation, the behavior of water molecules in and around IDPs, and the role of water in amyloid formation remains elusive.

In this thesis, the efforts were directed towards understanding the important role of chain-chain and chain-water interactions in amyloid formation using a wide array of biophysical techniques. The work described in this thesis provides novel insights into (i) the behavior of water molecules in and around IDPs (Chapter 2 and 4), (ii) the intriguing aspects

of water restructuring during amyloid formation (Chapter 3 and 4), and (ii) the mechanism of amyloid formation from a model amyloidogenic IDP and from a partially disordered state of a model globular protein (Chapter 5 and 6). The conformational changes (chain-chain interactions) during the amyloid formation were mainly studied using circular dichroism (CD), Raman spectroscopy and atomic force microscopy (AFM), and the behavior of water molecules in IDPs and amyloids (chain-solvent) was monitored using femtosecond- and picosecond-resolved fluorescence spectroscopy.

I believe that the work described in this thesis will improve our current understanding of amyloid formation process and provide important cues for the designing of potentially new amyloid detection tools as well as novel therapeutic approach to combat amyloid-related devastating human diseases.

1.9 References:

1. Dobson, C. M.; Šali, A.; Karplus, M. Protein Folding: A perspective from theory and experiment. *Angew. Chem. Int. Ed.* **1998,** *37*, 868-893.

2. Bartlett, A. I.; Radford, S. E. An expanding arsenal of experimental methods yields an explosion of insights into protein folding mechanisms. *Nat. Struct. Mol. Biol.* **2009,** *16*, 582-588.

3. Dunker, A. K.; Silman, I.; Uversky, V. N.; Sussman, J. L. Function and structure of inherently disordered proteins. *Curr. Opin. Struct. Biol.* **2008,** *18*, 756-764.

4. Dobson, C. M. Protein folding and misfolding. *Nature* **2003,** *426*, 884-890.

5. Wolynes, P. G.; Onuchic, J. N.; Thirumalai, D. Navigating the folding routes. *Science* **1995,** *267*, 1619-1620.

6. Dill, K. A.; Chan, H. S. From Levinthal to pathways to funnels. *Nat. Struct. Mol. Biol.* **1997,** *4*, 10-19.

7. Nussinov, R.; Udgaonkar, J. B. Editorial overview: Folding and binding: Dynamic conformational heterogeneity is pivotal to cell life. *Curr. Opin. Struct. Biol.* **2016,** *36*, iv-vi.

8. Hartl, F. U.; Hayer-Hartl, M. Converging concepts of protein folding in vitro and in vivo. *Nat. Struct. Mol. Biol.* **2009,** *16*, 574-581.

9. Jahn, T. R.; Radford, S. E. The Yin and Yang of protein folding. *FEBS J.* **2005,** *272*, 5962-5970.

10. Hartl, F. U. Molecular chaperones in cellular protein folding. *Nature* **1996,** *381* , 571-580.

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

12. Geddes, A. J.; Parker, K. D.; Atkins, E. D. T.; Beighton, E. "Cross-β" conformation in proteins. *J. Mol. Biol.* **1968,** *32*, 343-358.

13. Wetzel, R. Kinetics and thermodynamics of amyloid fibril assembly. *Acc. Chem. Res.* **2006,** *39*, 671-679.

14. Gazit, E. The "correctly folded" state of proteins: is it a metastable state? *Angew. Chem. Int. Ed.* **2002,** *41*, 257-259.

15. Baldwin, A. J.; Knowles, T. P. J.; Tartaglia, G. G.; Fitzpatrick, A. W.; Devlin, G. L.; Shammas, S. L.; Waudby, C. A.; Mossuto, M. F.; Meehan, S.; Gras, S. L.; Christodoulou, J.; Anthony-Cahill, S. J.; Barker, P. D.; Vendruscolo, M.; Dobson, C. M. Metastability of native proteins and the phenomenon of amyloid formation. *J. Am. Chem. Soc.* **2011,** *133*, 14160-14163.

16. Sunde, M.; Blake, C. The structure of amyloid fibrils by electron microscopy and X-ray diffraction. *Adv. Prot. Chem.* **1997,** *50*, 59-123.

17. Westermark, G. T.; Johnson, K. H.; Westermark, P. Staining methods for identification of amyloid in tissue. In *Methods Enzymol.*, Academic Press: **1999,** *309*, 3-25.

18. Dobson, C. M. Protein misfolding, evolution and disease. *Trends Biochem. Sci.* **1999,** *24*, 329-332.

19. Cohen, A. S. General introduction and a brief history of amyloidosis. In *Amyloidosis*, Marrink, J.; Van Rijswijk, M. H., Eds. Springer Netherlands: Dordrecht, **1986,** 3-19.

20. Sipe, J. D.; Cohen, A. S. Review: history of the amyloid fibril. *J. Struct. Biol.* **2000,** *130*, 88-98.

21. Wilks, S. *Guy's Hospital Reports.* Рипол Классик**: 1865.**

22. Kyle, R. A. Amyloidosis: a convoluted story. *Br. J. Haematol.* **2001,***114*, 529-538.

23. Cohen, A. S.; Calkins, E. Electron microscopic observations on a fibrous component in amyloid of diverse origins. *Nature* **1959,** *183*, 1202-1203.

24. Eanes, E. D.; Glenner, G. G. X-ray diffraction studies on amyloid filaments. *J. Histochem. Cytochem.* **1968,** *16*, 673-677.

25. Glenner, G. G.; Eanes, E. D.; Bladen, H. A.; Linke, R. P.; Termine, J. D. Beta-pleated sheet fibrils. A comparison of native amyloid with synthetic protein fibrils *J. Histochem. Cytochem.* **1974,** *22*, 1141-1158.

26. Astbury, W. T.; Dickinson, S.; Bailey, K. The X-ray interpretation of denaturation and the structure of the seed globulins. *Biochem. J.* **1935,** *29*, 2351-2360.

27. Glenner, G. G.; Ein, D.; Eanes, E. D.; Bladen, H. A.; Terry, W.; Page, D. L. Creation of "amyloid" fibrils from Bence Jones proteins in vitro. *Science* **1971,** *174*, 712-714.

28. Glenner, G. G.; Terry, W.; Harada, M.; Isersky, C.; Page, D. Amyloid fibril proteins: proof of homology with immunoglobulin light chains by sequence analyses. *Science* **1971,** *172*, 1150-1151.

29. Rambaran, R. N.; Serpell, L. C. Amyloid fibrils: Abnormal protein assembly. *Prion* **2008,** *2*, 112-117.

30. Prusiner, S. B. Novel proteinaceous infectious particles cause scrapie. *Science* **1982,** *216*,136-144.

31. Glenner, G. G.; Wong, C. W. Alzheimer's disease: Initial report of the purification and characterization of a novel cerebrovascular amyloid protein. *Biochem. Biophys. Res. Commun.* **1984,** *120*, 885-890.

32. Wickner, R. B. [URE3] as an altered URE2 protein: evidence for a prion analog in Saccharomyces cerevisiae. *Science* **1994,** *264*, 566-569.

33. Taylor, K. L.; Cheng, N.; Williams, R. W.; Steven, A. C.; Wickner, R. B. Prion domain initiation of amyloid formation in vitro from native Ure2p. *Science* **1999,** *283*, 1339-1343.

34. Barnhart, M. M.; Chapman, M. R. Curli biogenesis and function. *Annu. Rev. Microbiol.* **2006,** *60*, 131-147.

35. Fowler, D. M.; Koulov, A. V.; Alory-Jost, C.; Marks, M. S.; Balch, W. E.; Kelly, J. W.; Functional amyloid formation within mammalian tissue. *PLoS Biol.* **2006,** *4*, e6.

36. Lansbury, P. T. In pursuit of the molecular structure of amyloid plaque: new technology provides unexpected and critical information. *Biochemistry* **1992,** *31*, 6865-6870.

37. Blake, C.; Serpell, L. Synchrotron X-ray studies suggest that the core of the transthyretin amyloid fibril is a continuous beta-sheet helix. *Structure* **1996,** *4*, 989-998.

38. Jiménez, J. L.; Guijarro, J. I.; Orlova, E.; Zurdo, J.; Dobson, C. M.; Sunde, M.; Saibil, H. R. Cryo-electron microscopy structure of an SH3 amyloid fibril and model of the molecular packing. *EMBO J.* **1999,** *18*, 815-821.

39. Schmidt, M.; Sachse, C.; Richter, W.; Xu, C.; Fändrich, M.; Grigorieff, N. Comparison of Alzheimer Aβ(1–40) and Aβ(1–42) amyloid fibrils reveals similar protofilament structures. *Proc. Natl. Acad. Sci. U. S. A.* **2009,** *106*, 19813-19818.

40. Török, M.; Milton, S.; Kayed, R.; Wu, P.; McIntire, T.; Glabe, C. G.; Langen, R. Structural and dynamic features of Alzheimer's Abeta peptide in amyloid fibrils studied by site-directed spin labeling *J. Biol. Chem.* **2002,** *277*, 40810-40815.

41. Tycko, R. Solid State NMR Studies of Amyloid Fibril Structure. *Annu. Rev. Phys. Chem.* **2011,** *62*, 279-299.

42. Nelson, R.; Sawaya, M. R.; Balbirnie, M.; Madsen, A. O.; Riekel, C.; Grothe, R.; Eisenberg, D. Structure of the cross-beta spine of amyloid-like fibrils. *Nature* **2005,** *435*, 773-778.

43. Makin, O. S.; Atkins, E.; Sikorski, P.; Johansson, J.; Serpell, L. C. Molecular basis for amyloid fibril formation and stability. *Proc. Natl. Acad. Sci. U. S. A.* **2005,** *102*, 315-320.

44. Nelson, R.; Eisenberg, D. Recent atomic models of amyloid fibril structure. *Curr. Opin. Struct. Biol.* **2006,** *16*, 260-265.

45. Sawaya, M. R.; 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. Atomic structures of amyloid cross-beta spines reveal varied steric zippers. *Nature* **2007,** *447*, 453-457.

46. Greenwald, J.; Riek, R. Biology of amyloid: structure, function, and regulation. *Structure* **2010,** *18*, 1244-1260.

47. Wasmer, C.; Lange, A.; Van Melckebeke, H.; Siemer, A. B.; Riek, R.; Meier, B. H. Amyloid fibrils of the HET-s(218-289) prion form a beta solenoid with a triangular hydrophobic core. *Science* **2008,** *319*,1523-1526.

48. Fandrich, M.; Meinhardt, J.; Grigorieff, N. Structural polymorphism of Alzheimer Abeta and other amyloid fibrils. *Prion* **2009,** *3*, 89-93.

49. Petkova, A. T.; Leapman, R. D.; Guo, Z.; Yau, W. M.; Mattson, M. P.; Tycko, R. Self-propagating, molecular-level polymorphism in Alzheimer's beta-amyloid fibrils. *Science* **2005,** *307*, 262-265.

50. Hu, K. N.; McGlinchey, Ryan P.; Wickner, Reed B.; Tycko, R. Segmental polymorphism in a functional amyloid. *Biophys. J.* **2011,** *101*, 2242-2250.

51. Toyama, B. H.; Kelly, M. J. S.; Gross, J. D.; Weissman, J. S. The structural basis of yeast prion strain variants. *Nature* **2007,** *449*, 233-237.

52. Safar, J.; Wille, H.; Itri, V.; Groth, D.; Serban, H.; Torchia, M.; Cohen, F. E.; Prusiner, S. B. Eight prion strains have PrPSc molecules with different conformations. *Nat. Med.* **1998,** *4*, 1157-1165.

53. Jones, E. M.; Surewicz, W. K. Fibril Conformation as the basis of species- and strain-dependent seeding specificity of mammalian prion amyloids. *Cell* **2005,** *121*, 63-72.

54. Eisenberg, D.; Jucker, M. The amyloid state of proteins in human diseases. *Cell* **2012,** *148*, 1188-1203.

55. Sunde, M.; Serpell, L. C.; Bartlam, M.; Fraser, P. E.; Pepys, M. B.; Blake, C. C. F. Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *J. Mol. Biol.* **1997,** *273*, 729-739.

56. Chen, M.; Margittai, M.; Chen, J.; Langen, R. Investigation of α-synuclein fibril structure by site-directed spin labeling. *J. Biol. Chem.* **2007,** *282*, 24970-24979.

57. Iwata, K.; Fujiwara, T.; Matsuki, Y.; Akutsu, H.; Takahashi, S.; Naiki, H.; Goto, Y. 3D- structure of amyloid protofilaments of β2-microglobulin fragment probed by solid-state NMR. *Proc. Natl. Acad. Sci. U. S. A.* **2006,** *103*, 18119-18124.

58. Shewmaker, F.; McGlinchey, R. P.; Wickner, R. B. Structural insights into functional and pathological amyloid. *J. Biol. Chem.* **2011,** *286*, 16533-16540.

59. Petkova, A. T.; Buntkowsky, G.; Dyda, F.; Leapman, R. D.; Yau, W. M.; Tycko, R. Solid-state NMR reveals a ph-dependent antiparallel β-sheet registry in fibrils formed by a β-amyloid peptide. *J. Mol.Biol.* **2004,** *335*, 247-260.

60. Perutz, M. F.; Finch, J. T.; Berriman, J.; Lesk, A. Amyloid fibers are water-filled nanotubes. *Proc. Natl. Acad. Sci.* **2002,** *99*, 5591-5595.

61. Krishnan, R.; Lindquist, S. L. Structural insights into a yeast prion illuminate nucleation and strain diversity. *Nature* **2005,** *435*, 765-772.

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

63. Yamasaki, M.; Li, W.; Johnson, D. J. D.; Huntington, J. A. Crystal structure of a stable dimer reveals the molecular basis of serpin polymerization. *Nature* **2008,** *455*, 1255-1258.

64. Yamasaki, M.; Sendall, T. J.; Pearce, M. C.; Whisstock, J. C.; Huntington, J. A. Molecular basis of α1-antitrypsin deficiency revealed by the structure of a domain-swapped trimer. *EMBO reports* **2011,** *12*, 1011.

65. Knowles, T. P. J.; Vendruscolo, M.; Dobson, C. M. The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell Biol.* **2014,** *15*, 384-396.

66. Powers, E. T.; Powers, D. L. Mechanisms of protein fibril formation: nucleated polymerization with competing off-pathway aggregation. *Biophys. J.* **2008,** *94*, 379-391.

67. Buxbaum, J. N. Diseases of protein conformation: what do in vitro experiments tell us about in vivo diseases? *Trends Biochem. Sci.* **2003,** *28*, 585-592.

68. Frieden, C. Protein aggregation processes: In search of the mechanism. *Protein Sci.* **2007,** *16*, 2334-2344.

69. Serio, T. R.; Cashikar, A. G.; Kowal, A. S.; Sawicki, G. J.; Moslehi, J. J.; Serpell, L.; Arnsdorf, M. F.; Lindquist, S. L. Nucleated conformational conversion and the replication of conformational information by a prion determinant. *Science* **2000,** *289*, 1317.

70. Lee, J.; Culyba, E. K.; Powers, E. T.; Kelly, J. W. Amyloid-β forms fibrils by nucleated conformational conversion of oligomers. *Nat. Chem. Biol.* **2011,** *7*, 602-609.

71. Ferrone, F. Analysis of protein aggregation kinetics. In *Methods Enzymol.*, Academic Press: **1999**; *309*, 256-274.

72. Harper, J. D.; Lansbury, P. T. Models of amyloid seeding in alzheimer's disease and scrapie: mechanistic truths and physiological consequences of the time-dependent solubility of amyloid proteins. *Annu. Rev. Biochem.* **1997,** *66*, 385-407.

73. Wolde, P. R. t.; Frenkel, D. Enhancement of protein crystal nucleation by critical density fluctuations. *Science* **1997,** *277*, 1975.

74. Kumar, S.; Udgaonkar, J. B. Mechanisms of amyloid fibril formation by proteins. *Curr. Sci.* **2010,** *98*, 17.

75. Fischer, E. Einfluss der Configuration auf die Wirkung der Enzyme. *Berichte der deutschen chemischen Gesellschaft* **1894,** *27*, 2985-2993.

76. Kriwacki, R. W.; Hengst, L.; Tennant, L.; Reed, S. I.; Wright, P. E. Structural studies of p21Waf1/Cip1/Sdi1 in the free and Cdk2-bound state: conformational disorder mediates binding diversity. *Proc. Natl. Acad. Sci. U. S. A.* **1996,** *93*, 11504-11509.

77. Daughdrill, G. W.; Chadsey, M. S.; Karlinsey, J. E.; Hughes, K. T.; Dahlquist, F. W. The C-terminal half of the anti-sigma factor, FlgM, becomes structured when bound to its target, [sigma]28. *Nat. Struct. Mol. Biol.* **1997,** *4*, 285-291.

78. Wright, P. E.; Dyson, H. J. Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J. Mol. Biol.* **1999,** *293*, 321-331.

79. Uversky, V. N.; Gillespie, J. R.; Fink, A. L. Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins: Structure, Function, and Bioinformatics* **2000,** *41*, 415-427.

80. Dunker, A. K.; Obradovic, Z.; Romero, P.; Garner, E. C.; Brown, C. J. Intrinsic protein disorder in complete genomes. *Genome Inform. Ser. Workshop Genome Inform.* **2000,** *11*, 161-171.

81. Dunker, A. K.; Lawson, J. D.; Brown, C. J.; Williams, R. M.; Romero, P.; Oh, J. S.; Oldfield, C. J.; Campen, A. M.; Ratliff, C. M.; Hipps, K. W. Intrinsically disordered protein. *J. Mol. Graph. Model.* **2001,** *19*, 26-59.

82. Dunker, A. K.; Brown, C. J.; Lawson, J. D.; Iakoucheva, L. M.; Obradovic, Z. Intrinsic disorder and protein function. *Biochemistry* **2002,** *41*, 6573-6582.

83. Uversky, V. N. Protein folding revisited. A polypeptide chain at the folding-misfolding-nonfolding cross-roads: which way to go? *Cell. Mol. Life Sci.* **2003,** *60*, 1852-1871.

84. Tompa, P. Intrinsically unstructured proteins. *Trends Biochem. Sci.* **2002,** *27*, 527-533.

85. Ward, J. J.; Sodhi, J. S.; McGuffin, L. J.; Buxton, B. F.; Jones, D. T. Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J. Mol. Biol.* **2004,** *337*, 635-645.

86. Fink, A. L. Natively unfolded proteins. *Curr. Opin. Struct. Biol.* **2005,** *15*, 35-41.

87. Dyson, H. J.; Wright, P. E. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* **2005,** *6*, 197-208.

88. Gsponer, J.; Madan Babu, M. The rules of disorder or why disorder rules. *Prog. Biophys. Mol. Biol.* **2009,** *99*, 94-103.

89. Uversky, V. N.; Dunker, A. K. Understanding protein non-folding. *Biochim. Biophys. Acta - Proteins and Proteomics* **2010,** *1804*, 1231-1264.

90. Tompa, P. Intrinsically disordered proteins: a 10-year recap. *Trends Biochem. Sci.* **2012,** *37*, 509-516.

91. Forman-Kay, Julie D.; Mittag, T. From sequence and forces to structure, function, and evolution of intrinsically disordered proteins. *Structure* **2013,** *21*, 1492-1499.

92. Dunker, A. K.; Babu, M. M.; Barbar, E.; Blackledge, M.; Bondos, S. E.; Dosztányi, Z.; Dyson, H. J.; Forman-Kay, J.; Fuxreiter, M.; Gsponer, J.; Han, K.-H.; Jones, D. T.; Longhi, S.; Metallo, S. J.; Nishikawa, K.; Nussinov, R.; Obradovic, Z.; Pappu, R. V.; Rost, B.; Selenko, P.; Subramaniam, V.; Sussman, J. L.; Tompa, P.; Uversky, V. N. What's in a name? Why these proteins are intrinsically disordered. *Intrinsically Disordered Proteins* **2013,** *1*, e24157.

93. Burger, M. V.; Gurry, T.; Stultz, M. C. Intrinsically disordered proteins: Where computation meets experiment. *Polymers* **2014,** *6*, 2684-2719.

94. Vladimir, N. U. Intrinsic disorder-based protein interactions and their modulators. *Curr. Pharm. Des.* **2013,** *19*, 4191-4213.

95. Monod, J.; Wyman, J.; Changeux, J.-P. On the nature of allosteric transitions: A plausible model. *J. Mol. Biol.* **1965,** *12*, 88-118.

96. Koshland, D. E. Application of a theory of enzyme specificity to protein synthesis. *Proc. Natl. Acad. Sci. U. S. A.* **1958,** *44*, 98-104.

97. Kumar, S.; Showalter, S. A.; Noid, W. G. Native-based simulations of the binding interaction between RAP74 and the disordered FCP1 peptide. *J. Phys. Chem. B* **2013,** *117*, 3074-3085.

98. Shoemaker, B. A.; Portman, J. J.; Wolynes, P. G. Speeding molecular recognition by using the folding funnel: the fly-casting mechanism. *Proc. Natl. Acad. Sci. U. S. A.* **2000,** *97*.

99. Levy, Y.; Onuchic, J. N.; Wolynes, P. G. Fly-casting in protein−dna binding: frustration between protein folding and electrostatics facilitates target recognition. *J. Am. Chem. Soc.* **2007,** *129*, 738-739.

100. Huang, Y.; Liu, Z. Kinetic advantage of intrinsically disordered proteins in coupled folding-binding process: a critical assessment of the "fly-casting" mechanism. *J. Mol. Biol.* **2009,** *393*, 1143-1159.

101. Uversky, V. N. The mysterious unfoldome: structureless, underappreciated, yet vital part of any given proteome. *J. Biomed. Biotech.* **2010,** *2010*, 14.

102. van der Lee, R.; Buljan, M.; Lang, B.; Weatheritt, R. J.; Daughdrill, G. W.; Dunker, A. K.; Fuxreiter, M.; Gough, J.; Gsponer, J.; Jones, D. T.; Kim, P. M.; Kriwacki, R. W.; Oldfield, C. J.; Pappu, R. V.; Tompa, P.; Uversky, V. N.; Wright, P. E.; Babu, M. M. Classification of intrinsically disordered regions and proteins. *Chem. Rev.* **2014,** *114*, 6589-6631.

103. Babu, M. M.; Kriwacki, R. W.; Pappu, R. V. Versatility from protein disorder. *Science* **2012,** *337*, 1460.

104. Tompa, P.; Fersht, A. *Structure and Function of Intrinsically Disordered Proteins*. CRC Press: 2009.

105. Gsponer, J.; Futschik, M. E.; Teichmann, S. A.; Babu, M. M. Tight regulation of unstructured proteins: from transcript synthesis to protein degradation. *Science* **2008,** *322*, 1365.

106. Babu, M. M.; van der Lee, R.; de Groot, N. S.; Gsponer, J. Intrinsically disordered proteins: regulation and disease. *Curr. Opin. Struct. Biol.* **2011,** *21*, 432-440.

107. Chua, G.; Morris, Q. D.; Sopko, R.; Robinson, M. D.; Ryan, O.; Chan, E. T.; Frey, B. J.; Andrews, B. J.; Boone, C.; Hughes, T. R. Identifying transcription factor functions and targets by phenotypic activation. *Proc. Natl. Acad. Sci. U. S. A.* **2006,** *103*, 12045-12050.

108. Chu, I. M.; Hengst, L.; Slingerland, J. M. The Cdk inhibitor p27 in human cancer: prognostic potential and relevance to anticancer therapy. *Nat. Rev. Cancer* **2008,** *8*, 253-267.

109. Mao, Albert H.; Lyle, N.; Pappu, Rohit V. Describing sequence–ensemble relationships for intrinsically disordered proteins. *Biochem. J.* **2012,** *449*, 307.

110. Das, R. K.; Ruff, K. M.; Pappu, R. V. Relating sequence encoded information to form and function of intrinsically disordered proteins. *Curr. Opin. Struct. Biol.* **2015,** *32*, 102-112.

111. Mao, A. H.; Crick, S. L.; Vitalis, A.; Chicoine, C. L.; Pappu, R. V. Net charge per residue modulates conformational ensembles of intrinsically disordered proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2010,** *107*, 8183-8188.

112. Das, R. K.; Pappu, R. V. Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proc. Natl. Acad. Sci. U. S. A.* **2013,** *110*, 13392-13397.

113. Sickmeier, M.; Hamilton, J. A.; LeGall, T.; Vacic, V.; Cortese, M. S.; Tantos, A.; Szabo, B.; Tompa, P.; Chen, J.; Uversky, V. N.; Obradovic, Z.; Dunker, A. K. DisProt: the Database of Disordered Proteins. *Nucleic Acids Res.* **2007,** *35* (Database issue), D786-D793.

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

115. Pappu, R. V.; Wang, X.; Vitalis, A.; Crick, S. L. A polymer physics perspective on driving forces and mechanisms for protein aggregation. *Arch. Biochem. Biophys.* **2008,** *469*, 132-141.

116. Rubinstein, M.; Colby, R. H. *Polymer Physics*. OUP Oxford: 2003.

117. Flory, P. J. *Principles of Polymer Chemistry*. Cornell University Press: 1953.

118. Tartaglia, G. G.; Cavalli, A.; Pellarin, R.; Caflisch, A. Prediction of aggregation rate and aggregation-prone segments in polypeptide sequences. *Protein Sci.* **2005,** *14*, 2723-2734.

119. Pawar, A. P.; DuBay, K. F.; Zurdo, J.; Chiti, F.; Vendruscolo, M.; Dobson, C. M. Prediction of "aggregation-prone" and "aggregation-susceptible" regions in proteins associated with neurodegenerative diseases. *J. Mol. Biol.* **2005,** *350*, 379-392.

120. Raos, G.; Allegra, G. Macromolecular clusters in poor-solvent polymer solutions. *J. Chem. Phys.* **1997,** *107*, 6479-6490.

121. Chuang, J.; Grosberg, A. Y.; Tanaka, T. Topological repulsion between polymer globules. *J. Chem. Phys.* **2000,** *112*, 6434-6442.

122. Szent-Gyorgyi, A. Welcoming address. In *Cell-Associated Water*, Academic Press: 1979; pp 1-2.

123. Ball, P. Water as an active constituent in cell biology. *Chem. Rev.* **2008,***108*, 74-108.

124. Bagchi, B. *Water in Biological and Chemical Processes: From Structure and Dynamics to Function*. Cambridge University Press: Cambridge, 2013.

125. Thompson, H.; Soper, A. K.; Ricci, M. A.; Bruni, F.; Skipper, N. T. The three-dimensional structure of water confined in nanoporous vycor glass. *J. Phys. Chem. B* **2007,** *111*, 5610-5620.

126. Ellis, R. J.; Minton, A. P. Cell biology: Join the crowd. *Nature* **2003,** *425*, 27-28.

127. Tsukahara, T.; Hibara, A.; Ikeda, Y.; Kitamori, T. NMR study of water molecules confined in extended nanospaces. *Angew. Chem. Int. Ed.* **2007,** *46*, 1180-1183.

128. Damadian, R. Tumor Detection by Nuclear Magnetic Resonance. *Science* **1971,** *171*, 1151.

129. Fung, B. M.; Wassil, D. A.; Durham, D. L.; Chesnut, R. W.; Durham, N. N.; Berlin, K. D. Water in normal muscle and muscle with a tumor. *Biochim. Biophys. Acta - General Subjects* **1975,** *385*, 180-187.

130. Bhattacharyya, K.; Bagchi, B. Slow dynamics of constrained water in complex geometries. *J. Phys. Chem. A* **2000,** *104*, 10603-10613.

131. Pal, S. K.; Peon, J.; Bagchi, B.; Zewail, A. H. Biological water: femtosecond dynamics of macromolecular hydration. *J. Phys. Chem. B* **2002,** *106*, 12376-12395.

132. Frauenfelder, H.; Chen, G.; Berendzen, J.; Fenimore, P. W.; Jansson, H.; McMahon, B. H.; Stroe, I. R.; Swenson, J.; Young, R. D. A unified model of protein dynamics. *Proc. Natl. Acad. Sci. U. S. A.* **2009,** *106*, 5129-5134.

133. Pethig, R. Protein-water interactions determined by dielectric methods. *Annu. Rev. Phys. Chem.* **1992,** *43*,177-205.

134. Nandi, N.; Bagchi, B. Dielectric relaxation of biological water. *J. Phys. Chem. B* **1997,** *101*, 10954-10961.

135. Kauzmann, W. Some factors in the interpretation of protein denaturation. In *Adv. Protein Chem.*, C.B. Anfinsen, M. L. A. K. B.; John, T. E., Eds. Academic Press: **1959**; *14*, 1-63.

136. Pal, S. K.; Zewail, A. H. Dynamics of water in biological recognition. *Chem. Rev.* **2004,** *104*, 2099-2124.

137. Nandi, N.; Bhattacharyya, K.; Bagchi, B. Dielectric relaxation and solvation dynamics of water in complex chemical and biological systems. *Chem. Rev.* **2000,** *100*, 2013-2046.

138. Bagchi, B. Water dynamics in the hydration layer around proteins and micelles. *Chem. Rev.* **2005,** *105*, 3197-3219.

139. Halle, B. Protein hydration dynamics in solution: a critical survey. *Philos. Trans. R. Soc. Lond. B: Biol. Sci.* **2004,** *359*, 1207.

140. Ebbinghaus, S.; Kim, S. J.; Heyden, M.; Yu, X.; Heugen, U.; Gruebele, M.; Leitner, D. M.; Havenith, M. An extended dynamical hydration shell around proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2007,** *104*, 20749-20752.

141. Otting, G.; Liepinsh, E.; Halle, B.; Frey, U. NMR identification of hydrophobic cavities with ow water occupancies in protein structures using small gas molecules. *Nat. Struct. Mol. Biol.* **1997,** *4*, 396-404.

142. Zhong, D.; Pal, S. K.; Zewail, A. H. Biological water: A critique. *Chem. Phys. Lett.* **2011,** *503*, 1-11.

143. Pal, S. K.; Peon, J.; Zewail, A. H. Biological water at the protein surface: dynamical solvation probed directly with femtosecond resolution. *Proc. Natl. Acad. Sci. U. S. A.* **2002,** *99*, 1763-8.

144. Kamal, J. K.; Zhao, L.; Zewail, A. H. Ultrafast hydration dynamics in protein unfolding: human serum albumin. *Proc. Natl. Acad. Sci. U. S. A.* **2004,** *101*, 13411-6.

145. Qin, Y..; Wang, L.; Zhong, D. P. Dynamics and mechanism of ultrafast water–protein interactions. *Proc. Natl. Acad. Sci. U. S. A.* **2016,** *113*, 8424–8429.

146. Arya, S.; Mukhopadhyay, S. Ordered water within the collapsed globules of an amyloidogenic intrinsically disordered protein. *J. Phys. Chem. B* **2014,** *118*, 9191-8.

147. Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*. Springer US: 2007.

148. Hallidy, L. A.; Topp, M. R. Picosecond luminescence detection using type-II phase-matched frequency conversion. *Chem. Phys. Lett.* **1977,** *46*, 8-14.

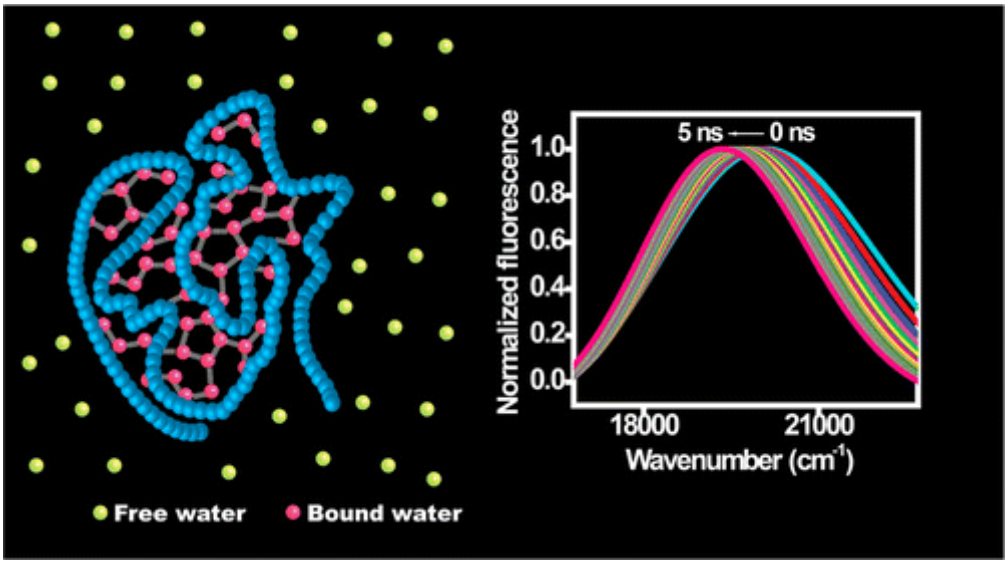
149. Mahr, H.; Hirsch, M. D. An optical up-conversion light gate with picosecond resolution. *Optics Commun.* **1975,** *13*, 96-99.

150. Shen, Y. R. *The Principles of Nonlinear Optics*. Wiley: 1984.

151. Xu, J.; Knutson, J. R. Ultrafast fluorescence spectroscopy via upconversion: applications to biophysics. In *Methods Enzymol.*, Academic Press: **2008**; *450*, 159-183.

152. Chosrowjan, H.; Taniguchi, S.; Tanaka, F. Ultrafast fluorescence upconversion technique and its applications to proteins. *FEBS J.* **2015,** *282*, 3003-3015.

Ordered Water within the Collapsed Globules of an Amyloidogenic Intrinsically Disordered Protein



The work described in this chapter has been published in the *Journal of Physical Chemistry*.

Reference: Arya, S. and Mukhopadhyay, S. *J. Phys. Chem. B* **2014,** *118* (31), 9191-8.

2.1 Introduction

The traditional sequence-structure-function paradigm cannot account for the existence of a large number of proteins that lack a well-defined 3D structure under physiological conditions. Such proteins possessing high net charge and low hydrophobicity belong to a distinct class of proteins commonly known as natively unfolded or intrinsically disordered proteins (IDPs).[1-11] Unlike the folded (globular) proteins, IDPs lack the ability to undergo autonomous folding[1] and usually exist as dynamic ensembles having rapidly fluctuating backbone Ramachandran dihedral angles.[4] Many IDPs upon binding to their respective biological partners either undergo disorder-to-order transition or form "fuzzy complexes",[6,7] thereby maintaining their chameleon state to perform various specialized functions like signaling, molecular recognition, regulation of transcription, and translation, etc.[9-11] However, the astonishing conformational plasticity of IDPs can also be detrimental because it allows the polypeptide chain to adopt aberrant toxic amyloidogenic conformations that are associated with a range of neurodegenerative disorders like Alzheimer's, Parkinson's, and Huntington's diseases.[12,13] Many theoretical and experimental approaches have been successfully applied to gain insight into the complex conformational behavior of IDPs.[14–27] These studies have revealed that most IDPs have a tendency to form collapsed globules in water,[14,17,19] especially the IDPs that are rich in polar, uncharged amino acid residues albeit devoid of hydrophobic amino acid residues. These IDPs include polyglutamine (polyQ),[19–21] N-terminal domain of the yeast prion protein Sup35,[14] and glycine-serine block copolymers.[22,27] An important factor that can control the intrinsic preference of the polypeptide chain to undergo collapse is net charge per residue.[15] Many IDP sequences have been found to have a high net charge per residue,[23] and it has been postulated that an increase in net charge per residue makes the globule-to-coil transition more favorable.[15,24,25] Furthermore, it has been observed that the polypeptide chain devoid of either nonpolar or polar side chains, such as polyglycine, also prefers an ensemble of collapsed structures in aqueous milieus, suggesting that the collapse of IDPs might originate, at least partially, from the conformational preferences of amide backbones in water.[22] Thus, water is believed to act as a poor solvent for polypeptide chains and favors the intrachain interactions over the chain-solvent interactions leading to the formation of conformationally labile collapsed globules that are not capable of undergoing autonomous folding.[15,26] The collapsed globule is at the crossroad of several biologically significant processes and can either undergo binding-induced folding in the presence of a binding partner or associate to form oligomers that can eventually mature into ordered amyloid fibrils.[26] Therefore, to understand how the

intricate balance between the extended, collapsed, folded, and misfolded conformations of the IDPs is maintained, it is important to assess the conformational propensities of IDPs in the presence of different solutes that can modulate the conformational preferences of the polypeptide chain in water. A novel approach for elucidating the conformational attributes of IDPs would be to monitor the distribution and dynamics of water molecules present in the vicinity of the polypeptide chain. Bovine κ-casein is an amyloidogenic isoform of caseins that belongs to the class of functional IDPs known as scavengers.[11] Aggregation of κ-casein results in the formation of amyloid fibrils that get deposited in the mammary glands.[28-31] We have previously shown that κ-casein under native monomeric condition undergoes a chain collapse in water to form unordered globules.[17] In this work, we characterize the dynamics of water molecules entrapped within the globules. These water molecules exhibit profoundly restrained relaxation dynamics indicating the formation of highly structured water cluster resembling nanoconfined water.

2.2 Experimental Section

2.2.1 Materials

κ-casein (from bovine milk), dithiothreitol (DTT), iodoacetic acid, guanidinium chloride (GdmCl), Tris HCl, sodium chloride (NaCl), sodium dodecyl sulfate (SDS) and sodium hydrogen phosphate (monobasic) were procured from Sigma (St. Louis, MO) and used as received. 6-Acryloyl-2- (dimethylamino)naphthalene (acrylodan) was obtained from Molecular Probes, Invitrogen Inc. All solutions were prepared in Milli-Q water. The pH of the buffers used was adjusted using a Metrohm pH meter at ~25 °C.

2.2.2 Circular dichroism (CD) measurements

The CD spectra were recorded in Chirascan Spectrophotometer (Applied Photophysics, UK) using a 1 mm path length quartz cell, and the spectra so obtained were corrected for buffer background signal. The buffer corrected spectra were then smoothened using Pro Data software.

2.2.3 Steady-state fluorescence measurements

Bovine κ-casein was reduced and carboxymethylated (RCM) using DTT and iodoacetic acid as described previously.[17] RCM κ-casein was stored under denatured condition (6 M GdmCl in pH 7, 50 mM phosphate buffer). For Trp steady-state fluorescence and CD measurements, the protein stock was diluted into the native buffer (pH 7.2, 50 mM phosphate buffer) to obtain

a final concentration of 20 μM. For studying the effect of detergent, a fresh stock of SDS was prepared in Milli-Q water. For acrylodan fluorescence measurements, reduced and denatured κ-casein was labeled with a 10 M excess of acrylodan under denatured condition (6 M GdmCl in pH 7.6, 100 mM Tris buffer, 2 h at 37 °C). The labeled protein was then passed through a PD-10 column to remove excess dye and was further concentrated using AMICON ultra (3 kDa cutoff; from Millipore). The concentration of the labeled protein was estimated using ϵ_{365} = 12800 M^{-1} cm^{-1} for acrylodan.[32] The acrylodan-labeled κ-casein was stored under denatured condition (6 M GdmCl in pH 7, 50 mM phosphate buffer). For steady-state fluorescence experiments, acrylodan-labeled protein was diluted into the native buffer (pH 7.2, 50 mM phosphate buffer) and denaturant (6 M GdmCl in pH 7, 50 mM phosphate buffer) to obtain a final concentration of 10 μM.

All the steady-state fluorescence measurements were made on Fluoromax-4 (Horiba Jobin Yvon, NJ). The samples were excited at 295 nm (Trp) and 375 nm (acrylodan). The steady-state fluorescence anisotropies were measured at 350 nm (Trp) and 500 nm (acrylodan). The steady-state fluorescence anisotropy (r_{ss}) is given by the following relationship:

$$r_{ss} = \frac{I_{\parallel} - I_{\perp} G}{I_{\parallel} + 2 I_{\perp} G} \quad (1)$$

where $I_{\parallel}$ and $I_{\perp}$ are fluorescence intensities collected using parallel and perpendicular geometry, respectively. The perpendicular components were always corrected using a G-factor. For the red-edge excitation shift (REES) experiments, the excitation wavelength for tryptophan was varied from 280 to 305 nm.

2.2.4 Stopped-flow fluorescence measurements

The fluorescence kinetics data were collected using a stopped-flow apparatus (Applied Photophysics, UK). The deadtime of mixing was ~5 ms. A 320LP (Trp) optical filter was used for collecting the total fluorescence. The mixing ratio was 10:1. The final concentrations of κ-casein and SDS were 10 and 100 μM, respectively. Native κ-casein was mixed with phosphate buffer (pH 7.2, 50 mM) to obtain the baseline signal. The stopped-flow data were acquired for 0.5 s with 10,000 samples per point. The kinetic traces were collected 10 times in triplicate.

2.2.5 Time-resolved fluorescence measurements

All time-resolved measurements were made using a time-correlated single photon counting (TCSPC) setup (Fluorocube, Horiba Jobin Yvon, NJ). The position of excitation and emission

polarizers was fixed at 54.7° (magic angle) and the peak count at 10,000 for fluorescence lifetime measurements. A 375 nm laser diode was used as excitation source for acrylodan. Ludox (colloidal silica) solution was used for collecting the instrument response function (IRF). The IRF had a full width at half maxima of ~265 ps. For time-resolved fluorescence depolarization measurements, the peak difference was set as 20,000 counts and the orientation of emission polarizer was toggled between 0° and 90° with respect to the fixed orientation of excitation polarizer for parallel ($I_{\parallel}$) and perpendicular fluorescence intensities ($I_{\perp}$), respectively. The emission wavelength was fixed at 500 nm with a bandpass of 6 nm. The perpendicular fluorescence intensities were corrected for G-factor. The anisotropy decays were analyzed using DAS6 software by globally fitting r(t) using the following relationship:

$$r(t) = \frac{I_{\parallel}(t) - GI_{\perp}(t)}{I_{\parallel}(t) + 2GI_{\perp}(t)} \quad (2)$$

The anisotropy decays were analyzed by biexponential decay function that describes the fast and slow rotational motion.[17,33,34]

$$r(t) = r_0[\beta_1 \exp\left(-\frac{t}{\Phi_1}\right) + \beta_2 \exp\left(-\frac{t}{\Phi_2}\right)] \quad (3)$$

where r_0 is the intrinsic fluorescence anisotropy, ϕ_1 and ϕ_2 are the short and long rotational correlation times and β_1 and β_2 are the amplitudes associated with short and long rotational correlation times.

The concentration of acrylodan-labeled κ-casein was 10 μM, and the concentrations of SDS used were 100, 250, and 500 μM. To construct the time-resolved emission spectra (TRES), the fluorescence decays were collected at different emission wavelengths from 440 to 600 nm after every 10 nm using an emission bandpass of 6 nm. The decays obtained at different emission wavelengths were then analyzed and fitted taking into account the IRF using DAS6 software provided with the TCSPC system. The fitted parameters were then used to simulate the fluorescence decays at different emission wavelength in Origin 8.5 software using $I_{\lambda}(t) = \Sigma\alpha_i e^{-t/\tau_i}$, where α_i and τ_i represent the contributions and lifetime of the different lifetime components, respectively. The time-resolved emission spectra (TRES) were constructed using

$$I(\lambda, t) = \frac{I_{\lambda}^{SS} I_{\lambda}(t)}{\Sigma\alpha_i\tau_i} \quad (4)$$

where I_{λ}^{SS} represents the steady-state fluorescence intensity at a fixed emission wavelength λ and $I_{\lambda}(t)$ as mentioned above is given by $I_{\lambda}(t) = \Sigma\alpha_i e^{-t/\tau_i}$. The spectra obtained

at different time points were then normalized and fitted using log-normal function to extract the peak frequency $\nu(t)$ as a function of time. The solvation correlation function is given by

$$C(t) = \frac{\nu(t)-\nu(\infty)}{\nu(0)-\nu(\infty)} \quad (5)$$

where $\nu(0)$, $\nu(t)$, and $\nu(\infty)$ are the emission peak frequencies at time 0, t, and ∞, respectively. The correlation function was then fitted using a sum of exponentials as the following:

$$C(t) = \Sigma \alpha_i e^{-t/\tau_{si}} \quad (6)$$

where α_i and τ_{si} represents the contributions and solvation time of the different solvation components, respectively.

2.3 Results and Discussion

2.3.1 Water accessibility within the compact disordered globules

Our previous bioinformatics analysis of bovine κ-casein has indicated that the amino acid sequence describes an intrinsically disordered state of the protein.[17] A closer look at the amino acid sequence of κ-casein revealed that the protein comprises at least two distinct segments having opposite net charge (Figure 2.1). The positively charged Nterminal segment comprising residues 1-100 contains fairly high occurrence of glutamine (Q) and asparagine (N) that contain amide side-chain (Q + N = 18). This segment also contains 9 tyrosines (Y), 6 lysines (K), and 5 arginines (R). The Q/N-rich segment of κ-casein is reminiscent of amyloidogenic segment of yeast prion proteins.[35,36] Therefore, this N-segment of κ-casein could potentially act as a prion-like amyloidogenic domain albeit the occurrence of Q/N/Y is much lower compared to that in yeast prion proteins. On the contrary, the C-terminal segment (101−169) is largely negatively charged and has a lower content of Q/N. The N-terminal segment adopts a collapsed premolten globule like structure under the physiological condition.[17] We hypothesized that this N-terminal segment will undergo further collapse in the presence of anionic detergent molecules. Upon binding to negatively charged detergent molecules, the charge repulsion in the N-terminal segment arising from K and R will be diminished and therefore the polypeptide chain is likely to undergo a transition from a labile collapsed globule to a more compact globule. We also conjectured that as a result of compaction of the globule, the water molecules within the globule will become more ordered.

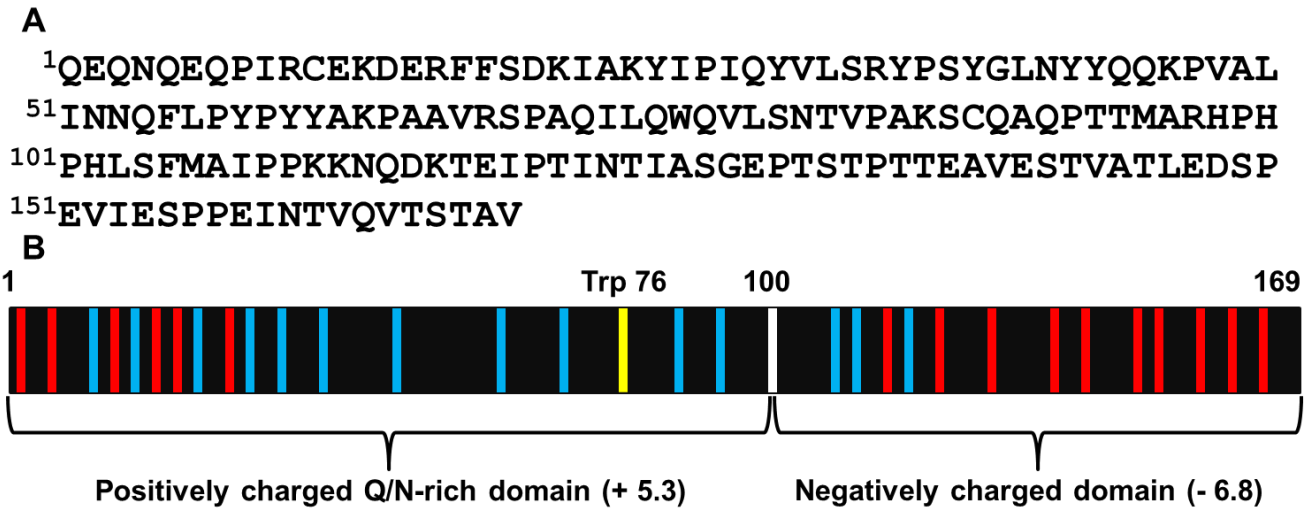


Figure 2.1 (A) Amino acid sequence of κ-casein. (B) Polypeptide divided into two distinct domains: positively charged Q/N-rich N-terminal segment (1−100) and negatively charged C-terminal segment (101−169). Red and blue lines represent negatively and positively charged residues, respectively. Trp 76 is shown by a yellow line.

First, to monitor the secondary structural changes, we carried out circular dichroism (CD) experiments. The CD spectrum of κ-casein under the physiological condition represents a disordered state.[17] Next we monitored the CD spectrum in the presence of an anionic detergent, SDS, at submicellar concentration. Although a little increase in the ellipticity at 222 nm was observed upon addition of 100 μM SDS (Figure 2.2A), the polypeptide did not undergo any large scale structural rearrangement and retained the premolten globule-like state as assessed by the CD double wavelength plot analysis [$\theta_{222}/\theta_{200}$].[10] Next, to further structurally characterize the SDS-induced state, we recorded the fluorescence of the single tryptophan (Trp 76) that is present in the Q/N-rich positively charged domain. In the absence of SDS, the Trp emission spectrum exhibited a maximum ~342 nm, suggesting that Trp experiences partial protection from the bulk water (Figure 2.2B). Upon addition of SDS, the fluorescence intensity increased with a concomitant blue shift to ~335 nm, suggesting further compaction of the unstructured globule in the presence of SDS (Figure 2.2B). The stopped-flow fluorescence experiments revealed that the compaction occurs on the 100 ms time scale with some amplitude of milliseconds and unresolved submillisecond kinetic phases (Figure 2.2B inset). We point out that this compaction of the disordered polypeptide chain does not cause Trp to experience

a significant protection from water because the emission maxima do not exhibit a substantial blue shift (< 330 nm). Therefore, the compact globules in both the absence and the presence of SDS represent wet globules containing water molecules within the collapsed yet disordered polypeptide chain. This interpretation is further corroborated by the acrylamide quenching experiments that demonstrated only a marginal protection of Trp from water in the compact disordered state (bimolecular quenching constant, $k_q \sim 4 \times 10^9$ M^{-1} s^{-1}). Taken together, this set of results reveals that unordered yet collapsed globule of κ-casein undergoes a further compaction in the presence an anionic detergent that diminishes charge repulsion within the polypeptide chain in a poor solvent such as water. Our Trp fluorescence readouts provide a support in favor of the presence of inner aqueous environment within the collapsed globules. We next performed experiments aiming at characterizing the nature of water molecules within the globules.

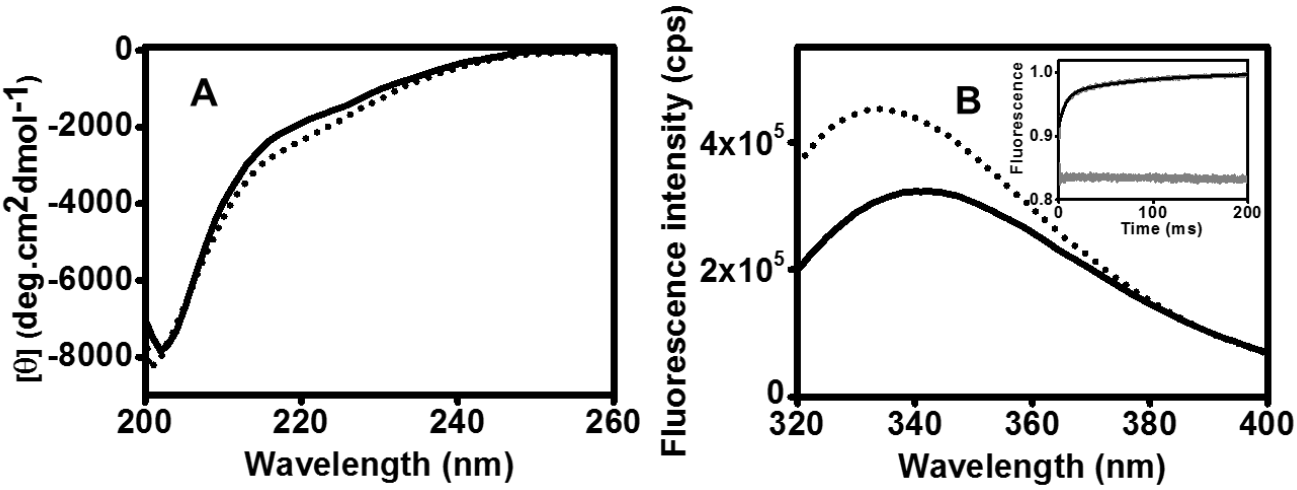


Figure 2.2 (A) CD and (B) Trp fluorescence spectra (λ_{ex} 295 nm) of κ-casein in the absence of SDS (solid line) and in the presence of 100 μM SDS (dotted line). The inset in (B) shows the stopped-flow kinetics of Trp fluorescence to monitor the structural changes from the native to SDS-induced state.

2.3.2 Slow dipolar relaxation of Trp observed by red-edge excitation shift

To delineate the characteristics of entrapped water molecules, we took advantage of a fluorescence readout, namely the red-edge excitation shift (REES) that represents a unique and sensitive approach to monitor the dynamics of restricted water molecules.[37-41] In bulk (nonviscous) media, the fluorescence emission peak maxima is independent of the excitation wavelength because the time scale of water reorientation (solvation dynamics in response to a transiently created excitation dipole) is orders of magnitude faster (picoseconds) compared to the fluorescence emission time scale (nanoseconds). In a highly ordered microenvironment, the water reorientation time gets impeded and often compares with the time scale of fluorescence emission (fluorescence lifetime). Under such conditions, the emission maxima exhibit a progressive red shift when the excitation

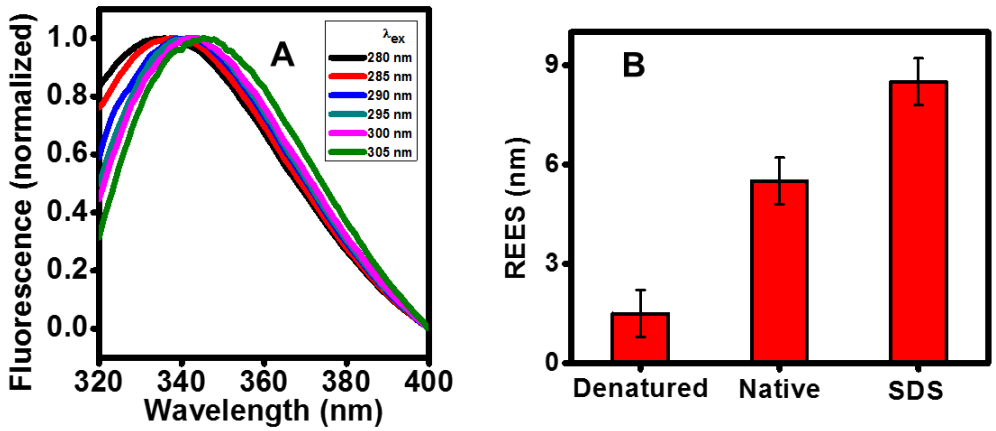


Figure 2.3 Red-edge excitation shift (REES) of Trp 76 in κ-casein. (A) The normalized fluorescence spectra obtained with different excitation wavelength from 280 to 305 nm. (B) The extent of REES (in nm) obtained using λ_{ex} from 290 to 305 nm plotted for different conformational states of the protein. Data are shown as mean ± SEM (n = 3).

wavelength is shifted to the red edge of the absorption spectrum.[37] Because a variety of spectral signatures suggested that the single tryptophan (Trp 76) of κ-casein resides in the wet globule, we hypothesized that the REES measurements on Trp 76 will report the dynamics of otherwise optically silent water molecules present within the globule. Under native condition in the absence of salt, Trp demonstrated ~6 nm REES upon changing the excitation wavelength from 290 to 305 nm (Figure 2.3A,B). In the presence of SDS, the extent of REES increased to ~9 nm (Figure 2.3B), suggesting that the water molecules become progressively ordered as the chain undergoes further compaction. On the contrary, the GdmCl denatured state showed a

much lower extent of REES (Figure 2.3B). Although the REES measurement is a reliable and sensitive indicator of slow solvent relaxation, it does not allow us to estimate the time scale of water relaxation. Therefore, to directly monitor the water dynamics, we next embarked on the study of the solvation dynamics using time-resolved emission spectra (TRES) that has previously been successfully employed in understanding the organization and dynamics around a fluorophore in complex chemical and biological systems.[42-52]

2.3.3 Dynamics of water within the globules

To measure the dynamics of water molecules within the unstructured globules, we covalently attached an environmentally sensitive dye, such as acrylodan, which reacts with the free thiol groups of Cys present in the protein. Acrylodan was chosen among the other thiol-active fluorescent dyes because it is a small environment-sensitive fluorophore, creates a highly stable thioether bond with Cys, demonstrates high fluorescence even at low (micromolar) protein concentrations and has been previously used for the solvation dynamics studies of proteins.[51,52]

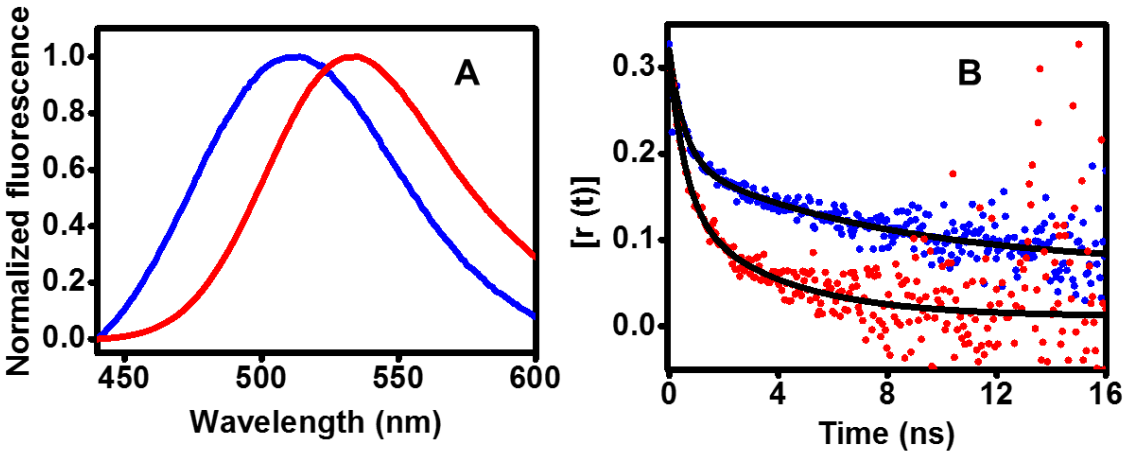


Figure 2.4 (A) Steady-state fluorescence spectra and (B) time-resolved fluorescence anisotropy decays of acrylodan-labeled κ-casein under denatured condition (red line) and native condition (blue line). Anisotropy fits have been represented with a black line. The recovered rotational correlation times are ϕ_1 = 0.4 ± 0.2 ns, ϕ_2 = 2.7 ± 1.0 ns (denatured) and ϕ_1 = 0.4 ± 0.2 ns, ϕ_2 = 7.2 ± 1.0 ns (native). The intrinsic time-zero anisotropy of acrylodan was 0.30.

As a prelude to our solvation dynamics experiments, we first recorded the fluorescence spectrum of acrylodan-labeled κ-casein under denatured and native conditions. The emission peak shifted from ~530 to ~510 nm upon changing the condition from denatured to native suggesting partial burial of the fluorophore in the collapsed state under the native condition

(Figure 2.4A). However, the blue shift was much lower compared to that seen when acrylodan is well protected from the aqueous environment (≤480 nm).[53] Additionally, the time-resolved fluorescence depolarization measurements indicated considerable local flexibility of the fluorophore along with the global tumbling of the globules with hydrodynamic radii of ~2 nm (Figure 2.4B).[17] These findings corroborate the intrinsic Trp fluorescence results, indicating that the polypeptide chain adopts a wet-globule state under the native condition.

We next embarked upon TRES experiments to estimate the time scale of water reorientation around the fluorophore within the globules under native condition by monitoring the time-dependent Stokes shift (TDSS). Figure 2.5A shows the time- resolved decays of acrylodan monitored at different emission wavelengths ranging from 440 to 600 nm. The fluorescence decays become progressively longer with an increase in the average fluorescence lifetime as a function of emission wavelength and thus provide a clear indication of slow nanosecond solvent relaxation.

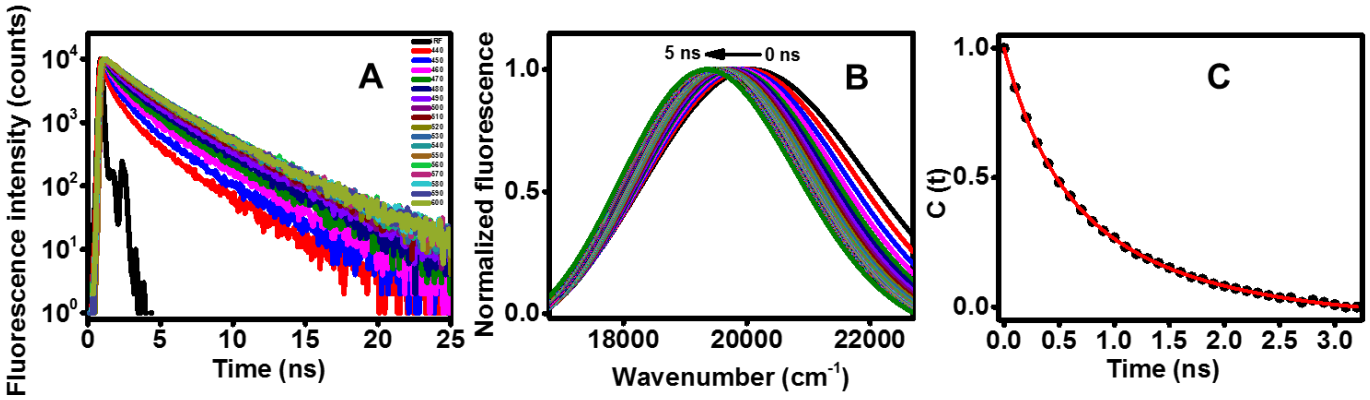


Figure 2.5 (A) Time-resolved fluorescence decays of acrylodan in native state of κ-casein obtained at different emission wavelengths from 440 to 600 nm. (B) Time-resolved emission spectra (TRES) constructed from the time-resolved fluorescence decays at different wavelengths. (C) Solvation correlation function (filled circle) and the biexponential fit (solid red line) to obtain the solvation time. From independent experiments, the total frequency shift was estimated to be 739 ± 60 cm^{-1} and the recovered solvation times are τ_{s1} = 0.38 ± 0.10 ns and τ_{s2} = 1.39 ± 0.15 ns.

The time-resolved spectra constructed from these decays exhibit gradual red shift in frequency, total shift of ~740 cm^{-1} (Figure 2.5B) (for details, see Experimental Methods). The hydration correlation function [C(t)] was then plotted as a function of time. The fitting of C(t) using a sum exponentials yielded two well-separated solvation time components: τ_{s1} ~ 0.4 ns and τ_{s2} ~ 1.4 ns (Figure 2.5C). The longer relaxation time is 3 orders of magnitude slower than

the slowest component (picoseconds) found in bulk water.[45] In the presence of SDS, the longer component of solvation time progressively increased from ~1.4 to ~2 ns as a function of SDS concentration (Figure 2.6 and Table 2.1). Taken together, the nanosecond reorientation of water molecules suggests that the compact yet unordered globules of κ-casein contain motionally restrained water molecules.

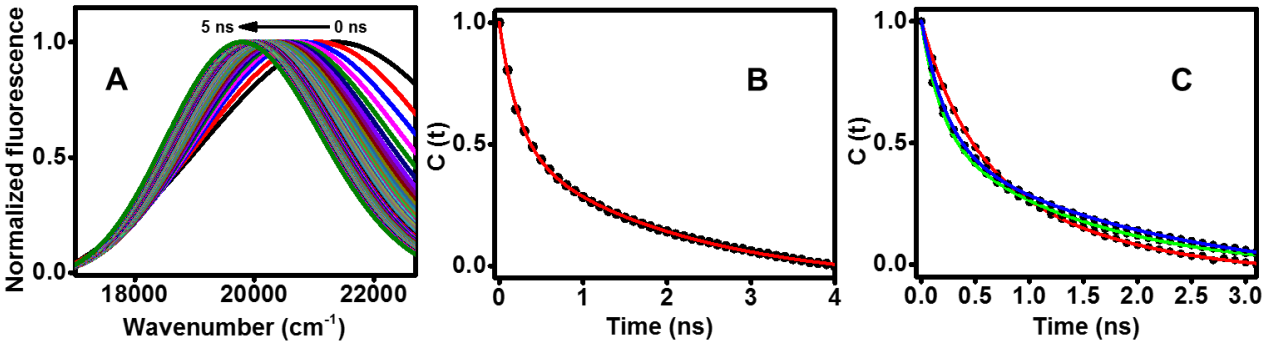


Figure 2.6 (A) Time-resolved emission spectra (TRES) constructed from the time-resolved fluorescence decays at different wavelengths for acrylodan in the SDS-induced state of κ-casein ([SDS] = 500 μM). The total frequency shift was 1565 cm^{-1}. (B) The solvation correlation function and the bi-exponential fit (solid red line). The longer component of solvation time (τ_{s2}) was 2.1 ns. (C) Comparison of the solvation correlation functions showing dampening of solvent relaxation as a function of SDS concentration. Red, green, and blue lines represent the fits for the correlation functions in the absence of SDS, in the presence of 100 and 500 μM, respectively.

Table 2.1

The parameters recovered upon fitting the solvation correlation functions constructed form TCSPC measurements:

[SDS]	τ_{s1} (ns)	α_{s1}	τ_{s2} (ns)	α_{s2}	$\Delta\nu$ (cm^{-1})
0 μM	0.4	0.4	1.27	0.6	734
100 μM	0.18	0.4	1.44	0.6	1329
250 μM	0.24	0.4	1.65	0.6	1322
500 μM	0.23	0.4	2.05	0.5	1565

2.3.4 Origin of ordered water within the collapsed globules of κ-casein

The N-terminal segment (1−100) of κ-casein possesses an amino acid composition that is reminiscent of a Q/N/Y-rich intrinsically disordered prion determinant of the yeast prion protein, Sup35. In this segment, the composition (Q + N + Y) is 27%, which is much higher compared to the average natural occurrence (∼11%) in proteins but much lower compared to the composition found in the N-domain (1-121) of Sup 35 (61%). We have previously shown that under native condition this N-region of κ-casein undergoes a chain collapse, a phenomenon that is also observed for Sup35, polyQ, and many other amyloidogenic IDPs.[17] The N-segment of κ-casein has a net charge of +5.3 at neutral pH. In the presence of low nondenaturing concentration of an anionic detergent, possibly because of the charge neutralization, the N-terminal segment demonstrates a transition to a more compact yet disordered wet globule. The coil-to-globule transition as a result of chain collapse of an IDP in a poor solvent, such as water, has important implications in amyloid formation. However, the nature of water present within an unstructured globule remained unexplored and elusive. Our results demonstrate that the water molecules within the globules are highly restrained. Many previous theoretical and experimental investigations on the relaxation of bulk water revealed that the solvation time ranges between 100s of femtoseconds to picoseconds.[45] On the contrary, water molecules present on the protein surface, also termed as "biological water", has much slower relaxation time ranging between 10 and 100 ps.[45,46] The mobility of water molecules at the protein surface has been implicated in various important events like protein folding, molecular recognition, and enzyme catalysis.[45,47–49] In our case, the water relaxation occurs on the nanosecond time scale and is 3 orders of magnitude slower than that of bulk water. Therefore, the water within the disordered globule is significantly more ordered than the "biological water" at the protein surface and resembles the constrained water pool found in nanoconfined systems such as reverse micelles and molten-globule intermediates.[42,49] Taken together, our results show the presence of a highly restrained water molecules within the IDP globule (Figure 2.7). A recent study has also indicated the restricted water motions on the surface of an amyloidogenic IDP, namely tau.[54]

What is the origin of highly ordered water within the compact disordered globule? The chain collapse leads to the formation of globules in which backbone amides participate in a hydrogen bonding network with the entrapped water molecules. The amide side chains from glutamines (Q) and asparagines (N) as well as polar hydroxyl groups from tyrosines can also

enrich the hydrogen-bonded network of water molecules in the collapsed globule. These water molecules become even more strongly networked and highly ordered when the globules undergo more compaction by decreasing the internal charge per residue. Additionally, the chain fluctuations can also contribute to the solvation dynamics.[47]

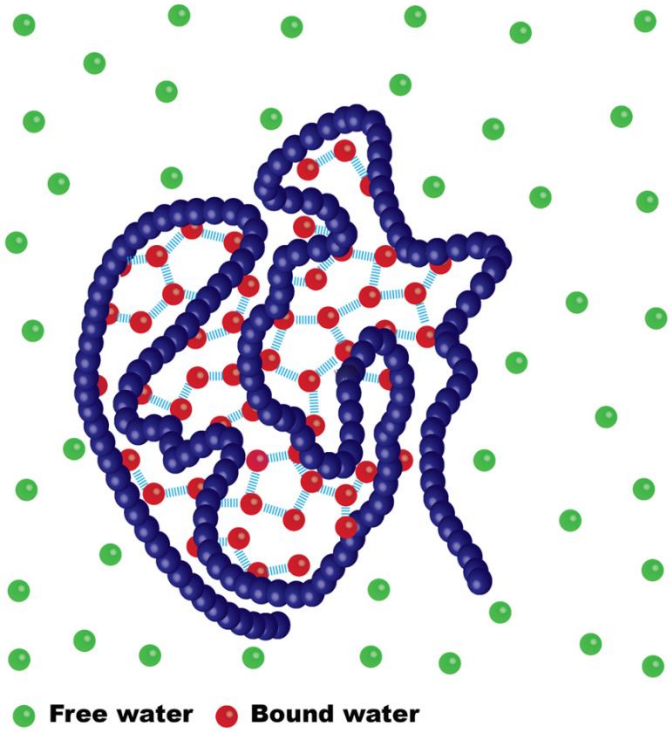


Figure 2.7 Schematic model showing ordered water cluster within the collapsed disordered state of the globule.

What is the implication of ordered water in the collapsed globules of an amyloidogenic protein? Experiments and simulations on amyloidogenic IDP consisting of Q/N-rich sequence have indicated that the formation of key oligomeric intermediates is crucial for amyloid assembly.[14,19,26,36] These obligatory oligomers are formed by coalescence of collapsed globules. We speculate that the process of oligomer formation will be favored by an entropic gain arising out of the release of ordered water molecules from the interior of the globule to the bulk milieu.[55] The conformational maturation of the polypeptide within the oligomers is likely to lead to the formation of ordered amyloid fibrils via the nucleated conformational conversion mechanism.[35,36,56]

2.4 Summary

Here, we have characterized the nature of water molecules entrapped (nanosecond solvation) within the collapsed globules of an amyloidogenic IDP, namely, κ-casein. These globules were found to undergo further compaction in the presence of SDS which is an anionic detergent and also a well-known lipid mimetic. The anionic detergent was able to diminish the intra-chain repulsion from the positively charged glutamine/asparagine-rich amyloidogenic N-terminal domain comprising 100 residues. Using picosecond-resolved fluorescence spectroscopy, we estimated the longer component of the solvation time to be ∼1.4 ns, which is 3 orders of magnitude slower than that in bulk water and more than an order of magnitude slower than the "biological water" present at the protein surface. These profoundly restrained water molecules present within the collapsed IDP globules resemble the ordered water cluster found under nanoconfinement. We suggest that the association of these globules would result in the release of ordered water molecules into the bulk milieu causing an entropic gain that would eventually drive the formation of the key (obligatory) oligomeric intermediates on the pathway to amyloids via nucleation-dependent polymerization. This particular aspect of amyloid aggregation along with the ultrafast solvation in the monomeric form of κ-casein are discussed in Chapter 3.

2.5 References:

1. Babu, M. M.; Kriwacki, R. W.; Pappu, R. V. Structural Biology. Versatility from Protein Disorder. *Science* **2012**, *337*, 1460–1461.

2. Dunker, A. K.; Babu, M. M.; Barbar, E.; Blackledge, M.; Bondos, S. E.; Dosztányi, Z.; Jane, H.; Forman-Kay, J.; Fuxreiter, M.; Gsponer, J. et al. What's in a Name? Why these Proteins are Intrinsically Disordered. *Intrinsically Disordered Proteins* **2013**, *1*, 1–5.

3. Dunker, A. K.; Obradovic, Z. The Protein Trinity-Linking Function and Disorder. *Nat. Biotechnol.* **2001**, *19*, 805–806.

4. Uversky, V. N.; Dunker, A. K. Understanding Protein Non-Folding. *Biochim. Biophys. Acta* **2010**, *1804*, 1231–1264.

5. Dyson, H. J.; Wright, P. E. Intrinsically Unstructured Proteins and their Functions. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 197–208.

6. Tompa, P.; Fuxreiter, M. Fuzzy complexes: Polymorphism and Structural Disorder in Protein-Protein Interactions. *Trends Biochem. Sci.* **2008**, *33*, 2–8.

7. Mittag, T.; Kay, L. E.; Forman-Kay, J. D. Protein Dynamics and Conformational Disorder in Molecular Recognition. *J. Mol. Recognit.* **2009**, *23*, 105–116.

8. Metallo, S. J. Intrinsically Disordered Proteins are Potential Drug Targets. *Curr. Opin. Chem. Biol.* **2010**, *14*, 481–488.

9. Tompa, P.; Csermely, P. The Role of Structural Disorder in the Function of RNA and Protein Chaperones. *FASEB J.* **2004**, *18*, 1169–1175.

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

11. Tompa, P.; Fersht, A. *Structure and Function of Intrinsically Disordered Proteins*; CRC Press, **2010**.

12. Chiti, F.; Dobson, C. M. Protein Misfolding, Functional Amyloid, and Human Disease. *Annu. Rev. Biochem.* **2006**, *75*, 333–366.

13. Uversky, V. N.; Oldfield, C. J.; Dunker, A. K. Intrinsically Disordered Proteins in Human Diseases: Introducing the D2 Concept. *Annu. Rev. Biophys.* **2008**, *37*, 215–246.

14. Mukhopadhyay, S.; Krishnan, R.; Lemke, E. A.; Lindquist, S.; Deniz, A. A. A Natively Unfolded Yeast Prion Monomer Adopts an Ensemble of Collapsed and Rapidly Fluctuating Structures. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 2649–2654.

15. Mao, A. H.; Crick, S. L.; Vitalis, A.; Chicoine, C. L.; Pappu, R. V. Net Charge Per Residue Modulates Conformational Ensembles of Intrinsically Disordered Proteins. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 8183–8188.

16. Maiti, N. C.; Apetri, M. M.; Zagorski, M. G.; Carey, P. R.; Anderson, V. E. Raman Spectroscopic Characterization of Secondary Structure in Natively Unfolded Proteins: Alpha-Synuclein. *J. Am. Chem. Soc.* **2004**, *126*, 2399–2408.

17. Jain, N.; Bhattacharya, M.; Mukhopadhyay, S. Chain Collapse of an Amyloidogenic Intrinsically Disordered Protein. *Biophys. J.* **2011**, *101*, 1720–1729.

18. Das, R. K.; Pappu, R. V. Conformations of Intrinsically Disordered Proteins are Influenced by Linear Sequence Distributions of Oppositely Charged Residues. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 13392-13397.

19. Crick, S. L.; Jayaraman, M.; Frieden, C.; Wetzel, R.; Pappu, R. V. Fluorescence Correlation Spectroscopy Shows that Monomeric Polyglutamine Molecules Form Collapsed Structures in Aqueous Solutions. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 16764–16769.

20. Walters, R. H.; Murphy, R. M. Examining polyglutamine peptide length: a connection between collapsed conformations and increased aggregation. *J. Mol. Biol.* **2009**, *393*, 978–992.

21. Wang, X.; Vitalis, A.; Wyczalkowski, M. A.; Pappu, R. V. Characterizing the Conformational Ensemble of Monomeric Polyglutamine. *Proteins* **2006**, *63*, 297–311.

22. Tran, H. T.; Mao, A.; Pappu, R. V. Role of Backbone-Solvent Interactions in Determining Conformational Equilibria of Intrinsically Disordered Proteins. *J. Am. Chem. Soc.* **2008**, *130*, 7380–7392.

23. Uversky, V. N.; Gillespie, J. R.; Fink, A. L. Why are "Natively Unfolded" Proteins Unstructured Under Physiologic Conditions? *Proteins* **2000**, *41*, 415–427.

24. Müller-Späth, S.; Soranno, A.; Hirschfeld, V.; Hofmann, H.; Rüegger, S.; Reymond, L.; Nettels, D.; Schuler, B. Charge Interactions can Dominate the Dimensions of

Intrinsically Disordered Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 14609–14614.

25. Yamada, J.; Phillips, J. L.; Patel, S.; Goldfien, G.; Calestagne-Morelli, A.; Huang, H.; Reza, R.; Acheson, J.; Krishnan, V. V; Newsam, S. et al. A Bimodal Distribution of Two Distinct Categories of Intrinsically Disordered Structures with Separate Functions in FG Nucleoporins. *Mol. Cell. Proteomics* **2010**, *9*, 2205–2224.

26. Vitalis, A.; Pappu, R. V. Assessing the Contribution of Heterogeneous Distributions of Oligomers to Aggregation Mechanisms of Polyglutamine Peptides. *Biophys. Chem.* **2011**, *159*, 14–23.

27. Möglich, A.; Joder, K.; Kiefhaber, T. End-to-End Distance Distributions and Intrachain Diffusion Constants in Unfolded Polypeptide Chains Indicate Intramolecular Hydrogen Bond Formation. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 12394–12399.

28. Reid, I. M. Corpora Amylacea of the Bovine Mammary Gland: Histochemical and Electron Microscopic Evidence for their Amyloid Nature. *J. Comp. Pathol.* **1972**, *82*, 409–413.

29. Ecroyd, H.; Thorn, D. C.; Liu, Y.; Carver, J. A. The Dissociated Form of Kappa-Casein is the Precursor to its Amyloid Fibril Formation. *Biochem. J.* **2010**, *429*, 251–260.

30. Ecroyd, H.; Koudelka, T.; Thorn, D. C.; Williams, D. M.; Devlin, G.; Hoffmann, P.; Carver, J. A. Dissociation from the Oligomeric State is the Rate-Limiting Step in Fibril Formation by Kappa-Casein. *J. Biol. Chem.* **2008**, *283*, 9012–9022.

31. Sokolovski, M.; Sheynis, T.; Kolusheva, S.; Jelinek, R. Membrane Interactions and Lipid Binding of Casein Oligomers and Early Aggregates. *Biochim. Biophys. Acta* **2008**, *1778*, 2341–2349.

32. Edmiston, P. L.; Wambolt, C. L.; Smith, M. K.; Saavedra, S. S. Spectroscopic Characterization of Albumin and Myoglobin Entrapped in Bulk Sol-Gel Glasses. *J. Colloid Interface Sci.* **1994**, *163*, 395–406.

33. Narang, D.; Sharma, P. K.; Mukhopadhyay, S. Dynamics and Dimension of an Amyloidogenic Disordered State of Human β(2)-Microglobulin. *Eur. Biophys. J.* **2013**, *42*, 767–776.

34. Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*; 3rd ed.; Springer Science, New York **2006**.

35. Krishnan, R.; Lindquist, S. L. Structural Insights into a Yeast Prion Illuminate Nucleation and Strain Diversity. *Nature* **2005**, *435*, 765–772.

36. Serio, T. R.; Cashikar, A. G.; Kowal, A. S.; Sawicki, G. J.; Moslehi, J. J.; Serpell, L.; Arnsdorf, M. F.; Lindquist, S. L. Nucleated Conformational Conversion and the Replication of Conformational Information by a Prion Determinant. *Science* **2000**, *289*, 1317–1321.

37. Demchenko, A. P. Site-Selective Red-Edge Effects. *Methods Enzymol.* **2008**, *450*, 59–78.

38. Chattopadhyay, A.; Rawat, S. S.; Kelkar, D. A.; Ray, S.; Chakrabarti, A. Organization and Dynamics of Tryptophan Residues in Erythroid Spectrin: Novel Structural Features of Denatured Spectrin Revealed by the Wavelength-Selective Fluorescence Approach. *Protein Sci.* **2003**, *12*, 2389–2403.

39. Haldar, S.; Chaudhuri, A.; Chattopadhyay, A. Organization and Dynamics of Membrane Probes and Proteins Utilizing the Red Edge Excitation Shift. *J. Phys. Chem. B* **2011**, *115*, 5693–5706.

40. Kelkar, D. A.; Chattopadhyay, A.; Chakrabarti, A.; Bhattacharyya, M. Effect of Ionic Strength on the Organization and Dynamics of Tryptophan Residues in Erythroid Spectrin: A Fluorescence Approach. *Biopolymers* **2005**, *77*, 325–334.

41. Chattopadhyay, A.; Haldar, S. Dynamic Insight into Protein Structure Utilizing Red Edge Excitation Shift. *Acc. Chem. Res.* **2014**, *47*, 12–19.

42. Bhattacharyya, K.; Bagchi, B. Slow Dynamics of Constrained Water in Complex Geometries. *J. Phys. Chem. A* **2000**, *104*, 10603–10613.

43. Jana, B.; Pal, S.; Bagchi, B. Hydrogen Bond Breaking Mechanism and Water Reorientational Dynamics in the Hydration Layer of Lysozyme. *J. Phys. Chem. B* **2008**, *112*, 9112–9117.

44. Balasubramanian, S.; Bagchi, B. Slow Solvation Dynamics Near an Aqueous Micellar Surface. *J. Phys. Chem. B* **2001**, *105*, 12529–12533.

45. Pal, S. K.; Peon, J.; Zewail, A. H. Biological Water at the Protein Surface: Dynamical Solvation Probed Directly with Femtosecond Resolution. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 1763–1768.

46. Nandi, N.; Bagchi, B. Dielectric Relaxation of Biological Water. *J. Phys. Chem. B* **1997**, *101*, 10954–10961.

47. Pal, S. K.; Peon, J.; Bagchi, B.; Zewail, A. H. Biological Water: Femtosecond Dynamics of Macromolecular Hydration. *J. Phys. Chem. B* **2002**, *106*, 12376–12395.

48. Bagchi, B. Water Dynamics in the Hydration Layer Around Proteins and Micelles. *Chem. Rev.* **2005**, *105*, 3197–3219.

49. Sen, P.; Mukherjee, S.; Dutta, P.; Halder, A.; Mandal, D.; Banerjee, R.; Roy, S.; Bhattacharyya, K. Solvation Dynamics in the Molten Globule State of a Protein. *J. Phys. Chem. B* **2003**, *107*, 14563–14568.

50. Pal, S. K.; Zewail, A. H. Dynamics of Water in Biological Recognition. *Chem. Rev.* **2004**, *104*, 2099–2123.

51. Jha, A.; Ishii, K.; Udgaonkar, J. B.; Tahara, T.; Krishnamoorthy, G. Exploration of the Correlation Between Solvation Dynamics and Internal Dynamics of a Protein. *Biochemistry* **2011**, *50*, 397–408.

52. Batabyal, S.; Mondol, T.; Pal, S. K. Picosecond-Resolved Solvent Reorganization and Energy Transfer in Biological and Model Cavities. *Biochimie* **2013**, *95*, 1127–1135.

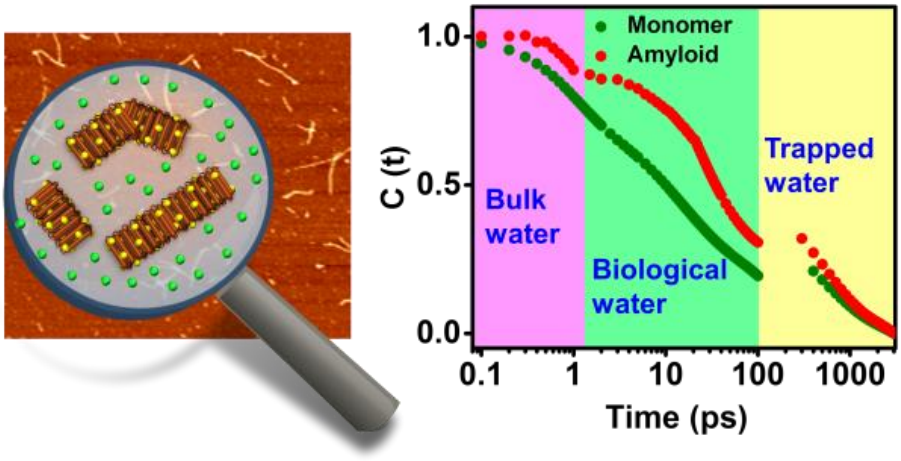
53. Boyd, A. E.; Marnett, A. B.; Wong, L.; Taylor, P. Probing the Active Center Gorge of Acetylcholinesterase by Fluorophores Linked to Substituted Cysteines. *J. Biol. Chem.* **2000**, *275*, 22401–22408.

54. Gallat, F.-X.; Laganowsky, A.; Wood, K.; Gabel, F.; van Eijck, L.; Wuttke, J.; Moulin, M.; Härtlein, M.; Eisenberg, D.; Colletier, J.-P.; Zaccai, G.; Weik, M. Dynamical Coupling of Intrinsically Disordered Proteins and their Hydration Water: Comparison with Folded Soluble and Membrane Proteins. *Biophys. J.* **2012**, *103*, 129–136.

55. Thirumalai, D.; Reddy, G.; Straub, J. E. Role of Water in Protein Aggregation and Amyloid Polymorphism. *Acc. Chem. Res.* **2012**, *45*, 83-92.

56. Lee, J.; Culyba, E. K.; Powers, E. T.; Kelly, J. W. Amyloid-β Forms Fibrils by Nucleated Conformational Conversion of Oligomers. *Nat. Chem. Biol.* **2011**, *7*, 602–609.

Water Rearrangements upon Disorder-to-Order Amyloid Transition



The work described in this chapter has been published in the *Journal of Physical Chemistry Letters*.

Reference: Arya et al. *J. Phys. Chem. Lett.* **2016,** *7* (20), 4105-4110.

3.1 Introduction

Water shares an intimate relationship with proteins and has a profound influence on the protein structure, dynamics and function.[1–7] The water present in the vicinity of proteins has significantly different dielectric and dynamical properties than the bulk water and is termed as the "biological water".[1] These motionally restrained water molecules are present at the protein interfaces, in the cavities or clefts, and play a critical role in the enzyme catalysis and protein function. The functional state of proteins is attained by the folding process; however, under certain circumstances, proteins misfold to an aberrant state that aggregates to form ordered amyloid fibrils comprising an extensively hydrogen bonded cross-β core. Amyloids epitomize a remarkable example of an exquisite supramolecular protein assembly and are associated with a plethora of fatal human diseases.[8–15] Although a wealth of information is available on the amyloids from the protein perspective, very little is understood about the intriguing role of water in amyloid formation. Because the intricate balance between chain-chain and chain-solvent interactions is critical for amyloid aggregation,[16,17] it is important to discern the distinct spatiotemporal characteristics of the water structure during the transformation from the functional native form to the pathological amyloid state. Various theoretical and experimental tools have emerged to study the properties of water in proteins and amyloids.[18–27] For instance, the atomic-level structural information obtained from X-ray crystallographic studies of amyloid peptides-based microcrystals revealed the presence of dry and wet β-sheet interfaces in amyloid assembly.[22] The dry interfaces are devoid of water molecules, whereas the wet interfaces entrap water molecules and participate in inter-strand hydrogen bonding.[22] However, these measurements are not amenable to the temporal assignments of water reorganization upon amyloid formation. Therefore, an experimental approach that offers an effective spatiotemporal resolution and a high sensitivity is the femtosecond time-resolved fluorescence spectroscopy using an environment-sensitive fluorophore attached to a specific location of the protein.[25,26] These measurements allow us to probe the hydration dynamics at the vicinity of the protein from femtoseconds to picoseconds during which hydrogen bond rearrangements occur within the water network.[1]

In this work, we have chosen a model amyloidogenic intrinsically disordered protein (IDP), bovine κ-casein,[28,29] the deposition of which is linked to mammary cancer.[30,31] Under the physiological condition, monomeric κ-casein adopts a (noncanonical) compact disordered state containing highly retarded water molecules.[28] Here, we delineate the role of water restructuring during the amyloid transition of disordered κ-casein.

3.2 Experimental Section

3.2.1 Materials

κ-casein (from bovine milk), iodoacetic acid, dithiothreitol (DTT), guanidium chloride, tris HCl, thioflavin-T (ThT) and sodium hydrogen phosphate (monobasic) were procured from Sigma (St. Louis, MO) and used without any further modification. 6-Acryloyl-2-(dimethylamino)naphthalene (acrylodan) was purchased from Molecular Probes, Invitrogen Inc. All solutions were prepared using Milli Q water. The pH of the buffers used was adjusted using on Metrohm (Herisau, Switzerland) 827 pH meter at ~25°C.

3.2.2 Aggregation reaction

Prior to initiating aggregation reaction, bovine κ-casein was monomerized for which the cysteines were reduced using DTT and iodoacetic acid, as described previously.[28] For following aggregation kinetics, aggregation reaction was initiated using 200 μM of reduced and carboxymethylated κ-casein. The aliquots were taken out at different time points (10 μM) and ThT fluorescence (10 μM) was measured. The fluorescence at different points was then plotted and fitted with an equation for nucleation-dependent polymerization (Arora et al. *FEBS Lett.* 2004, *564*, 121-125). For carrying out aggregation reaction with acrylodan labeled κ-casein, κ-casein was reduced using DTT and then labeled with acrylodan as described previously.[28] The aggregation reaction was initiated by mixing 100 μM of reduced (but unlabeled) κ-casein and 100 μM of acrylodan labeled κ-casein (stored under denatured condition) in 50 mM phosphate buffer (pH 7.2). The reaction mixture was continuously stirred at a speed of 300 rpm and the temperature was maintained at 37°C. After the reaction, the labeled protein was then passed through a PD-10 column to remove excess dye and was further concentrated using AMICON ultra (3 kDa cutoff; Millipore). The buffer used for dilution in all the experiments described below was same as the aggregation buffer (pH 7.2, 50 mM phosphate buffer). The monomeric/IDP (0 hour) and amyloid (45 hour) samples were prepared by taking aliquots of reaction mixture at the respective time points and diluting them with the dilution buffer to obtain the desired final concentration.

3.2.3 Circular dichroism (CD) measurements

For far-UV CD measurements, the samples were diluted 10 fold to get a final concentration of 20 μM. A 1 mm pathlength quartz cell was used. The CD spectra were collected in Chirascan Spectrophotometer (Applied Photophysics, UK) and a scan range of 200-260 nm was used with

a bandwidth of 1 nm. The spectra obtained were corrected for the buffer background and were further smoothened using Pro Data software. Finally, the spectra were plotted in Origin Pro 8.5 software.

3.2.4 Atomic force microscopy (AFM) imaging of amyloid fibrils

Prior to AFM measurements, the reaction mixture was diluted 200 fold using filtered dilution buffer and deposited on to the freshly cleaved and buffer washed muscovite mica (Grade V-4 mica from SPI, PA). Then, the sample was incubated for 30 minutes to ensure its adherence over the mica surface. After 30 minutes, the mica was washed with filtered buffer and dried under a gentle stream of nitrogen for about half an hour. The images were collected on MultiView 2000 instrument (Nanonics Imaging Ltd, Jerusalem, Israel) operating in intermittent contact mode based on the phase feedback technique. A Cr-coated cantilevered glass probe of diameter ~10 nm oscillating at a resonance frequency of ~32 kHz was used for AFM scanning. AFM height image of 310 x 310 pixels resolution, covering area of 3.3 µm x 3.3 µm with a sample delay of 4 ms was acquired using the Quartz software (provided with MultiView) and processed using the WSxM software (Horcas et al. *Rev. Sci. Instrum.* 2007, *78*, 013705).

3.2.5 Steady-state fluorescence measurements

All the steady-state fluorescence data were acquired in Fluoromax-4 (Horiba Jobin Yvon, NJ). The samples were excited at 375 nm (acrylodan) and 450 nm (ThT). The steady state fluorescence anisotropies were measured at 490 nm (acrylodan). The steady-state fluorescence anisotropy (r_{ss}) is given by the following equation:

$$r_{ss} = (I_{\parallel} - I_{\perp}G) / (I_{\parallel} + 2I_{\perp}G) \qquad (1)$$

where $I_{\parallel}$ and $I_{\perp}$ are fluorescence intensities collected using parallel and perpendicular geometry, respectively. The perpendicular components were always corrected using a G-factor. For the tryptophan red-edge excitation shift (REES) measurements, the excitation wavelength for tryptophan was varied from 280 nm to 305 nm.

3.2.6 Femtosecond fluorescence up-conversion measurements

For fluorescence up-conversion measurements (FOG 100, CDP, Russia), a second harmonic light of mode locked Ti:Sapphire laser (Tsunami, Spectra Physics, USA) with pulses of 100 fs (1 W, 80 MHz repetition rate), was used as the excitation source. In order to generate 800 nm

light, the self-mode-locked Ti:Sapphire laser was pumped by continuous wave second harmonic light of the output of 4.5 W Nd:YVO_4 laser (Millenia Pro, Spectra Physics, USA) which was further pumped by an electrically pumped diode laser. Around 25 mW of 400 nm light was finally used for exciting the samples, which were housed in a rotating cell. The polarization of the excitation light was fixed at the magic angle (54.7°) with respect to the gate light. For setting up the magic angle polarization, a set of Glan polarizer and Berek's variable waveplate was used. The upconversion of emission light was done using 800 nm gate light in 1 mm thick β-BBO (type I) crystal. The resultant up-converted light was passed through a UV bandpass filter which cuts excitation, emission and gate light and then passed through a double monochromator (DM2100, CDP, Russia). A photo-multiplier tube (PMT) having dark noise <5 cps was used for fluorescence detection. The up-conversion transients were collected from 475-585 nm at 10-nm intervals till 300 ps. An appropriate integration time and step size was used so that enough counts can be collected without photodegradation. The instrument response function (IRF) was estimated by measuring the cross-correlation between the gate light and Raman signal of buffer (400 nm excitation). The resulting profile when fitted with a Gaussian equation yielded a full width at half maxima (FWHM) of ~ 250 fs. The emission decay traces were fitted by reconvolution to two/three exponentials in a homebuilt program in Igor Pro 6.37 (Burai et al. *J. Chem. Phys.* 2009, *131*, 034504). The raw- and the fitted data were then plotted in Origin Pro 8.5 software. For anisotropy measurements, the fluorescence was collected at 515 nm and the orientation of the excitation light was 0° and 90° with respect to the gate light for parallel ($I_{\parallel}$) and perpendicular ($I_{\perp}$) fluorescence intensities, respectively.

3.2.7 Time-correlated single photon counting (TCSPC) measurements:

For time-resolved emission spectra (TRES): The fluorescence decays were collected at 54.7° (magic angle) polarization using time-correlated single photon counting (TCSPC) module (Fluorocube, Horiba Jobin Yvon, NJ). The peak count was fixed at 10,000. A laser diode was used as excitation source for acrylodan (375 nm). The colloidal silica (Ludox) solution was used for collecting the instrument response function (IRF). The IRF had a full width at half maxima (FWHM) of ~ 250 ps. The decay profiles were collected from 475-585 nm at 10-nm interval. The decays were then analyzed and fitted taking the IRF into account using DAS6 software provided with the TCSPC system.

For fluorescence anisotropy decay: For fluorescence anisotropy decay measurements, the peak difference was 10,000 counts and the orientation of emission polarizer was 0° and 90°

with respect to the excitation polarizer for parallel fluorescence intensities ($I_{\parallel}$) and perpendicular fluorescence intensities ($I_{\perp}$), respectively. The emission monochromator for acrylodan was fixed at 500 nm with a bandpass of 4 nm and 375 nm laser diode was used as excitation source for acrylodan. The width (FWHM) of instrument response function (IRF) was collected using Ludox (colloidal silica) and was found to be ~250 ps. The anisotropy decays were analyzed by globally fitting $I_{\parallel}(t)$ and $I_{\perp}(t)$, as described previously.[32]

3.2.8 Construction of TRES from up-conversion and TCSPC measurements and data analysis

The time-resolved emission spectra (TRES) analysis was done using our previously reported protocol.[28] Briefly, the fluorescence decay parameters recovered were used to simulate the fluorescence decays at different emission wavelengths in Origin Pro 8.5 software using $I_{\lambda}(t) = \sum_i \alpha_i \exp(-t/\tau_i)$, where α_i and τ_i represent the amplitude and lifetime of the different lifetime components, respectively. The time-resolved emission spectra (TRES) were constructed using:

$$I(\lambda, t) = I_{\lambda}^{SS} I_{\lambda}(t) / \sum_i \alpha_i \tau_i \qquad (2)$$

where I_{λ}^{SS} represents the steady-state fluorescence intensity at a fixed emission wavelength λ and $I_{\lambda}(t)$ is same as mentioned above. The spectra obtained at different time points were fitted using lorentzian function to recover the peak frequency $\nu(t)$ and were then normalized for the estimation of overall time-dependent fluorescence Stokes shift (TDFSS). The solvation correlation function was then constructed using:

$$C(t) = [\nu(t) - \nu(\infty)] / [\nu(0) - \nu(\infty)] \qquad (3)$$

where $\nu(0)$, $\nu(t)$ and $\nu(\infty)$ are the emission peak frequencies at time 0, t and ∞, respectively. $\nu(0)$ estimated from femtosecond measurements was 21,000 ± 150 cm^{-1}, which is close to the time zero reported previously[7] for acrylodan using the method desribed by Fee and Maroncelli.[37] The correlation function was finally fitted using a sum of exponentials as the following:

$$C(t) = \sum_i \alpha_{si} \exp(-t/\tau_{si}) \qquad (4)$$

where α_{si} and τ_{si} represent the amplitude and solvation time of the different solvation components, respectively. The solvation correlation function constructed from femtosecond upconversion data was fitted till 200 ps. For fitting C(t) from TCSPC, $\nu(0)$ obtained from our

femtosecond experiments was used. The ν(∞) used in both femtosecond up-conversion and TCSPC data was ν at 3 ns and was close to ν for the steady-state fluorescence spectrum. In order to extract the longest component of solvation, C(t) from TCSPC was fitted by fixing the initial two solvation components recovered from the femtosecond data analysis.

3.3 Results and Discussion

3.3.1 Characterization of amyloid state using a variety of biophysical tools

In order to directly monitor the reorientation dynamics of water during the amyloid assembly, we have utilized both femtosecond and picosecond time-resolved spectroscopy, in combination with other experimental tools. As a prelude, we first characterized amyloid formation from κ-casein under the physiological condition using a variety of biophysical tools. We used a well-known amyloid marker, thioflavin-T (ThT), to monitor the kinetics of amyloid aggregation that indicated a typical nucleation-dependent polymerization mechanism (Figure 3.1A). Amyloid formation was accompanied by a characteristic conformational change from random coil to β-sheet as demonstrated by the CD data (Figure 3.1B). AFM (atomic force microscopy) imaging revealed the nanoscale morphology of amyloid fibrils (Figure 3.1B inset). We further characterized the amyloid state using various environment-sensitive fluorescence readouts.

3.3.2 Solvent accessibility of the amyloid state probed by steady-state and time-resolved fluorescence

Since acrylodan is a highly solvent-sensitive fluorophore exhibiting a large Stokes shift and has been used previously for hydration studies in proteins,[32–34] we covalently labeled the thiol group of Cys residues in κ-casein using acrylodan (Figure 3.1C). The Cys residues are located in a highly amyloidogenic, (prion-like) glutamine (Q)- and asparagine (N)-rich, segment.[28] The disordered monomeric form exhibited an emission peak at ∼515 nm, indicating that the fluorophore is highly solvent exposed.[28] Upon amyloid formation, a small blue shift of only ∼5 nm suggested that the fluorophore experiences predominantly aqueous environment in the amyloid state (Figure 3.1C). Next, we performed fluorescence depolarization measurements using time-correlated single photon counting (TCSPC) to monitor the rotational dynamics of the fluorophore in both monomeric- and amyloid state. Typical biexponential fluorescence anisotropy decays indicated the (fast) local motion of the fluorophore attached to the protein and the (slow) global tumbling of the entire protein (Figure 3.1D). Upon amyloid aggregation, both local and global motions became slower, as expected (Table 3.1). However, a significant

amplitude (~30%) of local motion is retained in amyloid, suggesting that acrylodan is flexible and is likely to be exposed to water corroborating our steady-state fluorescence data. We also performed the femtosecond time-resolved anisotropy measurements (Figure 3.1D inset) that revealed no significant rotational mobility of the fluorophore on the sub-100 ps time scale (see later).

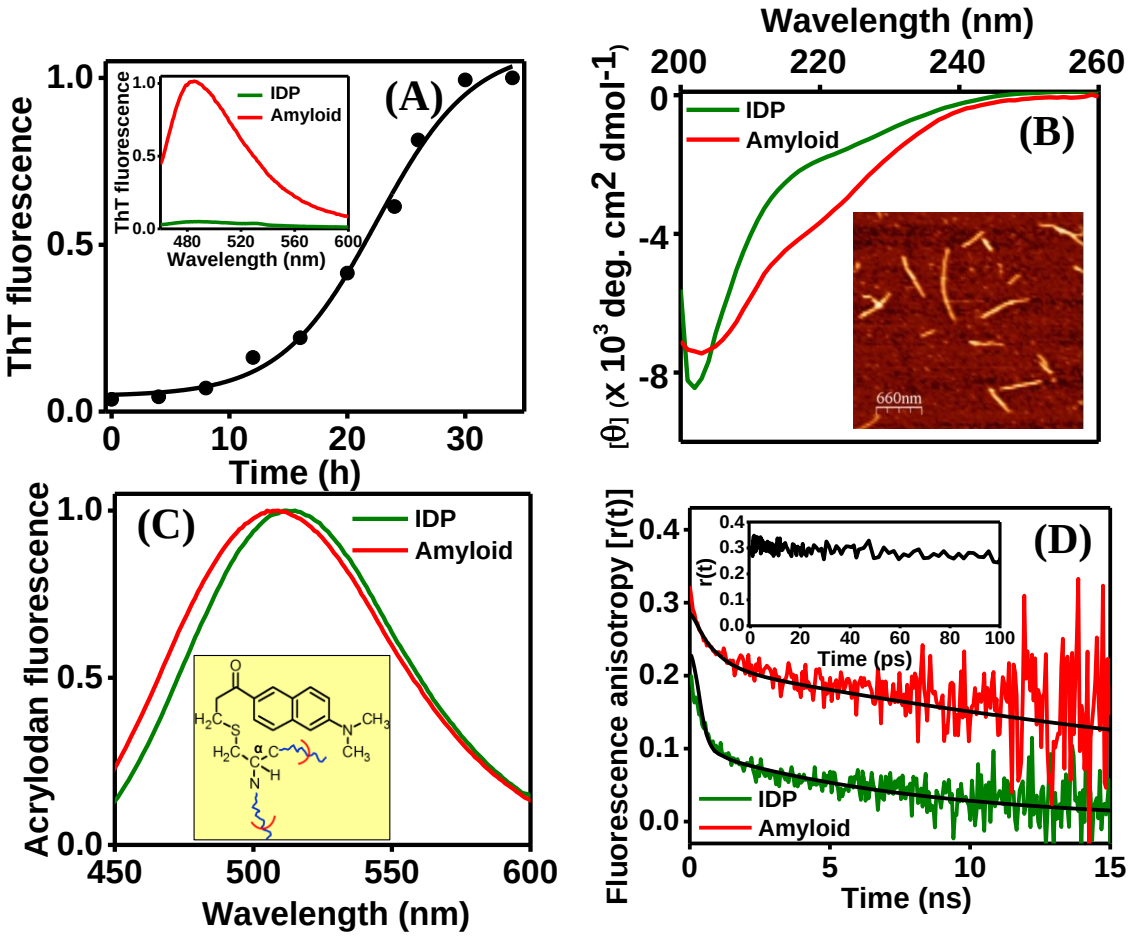


Figure 3.1 (A) Kinetics of κ-casein amyloid formation followed by thioflavin-T (ThT) fluorescence (fitted using the nucleation-dependent polymerization model) with an inset showing ThT fluorescence spectra for monomeric IDP (olive) and amyloid (red). (B) Far-UV circular dichroism (CD) spectra with an inset showing AFM image of amyloid fibrils. (C) Steady-state fluorescence spectra of acrylodan labeled protein in monomeric IDP (olive) and amyloid (red). The inset shows the structure of acrylodan labeled at Cys. (D) Fluorescence anisotropy decays of acrylodan in IDP (olive) and amyloid (red) with fits shown in black (See Table 3.1 for parameters). The inset shows the femtosecond time-resolved anisotropy decay of acrylodan for IDP from 0 to 100 ps.

Table 3.1

The typical parameters recovered from the fluorescence anisotropy decay analyses shown in Figure 3.1D: [§]

Sample	ϕ_{fast} (ns)	β_{fast}	ϕ_{slow} (ns)	β_{slow}	r_{ss}
IDP-state	0.20	0.61	7.9	0.39	0.10
Amyloid-state	0.70	0.27	>20	0.73	0.23

[§] The recovered rotational correlation times and their respective amplitudes are referred to as ϕ and β, respectively.

3.3.3 Hydration dynamics in the IDP and amyloid state of κ-casein probed by femtosecond upconversion and TCSPC

After confirming that acrylodan attached to the protein is appreciably solvent exposed in both the monomeric and amyloid forms, we embarked upon the hydration dynamics studies using time-resolved fluorescence spectroscopic measurements involving femtosecond up-conversion and TCSPC. The efficacy of the combination of these two techniques allowed us to capture a wide range of time scales from femtoseconds (up-conversion) to nanoseconds (TCSPC). These measurements allow us to quantify the extent and the time scale of solvent relaxation that are dictated by the rate of hydrogen bond rearrangements arising due to the diffusion of water molecules around the instantaneously created excited state dipole of the fluorophore.[35] We collected fluorescence upconversion and TCSPC transients separately at different emission wavelengths encompassing the entire spectral range of the acrylodan fluorescence. Subsequently, we constructed the time-resolved emission spectra (TRES) and analyzed the hydration correlation functions [C(t)] to estimate the time-dependent Stokes shift (TDSS) and the solvation times[25,34,36] (see Experimental Section for details of data acquisition and analysis).

In order to separate the bulk water from surface-bound water and to quantify the time scales of hydration with the femtosecond time resolution, we carried out fluorescence upconversion measurements on the IDP as well as on the amyloid state. The representative femtosecond transients at four emission wavelengths are shown in Figure 3.2. The intensity decays exhibited a typical rise component having negative amplitude as the emission wavelength was changed from the blue to the red end of the spectrum. This is a characteristic feature of solvent relaxation at the excited state.[34] The recovered time-zero frequency ($\nu(0)$)

from our femtosecond experiments was close to $\nu(0)$ estimated earlier,[33] using the method described previously.[37] The estimated TDSS within 200 ps [$\Delta\nu(0\rightarrow200ps)$] for the acrylodan-labeled κ-casein monomer and amyloid were found to be ~1400 cm^{-1} and ~1100 cm^{-1}, respectively. The TDSS data indicates a more extensive solvent relaxation for the monomer compared to amyloid. The recovered parameters of hydration dynamics are mentioned in Table 3.2 and are similar to those reported for other proteins.[25–27,34] A number of theoretical and experimental studies have shown that the time scale for bulk water relaxation ranges from hundreds of femtoseconds to picoseconds,[6] whereas "biological water" exhibits considerably slower relaxation time that spans from 10 to 100 ps.[6] Thus, the ultrafast component ($\tau_{s1} \sim 1$ ps) observed in both the cases can be assigned to the bulk water and the two slower components (of tens of picoseconds and longer) to surface bound water. We collectively denote these slower components (τ_{s2} and τ_{s3}) ranging between 10 and 100 ps as bound water type I. We would like to point out that the rotational dynamics of the dye is not likely to contribute to the solvation, since the femtosecond time-resolved anisotropy decay did not exhibit any significant drop within 100 ps (Figure 3.1D inset). A careful look at the hydration correlation plots [C(t) vs t] revealed that the solvation was nearly 90% complete in the IDP state, whereas it was ~78% complete in the amyloid state within 200 ps (Figure 3.2E and Table 3.2). Because the solvation is not 100% complete within the up-conversion experimental time window (200 ps) with respect to $\nu(\infty)$ in both cases, we refer to the unrelaxed (residual) solvation component as bound water type II which has a much longer relaxation time ($\gg$ 100 ps).

In order to delineate the roles of different types of water upon IDP-to-amyloid transition, contributions from the bulk water, bound water type I and type II for both IDP and the amyloid state were compared as shown in Figure 3.3A. Our results demonstrate that upon amyloid formation, the fraction of bulk water (α_{s1}) diminishes, whereas the contribution of bound water type II (indicated by a residual amplitude, α_r) increases almost by similar extent. However, the total fractional contribution of bound water type I ($\alpha_{s2} + \alpha_{s3}$) remains nearly unaltered although the hydration correlation time components, especially τ_{s2}, became much slower suggesting that the surface bound water is more restrained in the amyloid state compared to the monomeric state (Table 3.2).

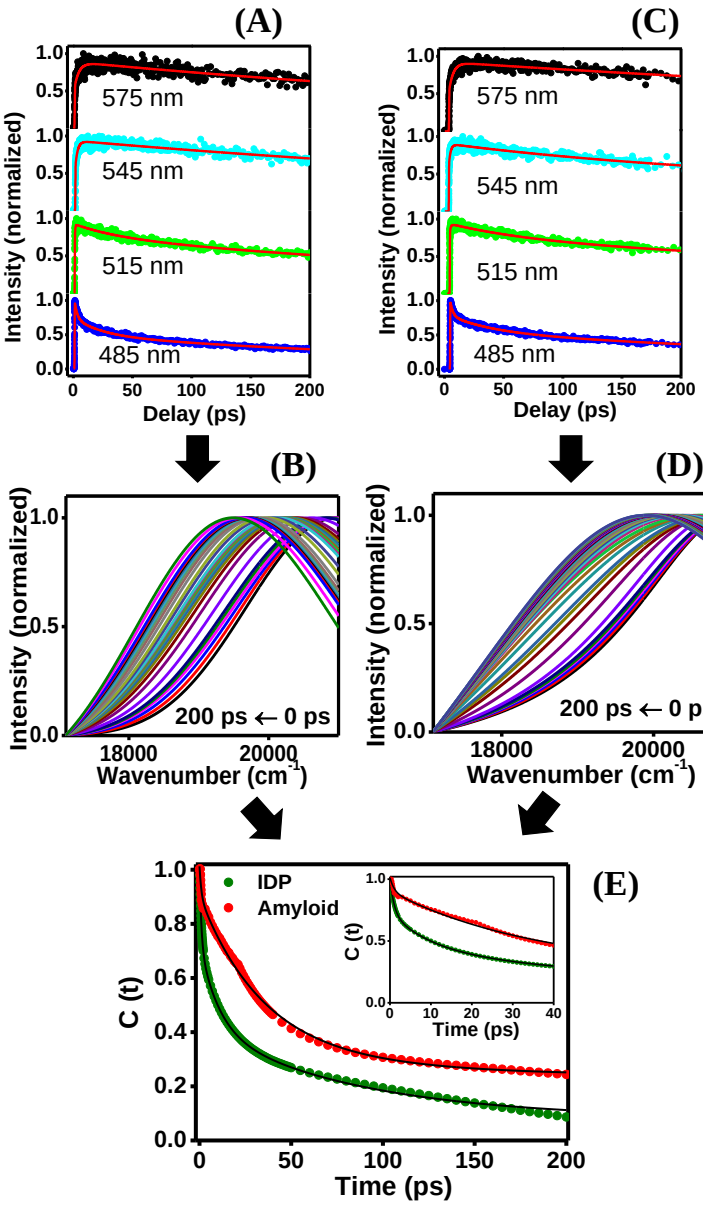


Figure 3.2 (A) Femtosecond transients for acrylodan at four emission wavelengths for IDP monomer: 485 nm (blue), 515 nm (green), 545 nm (cyan), and 575 nm (black). The fits are shown in red. (B) Timeresolved emission spectra (TRES) of IDP monomer. (C) Femtosecond transients for amyloid and (D) TRES for amyloid. (E) Hydration correlation function [C(t)] plots are shown for both monomer (olive) and amyloid (red) from 0 to 200 ps. The fits are shown by black solid lines (inset: 0-40 ps). Recovered solvation times are given in Table 3.2. See Experimental Section for data analysis.

Table 3.2

The parameters recovered upon fitting the solvation correlation functions constructed form femtosecond up-conversion measurements:[§]

Solvation times (amplitudes)	**IDP-state**	**Amyloid-state**
$\Delta\nu_{(0\rightarrow 200\ ps)}$	1433 ± 56 cm^{-1}	1123 ± 21 cm^{-1}
τ_{s1} (α_{s1})	1.0 ± 0.1 ps (0.27 ± 0.01)	1.0 ± 0.2 ps (0.11 ± 0.01)
τ_{s2} (α_{s2})	10.3 ± 0.5 ps (0.31 ± 0.01)	33.7 ± 1.2 ps (0.55 ± 0.05)
τ_{s3} (α_{s3})	85 ± 2 ps (0.34 ± 0.01)	90 ± 7 ps (0.12 ± 0.01)
α_r	0.10 ± 0.02	0.22 ± 0.02

[§]The up-conversion data were fitted till 200 ps as shown in Figure 3.2 $\Delta\nu_{(0\rightarrow 200\ ps)}$ is the estimated TDSS in 200 ps. The recovered solvation times and their respective amplitudes (fractional contributions) are referred to as τ_s and α_s, respectively. The amplitude of the (un-relaxed) residual solvation component is denoted as α_r. The (ultraslow) nanosecond solvation time was estimated from the TCSPC measurements (Figure 3.3B). See Experimental Section for the details of data analysis.

Next, in order to decipher the ultraslow time scale (≫100 ps) of type II water, we carried out TCSPC measurements that are more sensitive for slower solvation dynamics. Construction of the hydration correlation function [C(t)] from the TCSPC data revealed that the time-zero frequency of the amyloid state was closer to that recovered from our femtosecond upconversion experiments (~21 000 cm^{-1}) compared to the IDP state (Figure 3.3B). This reaffirms that the solvation is indeed slower in case of amyloid compared to the monomeric IDP. The longer component recovered for acrylodan in the IDP state of κ-casein was ~1.3 ns corroborating our previous report.[28] The (ultraslow) nanosecond solvation time scale of monomeric IDP state has been assigned to the ordered water network within the collapsed globules of natively unfolded κ-casein.[28] We would like to point out that such (ultraslow) nanosecond solvation dynamics is a characteristic of a compact disordered state and is absent in highly denatured state of proteins.[34] We next asked how this ultraslow solvation component of a compact disordered state changes during the conformational sequestration of the polypeptide chain into β-rich

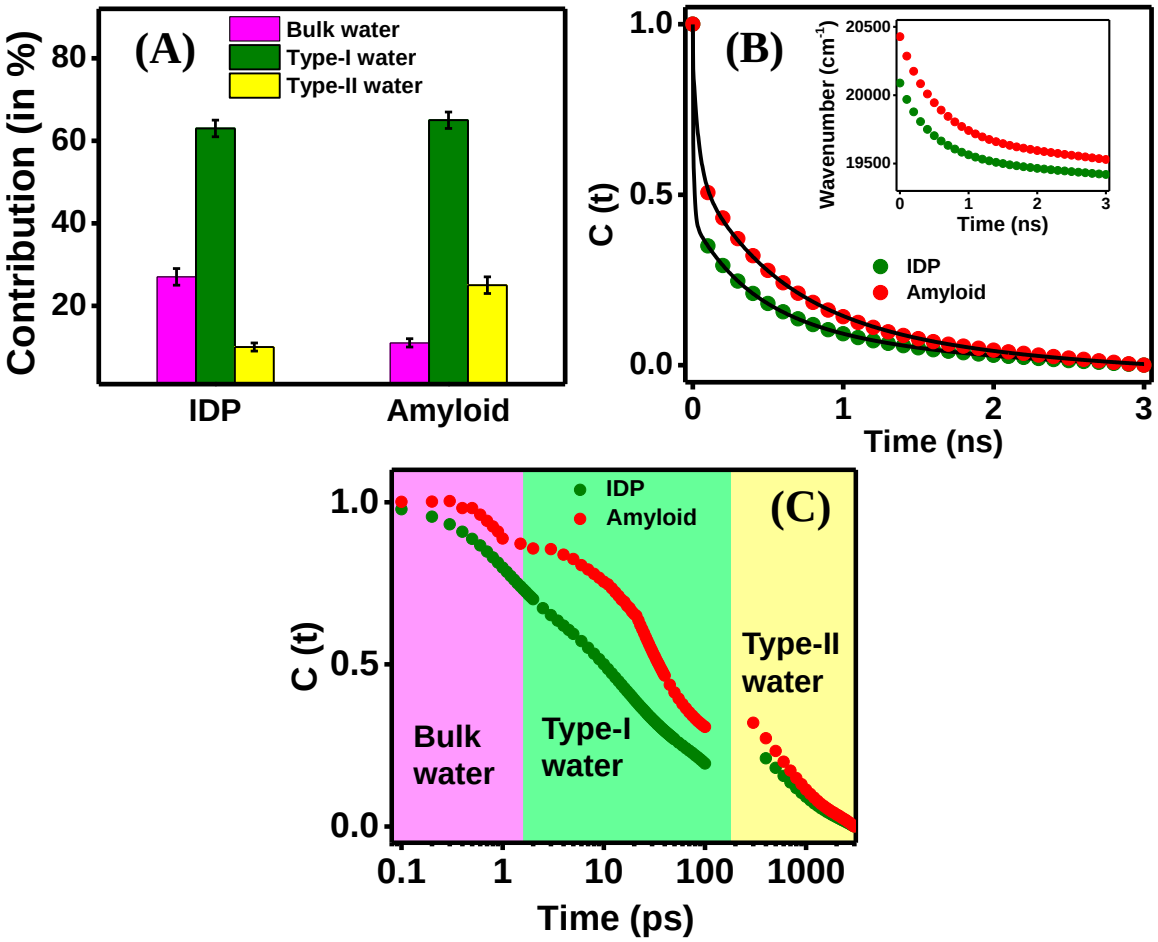


Figure 3.3 (A) Fractional contributions for bulk water (~1 ps), bound water type I (10-100 ps) and type II (unrelaxed residual component ≫100 ps) recovered from analysis of femtosecond data. (B) Hydration correlation function [C(t)] plot constructed from the TCSPC measurements with inset showing the drop in the wavenumber ν(0) recovered from TCSPC for IDP (olive) and amyloid (red). The higher time-zero wavenumber for amyloid state compared to the IDP state is indicative of slower solvation in the amyloid state. The longer solvation components were 1.3 ± 0.2 ns (IDP) and 2 ± 0.2 ns (amyloid). See SI for details of data analysis. (C) Complete hydration correlation plot ranging over four orders of magnitude time scale constructed by combining the femtosecond up-conversion data and the picosecond TCSPC data illustrating the presence of three distinct types of water in IDP (olive) and amyloid (red).

amyloid fibrils. Upon amyloid assembly, the time scale for longer solvation component increased to ~2 ns (Figure 3.3B). This result indicated that upon amyloid formation, the bound water type II becomes even more highly restrained exhibiting over three orders of magnitude slower dynamics than the bulk water. This could be possibly due to the entrapment of water within the interstitial spaces of amyloid architecture which resembles the trapped water pool

found under nanoconfinement such as in reverse micelles and cyclodextrin complexes.[7,38] We would like to mention here that the (ultraslow) nanosecond solvation component can potentially have contributions from the rotational dynamics. In our case, the nanosecond hydration dynamics in amyloid is unlikely to be influenced by the rotational mobility because the fluorescence anisotropy decay experiments suggested a much slower rotational correlation time (>20 ns; Figure 3.1D and Table 3.1). However, we do not completely rule out the possibility of minor contributions from the chain dynamics.

3.3.4 Nanosecond solvation confirmed by red-edge excitation shift (REES)

In addition to nanosecond solvation dynamics of acrylodan, the intrinsic tryptophan fluorescence also confirmed that there is an overall dampening in the (ultraslow) nanosecond solvation dynamics upon amyloid fibril formation. In the amyloid state, the single tryptophan at residue 76 exhibited a higher red-edge excitation shift (REES) that is a qualitative but robust indicator of ordering of surrounding dipolar environment (Figure 3.4). The extent of REES is estimated by monitoring the spectral shift of emission by changing the excitation wavelength at the red-edge of the absorption spectrum and is dependent on the extent of solvent relaxation competing with the fluorescence emission on the nanosecond time scale.[39] Our tryptophan REES data provide additional support in favor of water ordering upon amyloid formation.

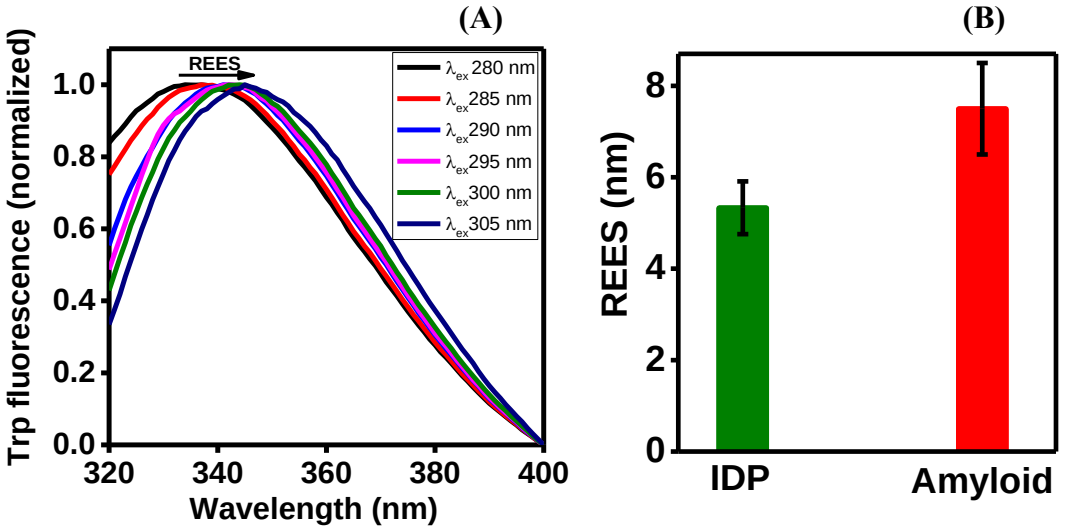


Figure 3.4 Red-edge excitation shift (REES) of Trp 76 in κ-casein amyloid. (A) The normalized fluorescence spectra obtained with different excitation wavelength from 280-305 nm. (B) REES (in nm) obtained using λ_{ex} from 290nm to 305 nm for the monomeric IDP state and the amyloid state. Data are shown as mean ± SEM (n = 3).

3.3.5 Femtosecond-to-nanosecond hydration dynamics reveals the nature of water rearrangements taking place during disorder-to-order amyloid transition

We next combined the hydration correlation functions ranging from femtoseconds to nanoseconds. Figure 3.3C shows the complete hydration correlation function that was constructed by combining both up-conversion and TCSPC data. The complete hydration correlation function depicts the overall dynamics of different types of water in both monomeric and amyloid state on the time scale that ranges across several orders of magnitude from femtoseconds to nanoseconds. Taken together, our results from the femtosecond upconversion and the TCSPC experiments reveal that upon amyloid formation: (i) there is an increase in the reorientation time of both bound water type I and type II suggestive of more ordering of water molecules, (ii) the contribution of bulk water decreases indicating chain desolvation, and (iii) the contribution of bound water type II increases suggesting more ordering. These findings indicate that bulk water and bound water type I and type II can potentially exchange during the entire aggregation cascade. Therefore, our results support the previous hypotheses where the entropy gain from the expulsion of water molecules has been proposed to drive the protein aggregation process.[17,19,40–43] The release of water molecules has also been considered a major driving force in protein crystallization. Additionally, the trapping of water molecules has been observed previously in $A\beta_{1-40}$ amyloid fibrils.[23] During the initial steps of aggregation, mediated by the coalescence of the disordered collapsed globules, the tightly held water molecules get entropically released. As the initial aggregates conformationally mature into β-rich amyloid fibrils, both bulk and surface-bound water molecules get progressively confined within the interstices of the amyloid fibers leading to the formation of hydrogen-bonded wet interfaces, and therefore, the contribution of bound water type II rises.

3.4 Summary

Here, we have monitored the water dynamics in the amyloid state around the Q/N-rich amyloidogenic N-terminal segment of κ-casein, which is an archetypal amyloidogenic IDP. IDPs have captivated enormous interest due to their association with a number of important functions and a variety of diseases that include cancer, cardiovascular diseases, and amyloid-related disorders like Alzheimer's and Parkinson's.[10] Conformational transition of IDPs has important implications in amyloid formation in which water is believed to influence the cascade of aggregation events.[19,20,23,24] We have previously demonstrated that the disordered, yet collapsed, globules of κ-casein contain highly retarded water molecules.[28] In the present work,

using femtosecond and picosecond spectroscopy, we have shown for the first time that different types of water undergo extensive reorganization during the disorder-to-order amyloid transition. The amyloid state exhibits much slower water mobility than the IDP state, which is accompanied by a decrease in the contribution of bulk water with a synchronous increase in the interstitially trapped water. Our novel findings suggest that mapping water dynamics at various segments of an amyloidogenic protein at different stages of amyloid formation will provide key insights into the evolution of water network at the residue-specific resolution. These studies are in progress in our laboratory. Additionally, it has been conjectured that the polymorphic diversity of amyloids and the strain phenomenon of prions can arise due to the differences in the water content of fibrils.[17] Moreover, it has been demonstrated that the magnetic response of water molecules present in the hydration layer can be detected by the magnetic resonance imaging technique.[24] Therefore, we believe that discerning the dynamics and organization of water in amyloids will open less frequently traversed avenues that will potentially lead to new amyloid detection tools as well as novel therapeutic approach to combat amyloid-related devastating human diseases.

3.5 References:

1. Bagchi, B. *Water in Biological and Chemical Processes: From Structure and Dynamics to Function*. Cambridge University Press: **2013**.

2. Ball, P. Water as an Active Constituent in Cell Biology. *Chem. Rev.* **2008,** *108*, 74-108.

3. Chaplin, M. Do We Underestimate the Importance of Water in Cell Biology? *Nat. Rev. Mol. Cell Biol.* **2006,** *7*, 861-866.

4. Bagchi, B. Water Dynamics in the Hydration Layer around Proteins and Micelles. *Chem. Rev.* **2005,** *105*, 3197-3219.

5. Frauenfelder, H.; Chen, G.; Berendzen, J.; Fenimore, P. W.; Jansson, H.; McMahon, B. H.; Stroe, I. R.; Swenson, J.; Young, R. D. A Unified Model of Protein Dynamics. *Proc. Natl. Acad. Sci. U. S. A.* **2009,** *106*, 5129-5134.

6. Pal, S. K.; Zewail, A. H. Dynamics of Water in Biological Recognition. *Chem. Rev.* **2004,** *104*, 2099-2124.

7. Bhattacharyya, K.; Bagchi, B. Slow Dynamics of Constrained Water in Complex Geometries. *J. Phys. Chem. A* **2000,** *104*, 10603-10613.

8. Tipping, K. W.; van Oosten-Hawle, P.; Hewitt, E. W.; Radford, S. E. Amyloid Fibres: Inert End-Stage Aggregates or Key Players in Disease? *Trends Biochem. Sci.* **2015,** *40*, 719-727.

9. Knowles, T. P.; Vendruscolo, M.; Dobson, C. M. The Amyloid State and its Association with Protein Misfolding Diseases. *Nat. Rev. Mol. Cell Biol.* **2014,** *15*, 384-396.

10. Uversky, V. N.; Dave, V.; Iakoucheva, L. M.; Malaney, P.; Metallo, S. J.; Pathak, R. R.; Joerger, A. C. Pathological Unfoldomics of Uncontrolled Chaos: Intrinsically Disordered Proteins and Human Diseases. *Chem. Rev.* **2014,** *114*, 6844-6879.

11. Tycko, R.; Wickner, R. B. Molecular Structures of Amyloid and Prion Fibrils: Consensus versus Controversy. *Acc. Chem. Res.* **2013,** *46*, 1487-1496.

12. Adamcik, J.; Mezzenga, R. Proteins Fibrils from a Polymer Physics Perspective. *Macromolecules* **2012,** *45*, 1137-1150.

13. Lee, J.; Culyba, E. K.; Powers, E. T.; Kelly, J. W. Amyloid-beta Forms Fibrils by Nucleated Conformational Conversion of Oligomers. *Nat. Chem. Biol.* **2011,** *7*, 602-609.

14. Greenwald, J.; Riek, R. Biology of Amyloid: Structure, Function, and Regulation. *Structure* **2010,** *18*, 1244-1260.

15. Nelson, R.; Sawaya, M. R.; Balbirnie, M.; Madsen, A. O.; Riekel, C.; Grothe, R.; Eisenberg, D. Structure of the Cross-[beta] Spine of Amyloid-like Fibrils. *Nature* **2005,** *435*, 773-778.

16. Chong, S. H.; Ham, S. Distinct Role of Hydration Water in Protein Misfolding and Aggregation Revealed by Fluctuating Thermodynamics Analysis. *Acc. Chem. Res.* **2015,** *48*, 956-965.

17. Thirumalai, D.; Reddy, G.; Straub, J. E. Role of Water in Protein Aggregation and Amyloid Polymorphism. *Acc. Chem. Res.* **2012,** *45*, 83-92.

18. Pavlova, A.; Cheng, C. Y.; Kinnebrew, M.; Lew, J.; Dahlquist, F. W.; Han, S. Protein Structural and Surface Water Rearrangement Constitute Major Events in the Earliest Aggregation Stages of Tau. *Proc. Natl. Acad. Sci. U. S. A.* **2016,** *113*, E127-36.

19. Schwierz, N.; Frost, C. V.; Geissler, P. L.; Zacharias, M. Dynamics of Seeded Abeta40-Fibril Growth from Atomistic Molecular Dynamics Simulations: Kinetic Trapping and Reduced Water Mobility in the Locking Step. *J. Am. Chem. Soc.* **2016,** *138*, 527-539.

20. Fichou, Y.; Schiro, G.; Gallat, F. X.; Laguri, C.; Moulin, M.; Combet, J.; Zamponi, M.; Hartlein, M.; Picart, C.; Mossou, E.; Lortat-Jacob, H.; Colletier, J. P.; Tobias, D. J.; Weik, M. Hydration Water Mobility is Enhanced Around Tau Amyloid Fibers. *Proc. Natl. Acad. Sci. U. S. A.* **2015,** *112*, 6365-6370.

21. Conti Nibali, V.; Havenith, M. New Insights into the Role of Water in Biological Function: Studying Solvated Biomolecules using Terahertz Absorption Spectroscopy in Conjunction with Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **2014,** *136*, 12800-12807.

22. Eisenberg, D.; Jucker, M. The Amyloid State of Proteins in Human Diseases. *Cell* **2012,** *148*, 1188-1203.

23. Kim, Y. S.; Liu, L.; Axelsen, P. H.; Hochstrasser, R. M. 2D IR provides Evidence for Mobile Water Molecules in Beta-Amyloid Fibrils. *Proc. Natl. Acad. Sci. U. S. A.* **2009,** *106*, 17751-17756.

24. Despa, F.; Fernandez, A.; Scott, L. R.; Berry, R. S. Hydration Profiles of Amyloidogenic Molecular Structures. *J. Biol. Phys.* **2008,** *34*, 577-590.

25. Pal, S. K.; Peon, J.; Zewail, A. H. Biological Water at the Protein Surface: Dynamical Solvation Probed Directly with Femtosecond Resolution. *Proc. Natl. Acad. Sci. U. S. A.* **2002,** *99*, 1763-1768.

26. Zhong, D.; Pal, S. K.; Zewail, A. H. Biological Water: A Critique. *Chem. Phys. Lett.* **2011,** *503*, 1-11.

27. Qin, Y.; Wang, L.; Zhong, D. Dynamics and Mechanism of Ultrafast Water–protein Interactions. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 8424-8429.

28. Arya, S.; Mukhopadhyay, S. Ordered Water within the Collapsed Globules of an Amyloidogenic Intrinsically Disordered Protein. *J. Phys. Chem. B* **2014,** *118*, 9191-9198.

29. Jain, N.; Bhattacharya, M.; Mukhopadhyay, S. Chain Collapse of an Amyloidogenic Intrinsically Disordered Protein. *Biophys. J.* **2011,** *101*, 1720-1729.

30. Ecroyd, H.; Thorn, David C.; Liu, Y.; Carver, John A. The Dissociated Form of κ-casein is the Precursor to its Amyloid Fibril Formation. *Biochem. J.* **2010,** *429*, 251-260.

31. Sokolovski, M.; Sheynis, T.; Kolusheva, S.; Jelinek, R. Membrane Interactions and Lipid Binding of Casein Oligomers and Early Aggregates. *Biochim. Biophys. Acta* **2008,** *1778*, 2341-2349.

32. Dalal, V.; Arya, S.; Mukhopadhyay, S. Confined Water in Amyloid-Competent Oligomers of the Prion Protein. *Chemphyschem* **2016**, *17*, 2804-2807.

33. Jha, A.; Ishii, K.; Udgaonkar, J. B.; Tahara, T.; Krishnamoorthy, G. Exploration of the Correlation between Solvation Dynamics and Internal Dynamics of a Protein. *Biochemistry* **2011,** *50*, 397-408.

34. Kamal, J. K.; Zhao, L.; Zewail, A. H. Ultrafast Hydration Dynamics in Protein Unfolding: Human Serum Albumin. *Proc. Natl. Acad. Sci. U. S. A.* **2004,** *101*, 13411-13416.

35. Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*. Springer US: **2007**.

36. Maroncelli, M.; Fleming, G. R. *J. Chem. Phys.* **1987,** *86*, 6221.

37. Fee, R. S.; Maroncelli, M. Estimating the Time-Zero Spectrum in Time-Resolved Emission Measurements of Solvation Dynamics. *Chem. Phys.* **1994,** *183*, 235-247.

38. Vajda, S.; Jimenez, R.; Rosenthal, S. J.; Fidler, V.; Fleming, G. R.; Castner, E. W. Femtosecond to Nanosecond Solvation Dynamics in Pure Water and Inside the γ-Cyclodextrin Cavity. *J. Chem. Soc. Faraday Trans.* **1995,** *91*, 867-873.

39. Chattopadhyay, A.; Haldar, S. Dynamic Insight into Protein Structure Utilizing Red Edge Excitation Shift. *Acc. Chem. Res.* **2014**, *47*, 12-19.

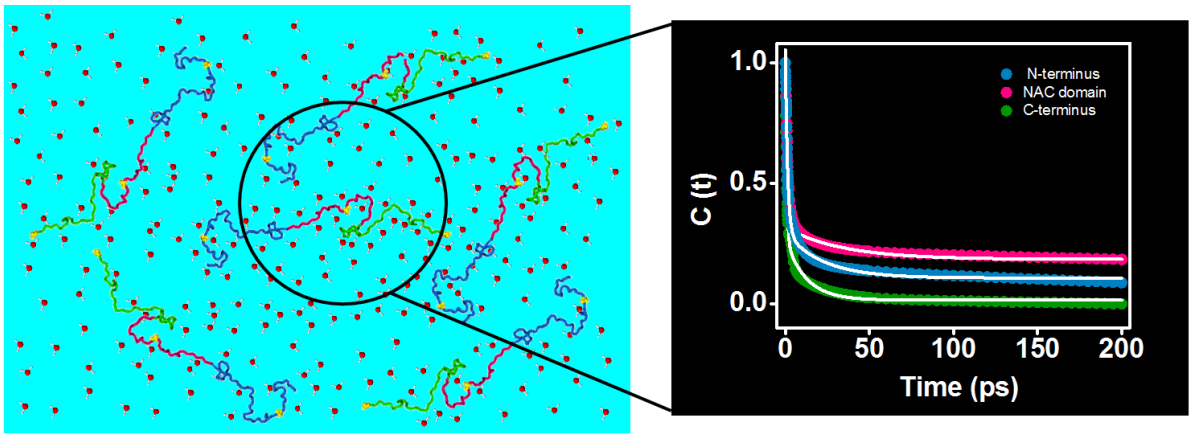
40. Krone, M. G.; Hua, L.; Soto, P.; Zhou, R.; Berne, B.J.; Shea, J. E. Role of Water in Mediating the Assembly of Alzheimer Amyloid-β Aβ16-22 Protofilaments. *J. Am. Chem. Soc.* **2008,** *130*, 11066-11072.

41. Reddy, G.; Straub, J. E.; Thirumalai, D. Dynamics of Locking of Peptides onto Growing Amyloid Fibrils. *Proc. Natl. Acad. Sci. U. S. A.* **2009,** *106,* 11948-11953.

42. Reddy, G.; Straub, J. E.; Thirumalai, D. Dry Amyloid Fibril Assembly in a Yeast Prion Peptide is Mediated by Long-lived Structures Containing Water Wires. *Proc. Natl. Acad. Sci. U. S. A.* **2010,** *107*, 11948-11953.

43. Vekilov, P. G. Phase Transitions of Folded Proteins. *Soft Matter* **2010,** *6*, 5254-5272.

Site-Specific Hydration Dynamics of α-synuclein at the Femtosecond Time Resolution



4.1 Introduction

The coupling between water and polypeptide chain is believed to have a profound influence on protein dynamics, folding, binding, assembly and function.[1-7] Water molecules residing in the hydration layer at the protein-water interface exhibit unique dielectric and dynamical characteristics that are fundamentally different from bulk water.[1] Femtosecond time-resolved spectroscopy offers a highly sensitive and selective method to monitor the reorientation dynamics of water molecules both in bulk phase and at the protein-water interface.[8-9] These ultrafast studies have shown that ordered interfacial water molecules display restrained mobility.[5, 7, 10] These restrained water molecules in the hydration layer of proteins are termed as "biological water" and play a critical role in the folding, enzyme catalysis and protein-DNA and protein-protein interactions. In case of globular proteins, water is proposed to act as a driving force for the hydrophobic collapse that occurs prior to the folding of protein to attain a well-defined 3-dimensional structure.[11] However, unlike in globular proteins, the behavior of water in natively unfolded or intrinsically disordered proteins (IDPs) is poorly understood. IDPs lack the ability to undergo autonomous folding under the native condition and exist as dynamic ensembles of interconverting structures.[12-15] The dynamic nature of IDPs confers them the flexibility to adopt diverse conformations depending upon the binding partner in order to perform multiple functions.[15-18] However, the conformational plasticity of IDPs is occasionally not dealt with appropriately by the cell and results in their aberrant aggregation to form cytotoxic amyloid aggregates,[19] that are implicated in a number of debilitating neurodegenerative diseases.[19-21] One such disease is Parkinson's disease (PD) which is the second most common movement disorder and also the second most common neurodegenerative disorder.[22-23] The pre-synaptic protein, α-synuclein is believed to play a central role in the pathogenesis of PD as the aggregates of α-synuclein are the main component of Lewy bodies, which epitomize the characteristic histological feature of PD.[24-25] Besides PD, the accumulation of α-synuclein in the form of misfolded oligomers and large aggregates is genetically as well as pathologically linked to multiple neurodegenerative diseases, collectively known as "synucleinopathies".[26] The precise cellular function of α-synuclein has remained largely unclear; however, it has been implicated in synaptic plasticity,[27] learning,[28] neurotransmitter release,[29-30] dopamine metabolism,[31] and vesicular trafficking.[32] α-synuclein is an IDP containing 140 amino acid residues and can be broadly divided into three distinct domains based on the amino acid composition: (i) positively charged N-terminal domain [1-60], (ii) hydrophobic and amyloidogenic non-amyloid-beta component (NAC) domain [61-95],

and (iii) negatively charged proline rich C-terminal domain [96-140]. The N-terminal domain is homologous to the lipid binding domain of apolipoproteins and is known to form amphipathic α-helices in the presence of synthetic lipids.[33] The interaction of α-synuclein with membranes is associated with the physiological function as well as misfolding and aggregation.[34-38] The NAC domain being hydrophobic in nature is known to play a crucial role in α-synuclein aggregation.[39-40] This region is also proposed to interact with heat shock proteins.[41] Studies have shown that the NAC domain forms the core of α-synuclein amyloid fibrils and can independently form aggregates that are cytotoxic in nature.[42] Moreover, there was a recent report that emphasized upon the importance of 68-78 residue stretch of the NAC domain in the context of cytotoxicity of α-synuclein aggregates.[43] The C-terminal domain which represents a putative Ca^{2+} binding region[44] plays a key regulatory role in α-synuclein aggregation by acting as a chaperone.[45-46] Thus, all domains of α-synuclein are very well orchestrated and have specific roles in functions and pathology. It is interesting to note that although α-synuclein belongs to the class of expanded IDP devoid of any significant collapse,[47-49] there are transient long-range interactions that make α-synuclein marginally more compact than a typical random-coil in solution.[50-52] Water, in general, is considered a poor solvent for polypeptide backbones and favors chain-chain interactions over chain-solvent interactions.[53] However, the amino acid composition has an extraordinary influence on the final fate of the polypeptide chain.[54] The fractional charge per residue (FCR) and net charge per residue (NCPR) analysis of α-synuclein suggests that it exists as chimera of globules and coils.[54] Plausibly, the intricate balance between solvent-mediated intra-chain attractions versus repulsions underlies the fundamental origin of these transient interactions in case of IDPs and plays a pivotal role in determining the most accessible conformational ensemble to a particular IDP sequence.[53] Since chain-water (solvent) interactions directly influence the structure of a protein and hence the function, we have utilized ultrafast laser spectroscopy to probe the behavior of water molecules within α-synuclein. Although a variety of techniques have emerged for monitoring water dynamics in proteins and other complex biological systems,[55-58] the approach utilized in this study allows us to watch the dynamic motions of water molecules in the monomeric state of the protein with femtosecond time resolution.[9] This timescale is significantly important for the detailed understanding of water present at the protein interfaces as hydrogen bond rearrangements in water occur on this time scale.[1] Additionally, it is important to emphasize here that the inherent tendency of IDPs to undergo intermolecular interactions[59] make them inaccessible to most techniques that offer a high time resolution but

require a very high protein concentration. Thus, in this aspect particularly, the ultrafast laser spectroscopy with high time resolution is quite appealing for probing the important dynamical changes of hydration in IDPs by the use of a highly solvent sensitive fluorophore and has not been explored until now for IDPs. Here, using ultrafast laser spectroscopy, we have been able to delineate the differences in the surface water dynamics within different domains of natively unfolded α-synuclein. The differences in water arrangement are likely to translate distinct roles of different domains in α-synuclein function and disease. In addition, we also observed a significant change in the hydration dynamics around NAC domain upon amyloid formation.

4.2 Experimental Section

4.2.1 Materials

Dithiothretiol (DTT), guanidium chloride, sodium chloride (NaCl) and tris HCl were procured from Sigma (St. Louis, MO) and used without any further modification. 6-Acryloyl-2-(dimethylamino) naphthalene (acrylodan) was purchased from Molecular Probes, Invitrogen Inc. All solutions were prepared using Milli Q water and pH of the buffers was adjusted using 1 N HCl or NaOH on Metrohm (Herisau, Switzerland) 827 pH meter at ~25°C.

4.2.2 Expression, purification and labeling of cysteine mutants of α-synuclein

The plasmid construct used for the recombinant expression of α-synuclein protein in *E.coli* BL21 (DE3) pLys S strain was same as reported previously. The cysteine mutations at 9, 18, 56, 78, 90 and 140 positions were created using QuikChange site-directed mutagenesis kit (Stratagene) and for all the mutations except for 9, alanine was changed to cysteine. For position 9, serine was mutated to cysteine. The protocol for expression and purification of the mutants was same as wild-type protein and has been reported previously.[60] The only difference was in the buffer used for dialysis which in this case was pH 7.4, 20 mM Tris buffer containing 1 mM DTT and 50 mM NaCl. The labeling of the cysteine mutants was carried out with a 5 M excess of acrylodan under denatured condition (6M Guanidinium chloride in pH 7.4, 20 mM Tris buffer containing 1 mM DTT and 50 mM NaCl). The labeled proteins were then passed through PD-10 column to remove excess dye. For all the experiments, the buffer exchange with the native buffer (pH 7.4, 20 mM Tris buffer containing 1mM DTT and 50 mM NaCl) was done on PD-10 column immediately before the experiments. The concentration of the labeled protein was estimated using ε_{365} = 12800 M^{-1} cm^{-1}.[61] The total protein concentration was estimated using ε_{275} = 5600 M^{-1} cm^{-1} [60] and ε_{365} = 12800 M^{-1} cm^{-1} , taking correction factor of acrylodan into account. The labeling efficiency was >85% for all the mutant proteins. The

concentration of the labeled proteins used for all our measurements except for circular dichroism was ~40 μM.

4.2.3 Circular dichroism (CD) spectroscopy

The CD spectra were recorded in Chirascan spectrometer (Applied Photophysics, UK) at room temperature using 1 mm path length quartz cuvette. The concentration of the unlabeled as well as labeled proteins used was 25 μM. The spectra obtained were corrected for buffer and were smoothened using Pro Data software. The spectra were finally plotted in Origin Pro 8.5 software.

4.2.4 Steady-state fluorescence measurements

The steady-state fluorescence measurements were carried out in Fluoromax-4 spectrometer (Horiba Jobin Yvon, NJ). The samples were excited at 400 nm wavelength and the spectra were collected from 440-600 nm. The steady-state fluorescence anisotropies were measured at 525 nm. The steady-state fluorescence anisotropy (r_{ss}) is given by the following equation:

$$r_{ss} = (I_{\parallel} - I_{\perp}G) / (I_{\parallel} + 2I_{\perp}G) \qquad (1)$$

where $I_{\parallel}$ and $I_{\perp}$ are fluorescence intensities collected using parallel and perpendicular geometry, respectively. The perpendicular components were always corrected using a G-factor.

4.2.5 Femtosecond fluorescence up-conversion measurements

A second harmonic light of mode locked 100 fs Ti:Sapphire laser (Tsunami, Spectra Physics, USA) centered at 800 (1 W, 80 MHz repetition rate), was used as excitation source for fluorescence upconversion measurements (FOG 100, CDP, Russia). For generating 800 nm light, the self-mode-locked Ti:Sapphire laser was pumped by continuous wave second harmonic light of the output of 4.5 W Nd:YVO4 laser (Millenia Pro, Spectra Physics, USA). The Nd:YVO4 laser was further pumped by an electrically pumped Diode laser. Finally, around 25 mW of 380 nm light was used to excite the sample, which was housed in a rotating cell. The excitation light polarization was set at magic angle (54.7°) with respect to the gate light. A set of Glan polarizer and Berek's variable waveplate was used for setting up the magic angle polarization. The emission light was upconverted using 800 nm gate light in 1 mm thick β-BBO (type I) crystal. The resulting upconverted light was passed through a UV bandpass filter (which cuts excitation, emission and gate light) and then through a double monochromator (DM2100, CDP, Russia). Finally, a photon counter PMT (having dark noise <5 cps) was used

for signal counting. The decays were collected from 450-560 nm, after every 10 nm till 300 ps. An appropriate integration time and step size was kept so that enough reliable data can be collected without much photodegradation. We carried out similar measurements using 400 nm as excitation wavelength and 475-585 nm emission scan range in order to monitor the effect of excitation wavelength on the solvation. For this, we chose three mutants: Cys-18, Cys-78 and Cys-140. The upconversion transients and time-resolved emission spectra (TRES), the individual correlation function plots and the overlayed correlation function plots (see below) corresponding to 400 nm excitation are shown in Figure 4.2, 4.3 and F4.5A, respectively. The instrument response function (IRF) was measured by measuring the cross correlation between the gate light and Raman signal of buffer. The resulting profile when fitted with a Gaussian equation (Origin Pro 8.5) gave a full width at half maxima (FWHM) around 254 fs. The emission decay traces were fitted by means of reconvolution in a desired number of exponential in a homebuilt program in Igor Pro 6.37 software.[62]

4.2.6 Time-correlated single photon counting (TCSPC) measurements

For TRES: The fluorescence lifetime data were collected at 54.7° (magic angle) polarization using time-correlated single photon counting (TCSPC) module (Fluorocube, Horiba Jobin Yvon, NJ). A peak count of 10,000 was used. The laser diode with 375 nm as excitation wavelength was used for acrylodan. For collecting the instrument response function (IRF), the colloidal silica (Ludox) solution was used. The IRF had a full width at half maxima (FWHM) at ~ 250 ps for 375 nm excitation source. The emission wavelength scan range used was 450-560 nm and the decay profiles were collected after every 10 nm. The decays were then analyzed and fitted taking into account the IRF using DAS6 software provided with the TCSPC system.

For fluorescence anisotropy decay: For time-resolved fluorescence anisotropy measurements, the peak difference was 10,000 counts and the orientation of emission polarizer was 0° and 90° with respect to the excitation polarizer for parallel fluorescence intensities ($I_{\parallel}$) and perpendicular fluorescence intensities ($I_{\perp}$), respectively. The emission wavelength for these measurements was fixed at 525 nm with a bandpass of 4 nm and 375 nm laser diode was used as excitation source for acrylodan. The width (FWHM) of instrument response function (IRF) was collected using Ludox (colloidal silica) and was found to be ~250 ps. The anisotropy decays were analyzed by globally fitting $I_{\parallel}(t)$ and $I_{\perp}(t)$, as described previously.[63]

4.2.7 Construction of TRES from femtosecond-resolved (fluorescence upconversion) and picosecond-resolved (TCSPC) data

The recovered fluorescence lifetime parameters were used to construct the fluorescence decays at different emission wavelengths in Origin Pro 8.5 software using $I_\lambda(t) = \sum_i \alpha_i \exp(-t/\tau_i)$, where α_i and τ_i represent the contributions and lifetime of the different lifetime components, respectively. The time-resolved emission spectra (TRES) were then constructed using:

$$I(\lambda, t) = I_\lambda^{SS} I_\lambda(t) / \sum_i \alpha_i \tau_i \qquad (2)$$

where I_λ^{SS} represents the steady-state fluorescence intensity at a fixed emission wavelength λ and $I_\lambda(t)$ is same as mentioned above. All the spectra (obtained at different time points) were fitted using log-normal function to extract the peak frequency ν(t) as a function of time and were then normalized to account for the overall extent of frequency shift. The solvation correlation function is given by:

$$C(t) = (\nu(t) - \nu(\infty)) / (\nu(0) - \nu(\infty)) \qquad (3)$$

where ν(0), ν(t) and ν(∞) are the emission peak frequencies at time 0, t and ∞, respectively. The recovered ν(0) from our measurements was ~21300 ± 150 cm^{-1}, which is close to ν(0) reported previously.[64] The recovered ν(2 ns) (from TCSPC TRES measurements) for all the mutants was similar (~19000 cm^{-1}) and ν(200 ps) for Cys-140 were close to the steady-state emission maximum. Therefore, we used 19000 cm^{-1} as ν(∞). The correlation function was then fitted using a sum of exponentials:

$$C(t) = \sum_i \beta_i \exp(-t/\tau_{si}) \qquad (4)$$

where β_i and τ_{si} represents the contributions and solvation time of the different solvation components, respectively.

4.2.8 Aggregation reaction

For setting up aggregation reaction, 50 μM of acrylodan labeled was added to 150 μM of wild-type α-synuclein in pH 7.4, 20 mM Tris buffer containing 100 mM NaCl. The temperature of the reaction was maintained at 37°C and the reaction was continuously stirred at 800 rpm. The reaction was monitored by scattering and atomic force microscopy (AFM). At the time when fibrils were observed, the steady-state fluorescence measurements were carried out. The parameters for acrylodan fluorescence were same as mentioned above.

4.2.9 Atomic force microscopy (AFM) imaging of α-synuclein fibrils

For AFM imaging experiment, the aliquot from the aggregation reaction mixture (10 µl) was taken out and deposited on the freshly cleaved mica surface (Grade V-4 from SPI, PA) and allowed to incubate for 2 minutes. The mica surface was then washed thrice with filtered Milli Q water and dried under a gentle stream of nitrogen for about 20 minutes before it was scanned on an Innova AFM instrument (Bruker corporation) operating in the tapping mode (Scan rate: 0.5Hz, Pixels:512). The image was processed in WSxM software (Nanonics Imaging Ltd, Jerusalem, Israel).[65]

4.3 Results and Discussion

Taking into account the significance of all three domains in α-synuclein function and disease, and advantage of the fact that α-synuclein is devoid of cysteines, we created six single cysteine mutations: Cys-9, 18 and 56, Cys-78 and 90, and Cys-140 in N-terminus, NAC domain and C-terminus, respectively (Figure 4.1A). The cysteines were covalently modified with acrylodan which is considered a suitable fluorophore for hydration dynamics studies due to its high sensitivity towards the polarity of the medium and a large Stokes shift.[66] In order to ensure that the cysteine mutation did not induce any structure in disordered monomeric α-synuclein, we monitored circular dichroism (CD) signature for all the mutants. The CD spectra for all the mutants overlapped and showed signature of a typical disordered monomeric protein akin to the wild type α-synuclein (Figure 4.1B). In order to probe ultrafast hydration, it is extremely important to first confirm that the fluorophore is sufficiently exposed to water. Therefore, as a prelude to the ultrafast hydration studies, we carried out steady-state fluorescence experiments with acrylodan labeled mutants.

4.3.1 Site-specific fluorescence reveals that the cysteine residues in natively unfolded α-synuclein are highly solvent exposed

The emission maximum for all acrylodan labeled mutants was found to be ~525 nm, which clearly indicates that acrylodan is completely solvent exposed (Figure 4.1C). The steady-state fluorescence anisotropy was also comparable for all the mutants but slightly lower for Cys-140 due to it being the last residue of the polypeptide chain (C-terminus) (Figure 4.1D). Thus, the steady-state fluorescence attributes of acrylodan labeled mutants suggested that the microenvironment is similar for all the probed residues in the monomeric state of α-synuclein. The solvent exposure of acrylodan is not very surprising as the surface of the disordered proteins, especially for the expanded IDPs, is expected to be significantly hydrated by the

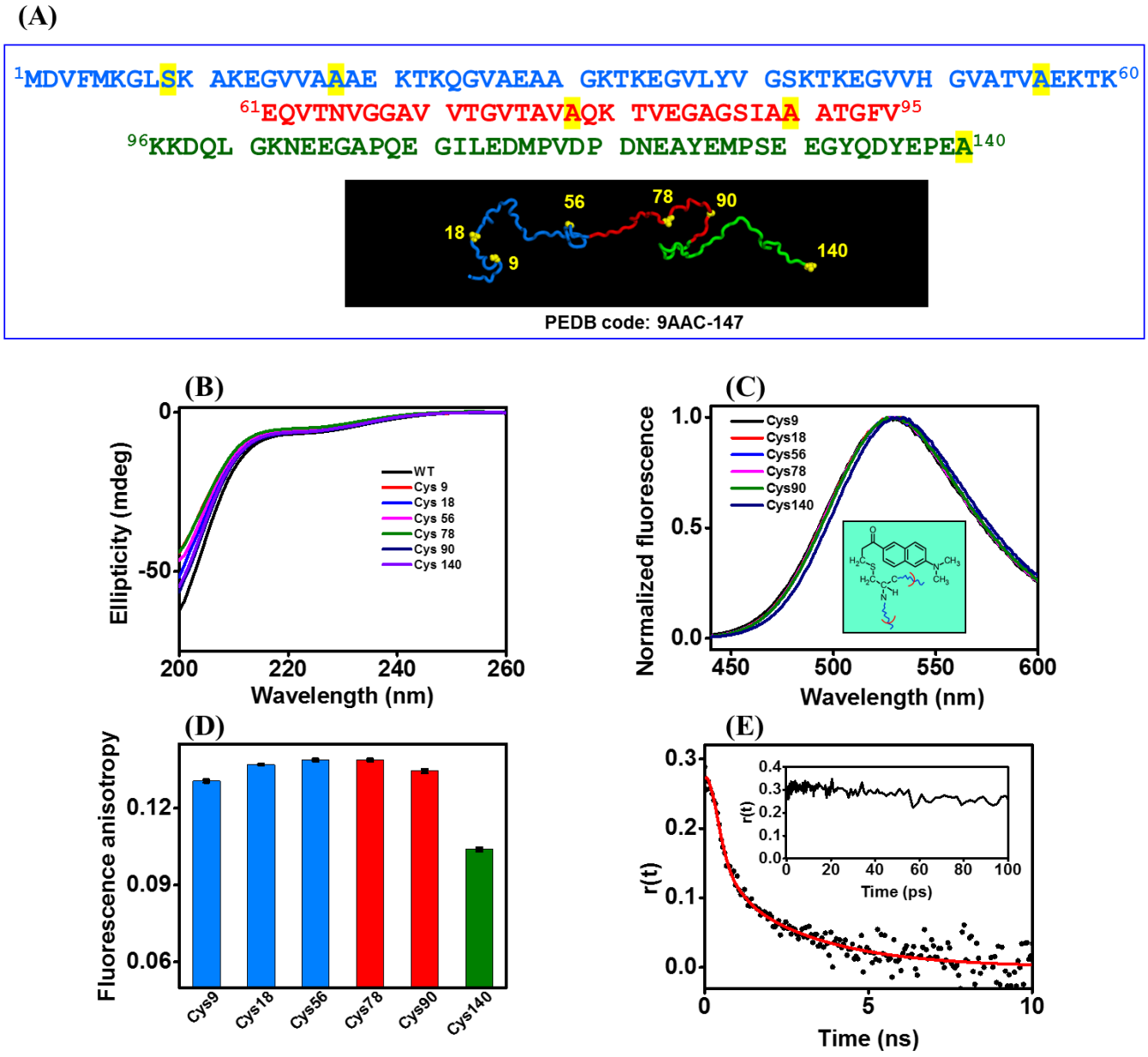


Figure 4.1 (A) The amino acid sequence of α-synuclein along with one of the NMR structures taken from the protein ensemble database (PeDB: 9AAC, conformation no. 147) generated using PyMOL are shown. The blue color corresponds to the N-terminus, red to the NAC domain and olive to the C-terminus. The single cysteine mutation positions are highlighted in yellow. (B) CD spectra for all cysteine variants of monomeric α-synuclein. (C) The steady-state fluorescence spectra and (D) fluorescence anisotropy for all acrylodan labeled cysteine positions (highlighted in Figure 4.1A) in the monomeric state of α-synuclein are shown. (E) Fluorescence anisotropy decay of acrylodan labeled at 78 position (Cys-78) of α-synuclein (fit shown in red) (Recovered parameters: ϕ_{fast} ~ 0.8 ns, β_{fast} ~ 0.6 and ϕ_{slow} ~ 2.3 ns, β_{slow} ~ 0.4) with the inset showing its femtosecond time-resolved anisotropy decay from 0-100 ps.

solvent molecules. Next, we performed fluorescence depolarization measurements in the monomeric state of α-synuclein using time correlated single photon counting (TCSPC) to monitor the rotational dynamics of the fluorophore. We obtained a typical bi-exponential fluorescence anisotropy decay that indicated the (fast) local motion of the fluorophore attached to the protein and the (slow) global tumbling of the entire protein (Figure 4.1E). A significant amplitude (~ 60%) of local motion was observed which suggests that acrylodan is flexible and is certainly exposed to water. We also carried out the femtosecond time-resolved anisotropy measurements (Figure 4.1E inset) that revealed no significant rotational mobility on the sub-100 ps timescale (see below).

4.3.2 Residue-specific time-dependent fluorescence Stokes shift for different domains of α-synuclein

After confirming that all the mutants are appreciably water exposed, we carried out ultrafast hydration dynamics studies using femtosecond fluorescence up-conversion technique in order to probe the water structure in different domains of α-synuclein. Using this technique, we can probe the dielectric response of proteins by monitoring the time-dependent fluorescence Stokes shift (TDFSS) of the fluorophore which is either part of the protein or is covalently linked to it.[67] Immediately upon excitation, the non-equilibrated hydration water network around the fluorophore relaxes *via* hydrogen bond rearrangement to a new configuration. The reorganization of hydration water network around the excited state of fluorophore results in an emission shift to longer wavelength (lower energy) and this shift in the emission spectrum during the process of reorganization is known as TDFSS. The extent of the shift within a particular experimental time window is an indicator of how much solvent relaxed the fluorophore is with respect to the initial unrelaxed state which in turn is dependent on how water molecules are organized and how fast/slow hydrogen bond rearrangements take place around a fluorophore. In order to estimate TDFSS and the hydration timescales, we constructed time-resolved emission spectra (TRES) and solvation correlation function (C (t)), respectively using the femtosecond-resolved fluorescence up-conversion transients collected at different wavelengths that traversed the whole spectral range of the acrylodan fluorescence (For details, refer to Experimental Section). The representative transients at four different wavelengths and TRES (with arrow indicating the total TDFSS) for three mutants, Cys-18 (N-terminus), Cys-78 (NAC domain) and Cys-140 (C-terminus) are shown in (Figure 4.2). The signal was found to decay on the blue edge and rise on the red edge, which is an archetypal feature of solvent relaxation [7, 10, 68-70]. The TDFSS within 200 ps was found to be ~2200 cm^{-1}, ~2100 cm^{-1} and

~2300 cm^{-1} for Cys-9,18 and 56 (N-terminus), Cys-78 and 90 (NAC domain), and Cys-140 (C-terminus), respectively. The C(t) for Cys-18, 78 and 140 are shown in (Figure 4.3). The time-zero frequency (ν(0)) recovered from our femtosecond experiments for all the mutants was

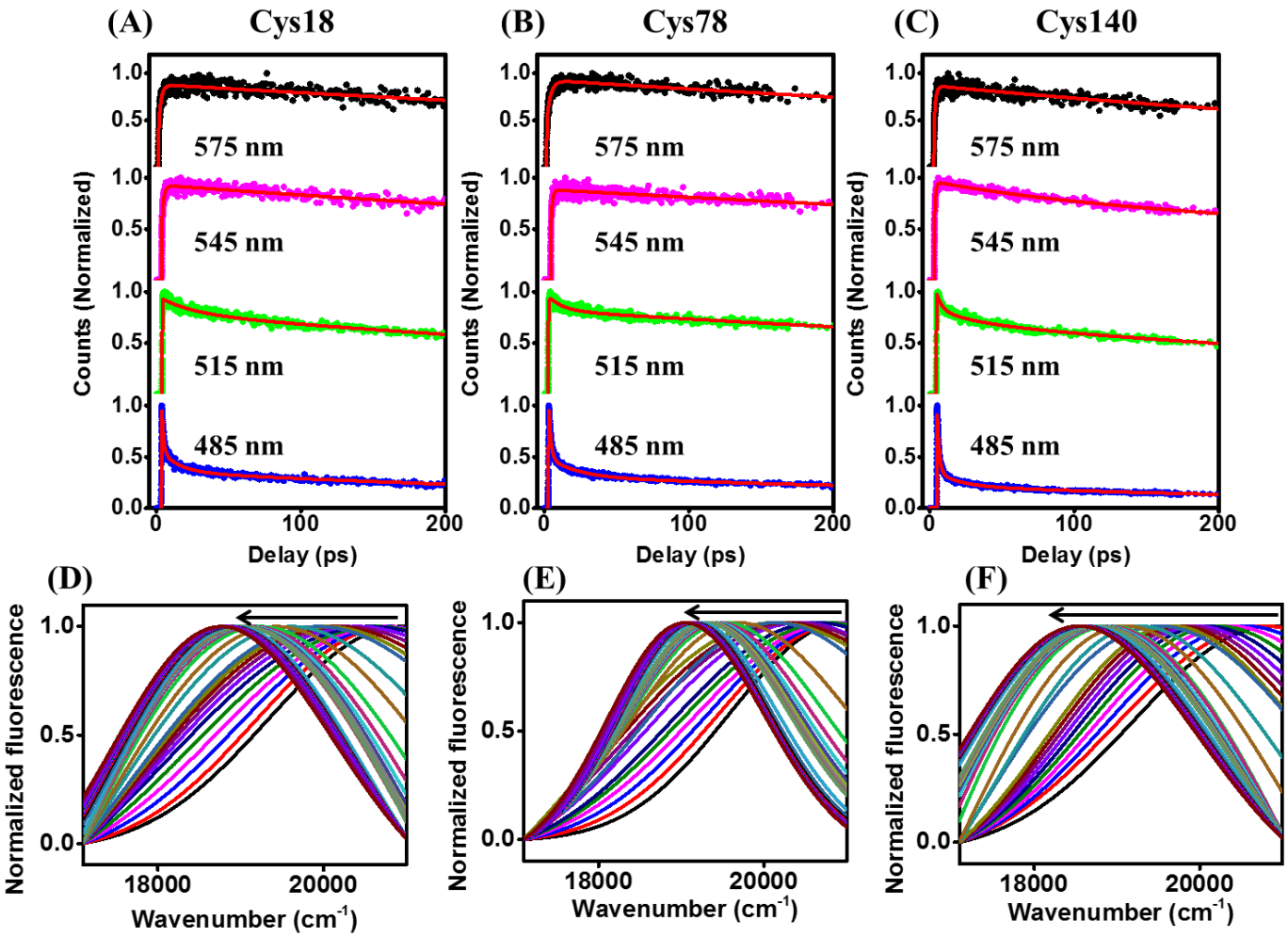


Figure 4.2 The femtosecond up-conversion transients (with fits shown in red) at four emission wavelengths: 485 nm (blue), 515 nm (green), 545 nm (pink) and 575 nm (black), and the time-resolved emission spectra (TRES) with arrow indicating the time-dependent fluorescence Stokes shift (TDFSS) from 0 to 200 ps are shown for acrylodan labeled (A and D) Cys-18, (B and E) Cys-78 and (C and F) Cys-140 α-synuclein.

close to ν(0) estimated previously,[64] using a method described for theoretical estimation of time-zero spectrum.[71] It is interesting to note that the extent of TDFSS observed for all acrylodan labeled mutants of α-synuclein in general and Cys-140 in particular, is close to the typical shift observed for acrylodan in bulk water and is significantly different from the shift observed for acrylodan present in relatively hydrophobic environment.[10] However, the

timescales of hydration dynamics (discussed below) recovered are appreciably different from the bulk and are similar to the timescales reported for proteins.[7, 10]

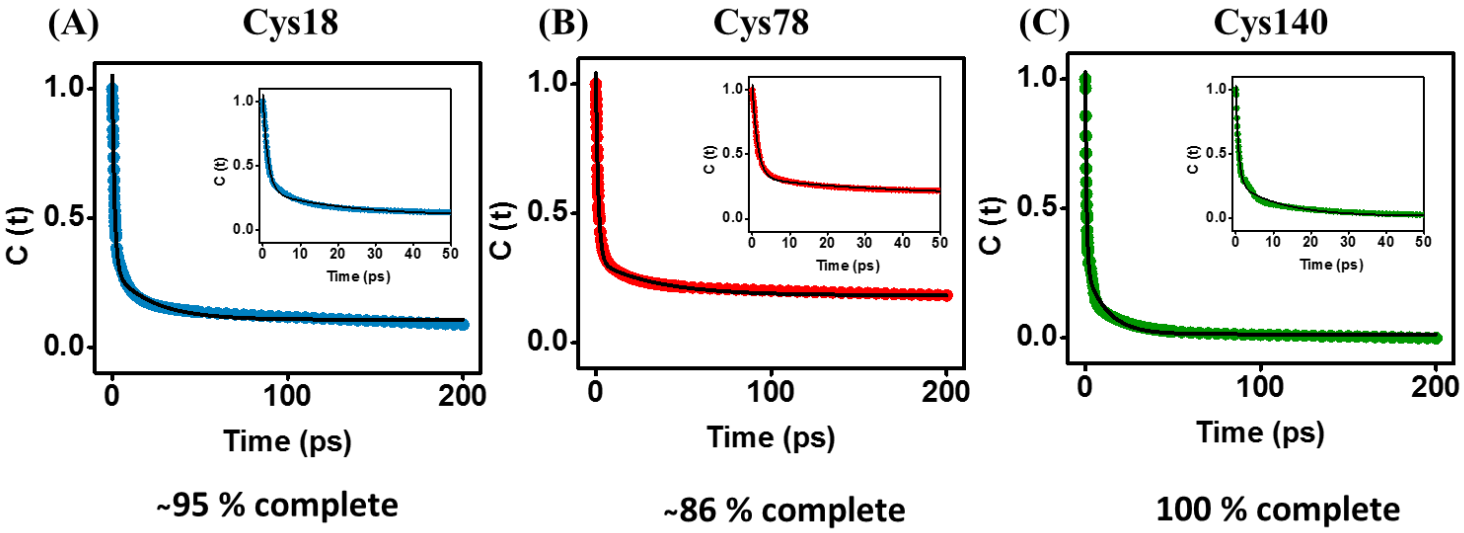


Figure 4.3 The hydration correlation function [C(t)] plots (from 0 to 200 ps) are shown for acrylodan labeled (A) Cys-18 (blue), (B) Cys-78 (red) and (C) Cys-140 (olive) α-synuclein. The fits are shown by black solid lines (inset: 0-50 ps). Recovered solvation times are ~ 1 ps and 22 ps for Cys-18, ~ 1 ps and 34 ps for Cys-78, and ~ 1 ps and 12 ps for Cys-140. The residual solvation amplitudes were ~ 5 % and 15 % for Cys-18 and Cys-78, respectively for the data shown here. See Table 4.1 for the recovered parameters along with error from the analyses of two independent datasets for all the acrylodan labeled single cysteine positions.

4.3.3 Resolving the contributions of bulk- and bound water

Given the extremely sensitive nature of femtosecond fluorescence up-conversion, we were able to clearly separate the bulk water and surface bound water dynamics in monomeric α-synuclein, around the chosen loci. In case of Cys-9, 18 and 56 (N-terminus), the solvation with respect to ν(∞) was ~95% complete within 200 ps. In case of NAC domain, the solvation was found to be ~86% and ~90% complete for Cys-78 and Cys-90, respectively within 200 ps. Unlike N-terminus and NAC domain, the solvation in case of Cys-140 (C-terminus) was almost 100% complete within 200 ps. The recovered solvation parameters for all the mutants are given in Table 4.1 and are shown in Figure 4.5 for comparison. Several theoretical and experimental investigations have revealed that the solvation time for bulk water relaxation spans from 100 femtoseconds to picoseconds.[68] In contrast, water molecules present in the proximity of the protein surface, also known as "biological water", has considerably slower relaxation time

ranging between 10 and 100 ps.[68, 72] Thus, the ultrafast component (~1 ps) represents the bulk water and the slower component (of tens of picosecond), which we refer to as "bound water type-I", can be assigned to the biological water. We refer to the unrelaxed component of solvation as "bound water type-II". It is important to emphasize here that the rotational dynamics of the fluorophore did not contribute to the solvation, since the femtosecond time-

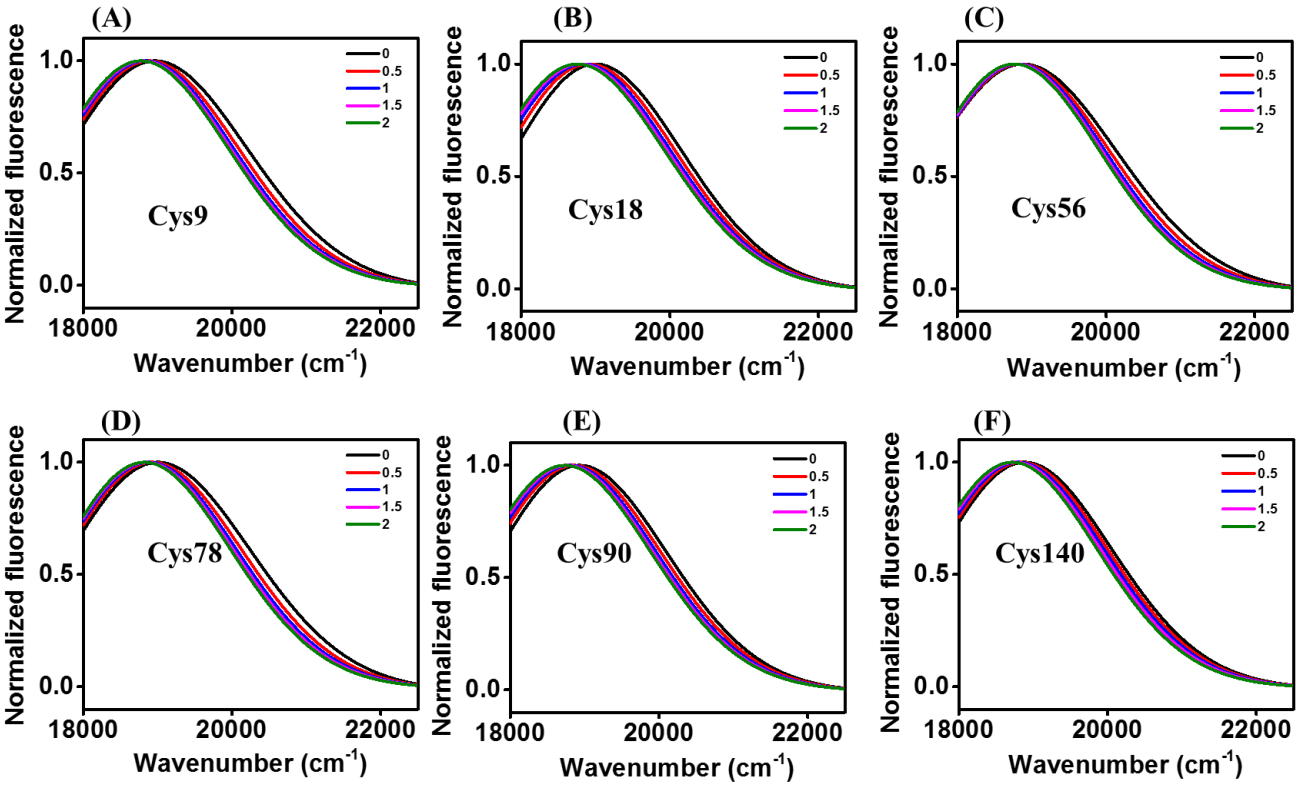


Figure 4.4 TRES at 5 different time points (in ns) for all acrylodan labeled mutants: (A) Cys-9, (B) Cys-18, (C) Cys-56, (D) Cys-78, (E) Cys-90 and (F) Cys-140, are shown.

resolved anisotropy decay did not exhibit any significant drop within 100 ps (Figure 4.1E inset). Further, in order to monitor ps-ns hydration dynamics (if any) in these mutants, we carried out TRES measurements using time-correlated single photon counting (TCSPC) technique as reported previously.[61] The TRES profiles for all the mutants were indistinguishable and did not exhibit significant TDFSS (Figure 4.4). The insignificant (100-150 cm^{-1}) TDFSS from TCSPC TRES suggests that there is very little/no nanosecond solvation component. It is important to emphasize here that the differences in the water structure within α-synuclein are inaccessible by picosecond/nanosecond measurements and can only be resolved using high resolution spectroscopy. Hence, it is clear that even though α-synuclein is largely disordered; there might be some local structural differences due to variation in the

arrangement of bound water within N-terminus, NAC domain and C-terminus because the solvation component of tens of picoseconds is attributed to the hydration water network in the inner layers close to the protein surface and is likely to drive the local side chain motions on the ps timescale or influence the local structure.[73]

Table 4.1

The parameters recovered upon fitting the solvation correlation functions constructed form femtosecond up-conversion measurements:[§]

Position	τ_{s1} (ps) (Bulk water)	α_{s1} (Bulk water)	τ_{s2} (ps) (Bound water type-I)	α_{s2} (Bound water type-I)	α_r, r: Residual solvation (Bound water type-II/ Unrelaxed component)
9	1 ± 0.1	0.74 ± 0.02	19 ± 2	0.20 ± 0.02	0.06 ± 0.01
18	1.4 ± 0.2	0.74 ± 0.02	18 ± 2	0.20 ± 0.01	0.06 ± 0.01
56	1.4 ± 0.1	0.75 ± 0.04	15 ± 2	0.19 ± 0.03	0.07 ± 0.01
78	1.8 ± 0.1	0.77 ± 0.02	31 ± 4	0.10 ± 0.02	0.12 ± 0.02
90	1.6 ± 0.1	0.76 ± 0.02	21 ± 2	0.15 ± 0.02	0.09 ± 0.02
140	1.1 ± 0.1	0.79 ± 0.03	11 ± 1	0.20 ± 0.02	0.03 ± 0.01

[§]The up-conversion data were fitted till 200 ps as shown in Figure 4.3. The recovered solvation times and their respective amplitudes (fractional contributions) are referred to as τ_s and α_s, respectively. The amplitude of the (un-relaxed) residual solvation component is denoted as α_r. See Experimental Section for the details of data analysis.

4.3.4 The presence of restrained water molecules in the NAC domain

Taken together, the above set of data demonstrates that: (i) albeit the extent of steady-state Stokes shift is same for all the mutants (Figure 4.1C), the timescales of bound water type-I relaxation are markedly different (Figure 4.5B). It is important to highlight here that the timescales of hydration are not essentially manifested in the magnitude of steady-state Stokes shift.[7] (ii) The contribution of the bulk water is similar for all the mutants (~75%) (Figure 4.5C) which signifies that the ultrafast component dominates the slower component of hydration for all the mutants of α-synuclein. This observation is in agreement with a previous report on the unfolded state of a globular protein where the bulk water contribution was found to be ~75%.[10] (iii) The bound water type-I relaxation is fastest for Cys-140 (C-terminus) and is slowest for

Cys-78 (NAC domain). Additionally, the contribution of the unrelaxed component of hydration is most for Cys-78 (~15%) and least for Cys-140 (<5%). (iv) Cys-9, 18 and 56 of N-terminus exhibit similar bound water type-I relaxation dynamics and the dynamics is slower than Cys-140 but faster than Cys-78 and 90 of the NAC domain. (v) Cys-90, which is present towards the C-terminal end of NAC domain, exhibits faster bound water type-I relaxation dynamics than Cys-78.

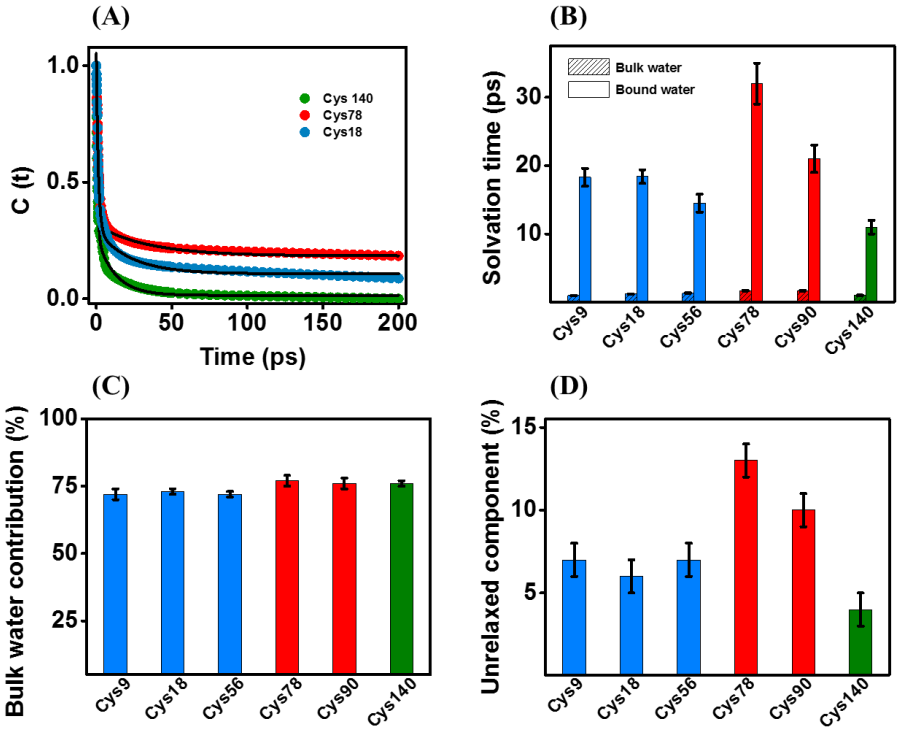


Figure 4.5 (A) The [C(t)] plots (from 0 to 200 ps) for acrylodan labeled Cys-18, 78 and 140 α-synuclein are shown for comparison. (B) The solvation times recovered from analysis of femtosecond data and the contributions from (C) the bulk water and (D) the unrelaxed component (bound water type-II) are shown for all the acrylodan labeled single cysteine positions of α-synuclein.

The contributions from the unrelaxed component for all the mutants are shown in Figure 4.5D.

Collectively, these observations propose that the contribution and the timescale of bulk water is similar for all the mutants. However, bound water type-I around Cys-78 (NAC domain) is associated with a longer residence time and hence these water molecules are more ordered compared to water molecules present in N- and C-terminus. Overall, the bound water type-I relaxation time follows this order: Cys-78 (NAC) > Cys-90 (NAC)> Cys-9 (N) ≈ Cys-18 (N) ≈ Cys-56 (N) > Cys-140 (C).

4.3.5 Variation in the timescale of bound water dynamics within α-synuclein

The similar contribution and the timescale of bulk water relaxation for all the mutants of α-synuclein suggests that the polypeptide chain is homogeneously solvated and is devoid of any prominent secondary structure in the native state. However, the discrepancies in the timescale of bound water type-I relaxation are indicative of some unique differences within different domains.

In order to interpret the slowest bound water relaxation around Cys-78 (NAC domain), it is important to appreciate that within α-synuclein, the NAC domain is the most hydrophobic domain (~66% of the total amino acid residues present in the NAC domain are hydrophobic in nature). Though a large number of models have been proposed to explain the hydrophobic interactions, none of them can provide a complete account for the hydrophobic hydration. One of the early models that were proposed and is very well-known is "iceberg" model of hydrophobic hydration, proposed by Franks and Evans.[74] A further explanation based on the entropic force of attraction that builds upon "iceberg" model was provided by Kauzmann.[75] According to Kauzmann, the accommodation of hydrophobic species would require the disruption of the integrity of the hydrogen bonded network of water and this enthalpic penalty that would be imposed upon by the loss of hydrogen bonds can be compensated by the creation of ordered structure of water around hydrophobes. The creation of such an ordered structure results in a decrease in the entropy which in turn causes the dampening of the rotational and translational motion of water. However, when two such hydrophobes enslaved by ordered water molecules come close to each other, there occurs an entropically favorable release of the water molecules to the bulk milieu. Although Kauzmann accepted the iceberg hypothesis, he anticipated the crystallinity of the structures of the Frank-Evans icebergs to be less perfect than what was proposed in the original iceberg model. Based on a number of studies that followed, it is now believed that the hydration layer around hydrophobes is not made up of clathrate water, rather, it is a dynamic shell formed by van der Waals attraction.[11]

The above discussion elucidates why the bound water relaxation dynamics gets dampened around Cys-78 residue, which is present in the middle of the most hydrophobic NAC domain of α-synuclein. In fact, the retarded bound water dynamics around Cys-78 can possibly account for the reduced intramolecular diffusion for [66-90] segment of the NAC domain, reported previously.[76] In addition to hydrophobic effect, it is interesting to note that Cys-78 is surrounded by polar residues like threonine and glutamine that can participate in extensive

hydrogen bonding which can further dampen the motion of water molecules around it. The presence of ordered (motionally restrained) water molecules can account for the amyloidogenic nature of NAC domain, as the intermolecular association of NAC domain to form amyloids would result in the entropic liberation of ordered water molecules to the bulk milieu.

Interestingly, compared to Cys-78, the bound water relaxation is comparatively faster in case of Cys-90. This is probably due to Cys-90 being present towards the C-terminal end of NAC domain due to which it does not exhibit the hydrophobic effect as strongly as Cys-78 does. This observation emphasizes the importance of residue position 78 in α-synuclein fibrillation. Besides, there is a report where a 12-amino acid stretch (^{71}VTGVTAVAQKTV^{82}) in the hydrophobic NAC domain of α-synuclein has been proposed to be necessary and sufficient in α-synuclein fibrillation.[77]

Compared to both Cys-78 and 90 (NAC domain), Cys-9, 18 and 56 (N-terminus) exhibited faster bound water relaxation dynamics which, however, was slower than Cys-140 (C-terminus). The difference in the bound water solvation time for residues in N-terminus can be explained to certain extent on the basis of amino acid composition which is appreciably different from both NAC domain and C-terminus. The N-terminal domain is rich in amino acids that are hydrogen bond donors and it has been proposed that hydrogen bond donors induce a moderate slowdown in the hydrogen bond exchange dynamics.[78] Additionally, N-terminal domain contains a small stretch of hydrophobic residues. Thus, the presence of hydrophobic and hydrogen bond donating residues can account for the retarded water dynamics in N-terminus when compared to Cys-140. However, the percentage of hydrophobic residues in N-terminus is small compared to the amyloidogenic NAC domain.

The fastest bound water relaxation dynamics is demonstrated by Cys-140 which is the last residue of the polypeptide chain. In order to account for the fast motion of water molecules around Cys-140 in C-terminus, two important points must be taken into consideration: (i) 140 is the last residue and hence is highly solvated, (ii) there are significant number of proline residues in C-terminus. The proline residues are known to participate in strong hydrogen bonding with water and are also known to act as osmoprotectant in a manner that does not significantly perturb the hydrogen bonded network of bulk water.[79]

The contrasting nature of water molecules within different domains of α-synuclein might have important biological consequences, such as: (i) The bound water molecules in N-terminal domain with a markedly different hydration timescales than other two domains might

have a specific and selective role in biomolecular recognition during α-synuclein biomembrane interactions because hydration water is known to play a significant role in mediating protein-ligand binding,[60, 80] (ii) Although the ordered water molecules in NAC domain are anticipated to act as a driving force for α-synuclein fibrillation (discussed above), the distinct timescale might also be tailored to its interaction with heat shock proteins which is known to inhibit α-synuclein fibrillation.[41] (ii) The remarkably fast water dynamics around Cys-140 in C-terminus, might have an important role to play in its biological function as a chaperone wherein C-terminus protects the hydrophobic NAC domain.[46]

4.3.6 Hydration dynamics in the amyloid state of α-synuclein revealed the presence of highly restrained water around NAC domain

According to the above discussion on entropic driving force for amyloid formation, the tendency of NAC domain to undergo intermolecular interactions is expected to be higher compared to other two domains since the intermolecular association would result in the entropic liberation of structured water molecules. In order to validate this hypothesis, we

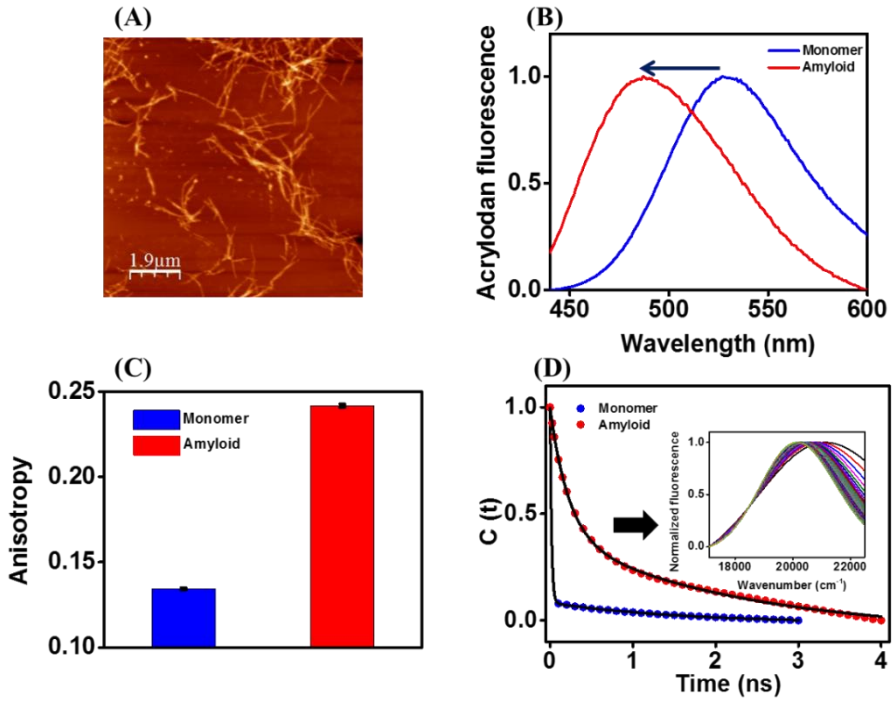


Figure 4.6 (A) AFM image of α-synuclein amyloid fibrils. The steady-state fluorescence (B) spectra and (C) anisotropy for acrylodan labeled Cys-78 α-synuclein in the monomeric and amyloid state, shown in blue and red, respectively. (D) The [C(t)] plot constructed from TCSPC data using time-zero from femtosecond experiments for acrylodan labeled Cys-78 α-synuclein in the monomeric (in blue) and amyloid (in red) state. TCSPC TRES for the amyloid state is shown as an inset. For TCSPC TRES for the monomeric state, see Figure 4.4D.

carried out the aggregation reaction using acrylodan labeled Cys-78 (For details, refer to Experimental Section). Upon fibril formation (Figure 4.6A), the emission maximum of acrylodan showed a huge blue shift of ~40 nm (from ~525 to ~495 nm) (Figure 4.6B). Additionally, an increase in fluorescence anisotropy was observed upon amyloid formation. Further, we carried out TRES measurements with these amyloid fibrils using TCSPC since we expected a nanosecond solvation component upon disorder-to-order amyloid transition, as observed for κ-casein fibrils (Chapter 3). The nanosecond component was indeed observed for α-synuclein fibrils and the time-zero recovered from our TCSPC experiment was comparable to the time-zero obtained from our femtosecond experiments, which suggests that there is no loss or probably a very little loss of ~1-5% (considering the error associated with estimation of time zero from TCSPC due to poor time resolution) in solvation within the time resolution of our TCSPC setup. This further indicates that upon disorder-to-order amyloid transition: (i) the bulk water contribution drops significantly from ~75% in the monomeric state to ~1-5% in the amyloid state, (ii) there is increase in the contribution of ultraslow solvation component (>100 ps) from ~15% in the monomeric state (for Cys-78) to ~50%. Taken together, these observations are clearly suggestive of the entropic release of ordered water molecules from NAC domain of α-synuclein, upon disorder-to-order amyloid transition. Therefore, the presence of structured water in NAC domain re-emphasizes the importance of NAC domain in the context of α-synuclein aggregation.

4.4 Summary

Our results provide novel insights into the potential role of water in α-synuclein function and disease. In this work, we elucidate the behavior of water in intrinsically disordered α-synuclein, aggregation of which underlies the pathogenesis in Parkinson's disease. To the best of our knowledge, this is the first experimental study on the hydration dynamics of α-synuclein in the monomeric state. Using ultrafast measurements, we have been able to detect the presence of ordered water molecules in the hydrophobic NAC domain of α-synuclein that are strikingly different from the water molecules in other two domains. The presence of ordered water molecules in the hydrophobic NAC domain underscores its importance in α-synuclein aggregation and provides important cues towards the designing of drugs for targeting Parkinson's disease as hydration water is known to strongly influence protein-ligand interactions. Additionally, by probing the water structure in different domains of α-synuclein under physiologically relevant condition, we have been able to identify the plausible differences in the secondary structure propensities within different regions. According to our

results, the NAC domain is anticipated to have a slightly compact local structure around it due to the presence of restrained water molecules.

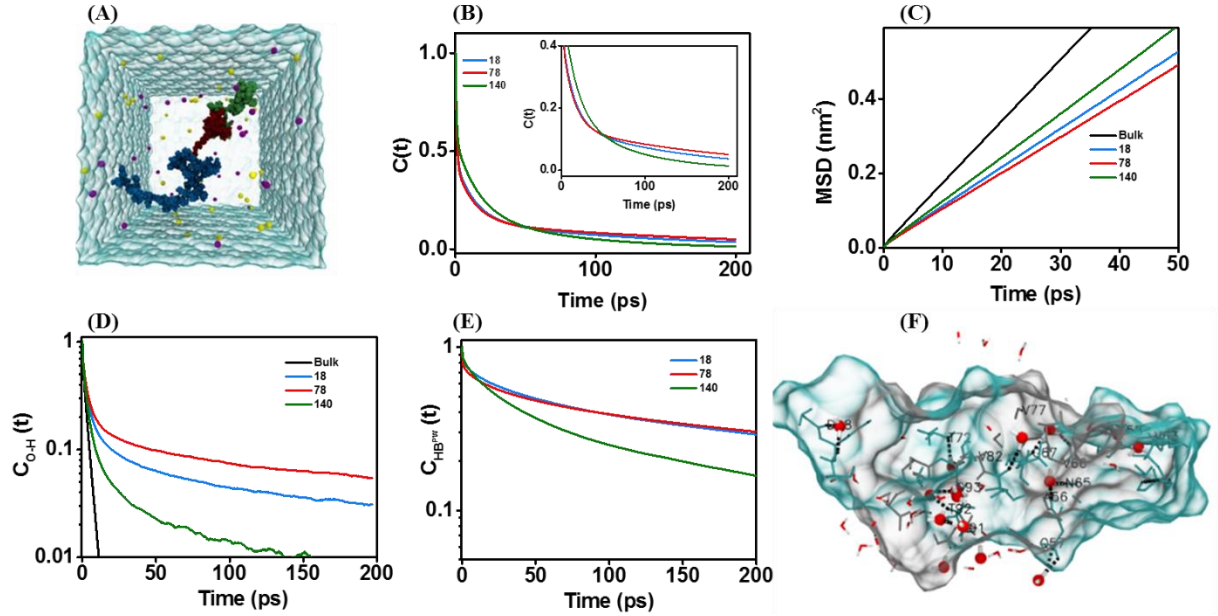


Figure 4.7 (A) α-synuclein in water. (B) Survival probability functions, C(t) and (C) Mean square displacemet (MSD) of water molecules in the α-synuclein hydration shell around three different sites. Linear-log plot of (D) second-order reorientational time correlation function C(t) for all O-H bonds of water molecules in the hydration shell (Black line represents data corresponding to bulk water) and (E) protein-water H-bond lifetime correlation functions, $C_{HB}(t)$. (F) A snapshot of hydration shell near Ala-78 and Ala-90. Water molecules constituting hydration shell of those two residues are shown using bond representation (oxygen in red and hydrogen in white). Water molecules that are longer-lived (at least 200 ps) are highlighted. Peptide residues that are in proximity of those water molecules are shown using both surface and bond representations (polar/charged in cyan and non-polar in white). Protein-water H-bonds are shown using black dashed line. A number of residues that are in contact of long-lived water molecules are labeled. The MD simulations (data shown above) were carried out by Dr. Payel Das at IBM Thomas J. Watson Research Center, Yorktown Heights, NY USA.

The slight compaction in the structure around the NAC domain is amenable to femtosecond time resolution and is sensed by the dynamic motion of water molecules present in the close proximity of the protein surface. It is possible that the specificity and selectivity of different domains in α-synuclein function and disease is tailored to the differences in the water organization. In addition, we have performed atomistic molecular dynamics (MD) simulations of α-synuclein in explicit water to provide further insights into the structure and dynamics of surface water molecules and the simulation data very well corroborates our experimental data. The simulation data also showed the presence of ordered water cluster around 78 residue in NAC domain (See Figure 4.7).

4.5 References:

1. Bagchi, B. *Water in Biological and Chemical Processes: From Structure and Dynamics to Function*. Cambridge University Press: 2013.

2. Ball, P. Water as an Active Constituent in Cell Biology. *Chem. Rev.* **2008,** *108,* 74-108.

3. Chaplin, M. Do we underestimate the importance of water in cell biology? *Nat. Rev. Mol. Cell Biol.* **2006,** *7,* 861-866.

4. Frauenfelder, H.; Chen, G.; Berendzen, J.; Fenimore, P. W.; Jansson, H.; McMahon, B. H.; Stroe, I. R.; Swenson, J.; Young, R. D. A unified model of protein dynamics. *Proc. Natl. Acad. Sci.* **2009,** *106,* 5129-5134.

5. Bagchi, B. Water Dynamics in the Hydration Layer around Proteins and Micelles. *Chem. Rev.* **2005,** *105,* 3197-3219.

6. Bhattacharyya, K.; Bagchi, B. Slow Dynamics of Constrained Water in Complex Geometries. *J. Phys. Chem. A* **2000,** *104,* 10603-10613.

7. Pal, S. K.; Zewail, A. H. Dynamics of water in biological recognition. *Chem. Rev.* **2004,** *104,* 2099-123.

8. Zhong, D.; Pal, S. K.; Zewail, A. H. Biological water: A critique. *Chem. Phys. Lett.* **2011,** *503,* 1-11.

9. Bhattacharyya, K. Nature of biological water: a femtosecond study. *Chem. Commun.* **2008,** 2848-2857.

10. Kamal, J. K.; Zhao, L.; Zewail, A. H. Ultrafast hydration dynamics in protein unfolding: human serum albumin. *Proc. Natl. Acad. Sci. U. S. A.* **2004,** *101,* 13411-6.

11. Baldwin, R. L. Dynamic hydration shell restores Kauzmann's 1959 explanation of how the hydrophobic factor drives protein folding. *Proc. Natl. Acad. Sci. U. S. A.* **2014,** *111,* 13052-6.

12. van der Lee, R.; Buljan, M.; Lang, B.; Weatheritt, R. J.; Daughdrill, G. W.; Dunker, A. K.; Fuxreiter, M.; Gough, J.; Gsponer, J.; Jones, D. T.; Kim, P. M.; Kriwacki, R. W.; Oldfield, C. J.; Pappu, R. V.; Tompa, P.; Uversky, V. N.; Wright, P. E.; Babu, M. M. Classification of Intrinsically Disordered Regions and Proteins. *Chem. Rev.* **2014,** *114,* 6589-6631.

13. Dunker, A. K.; Babu, M. M.; Barbar, E.; Blackledge, M.; Bondos, S. E.; Dosztányi, Z.; Dyson, H. J.; Forman-Kay, J.; Fuxreiter, M.; Gsponer, J.; Han, K.-H.; Jones, D. T.; Longhi, S.; Metallo, S. J.; Nishikawa, K.; Nussinov, R.; Obradovic, Z.; Pappu, R. V.; Rost, B.; Selenko, P.; Subramaniam, V.; Sussman, J. L.; Tompa, P.; Uversky, V. N. What's in a name? Why these proteins are intrinsically disordered. *Intrinsically Disordered Proteins* **2014,** *1*, e24157.

14. Uversky, V. N. A decade and a half of protein intrinsic disorder: biology still waits for physics. *Protein Sci.* **2013,** *22*, 693-724.

15. Dyson, H. J.; Wright, P. E. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* **2005,** *6*, 197-208.

16. Das, R. K.; Mittal, A.; Pappu, R. V. How is functional specificity achieved through disordered regions of proteins? *Bioessays* **2013,** *35*, 17-22.

17. Hsu, W. L.; Oldfield, C. J.; Xue, B.; Meng, J.; Huang, F.; Romero, P.; Uversky, V. N.; Dunker, A. K. Exploring the binding diversity of intrinsically disordered proteins involved in one-to-many binding. *Protein Sci.* **2013,** *22*, 258-273.

18. Uversky, V. N.; Dunker, A. K. The case for intrinsically disordered proteins playing contributory roles in molecular recognition without a stable 3D structure. *F1000 Biol. Rep.* **2013,** *5*, 1.

19. Uversky, V. N.; Dave, V.; Iakoucheva, L. M.; Malaney, P.; Metallo, S. J.; Pathak, R. R.; Joerger, A. C. Pathological unfoldomics of uncontrolled chaos: intrinsically disordered proteins and human diseases. *Chem. Rev.* **2014,** *114*, 6844-6879.

20. Knowles, T. P.; Vendruscolo, M.; Dobson, C. M. The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell Biol.* **2014,** *15*, 384-396.

21. Chiti, F.; Dobson, C. M. Protein Misfolding, Functional Amyloid, and Human Disease. *Annu. Rev. Biochem.* **2006,** *75*, 333-366.

22. Bertram, L.; Tanzi, R. E. The genetic epidemiology of neurodegenerative disease. *J. Clin. Invest.* **2005,** *115*, 1449-1457.

23. Alexander, G. E. Biology of Parkinson's disease: pathogenesis and pathophysiology of a multisystem neurodegenerative disorder. *Dialogues Clin. Neurosci.* **2004,** *6*, 259-280.

24. Spillantini, M. G.; Schmidt, M. L.; Lee, V. M. Y.; Trojanowski, J. Q.; Jakes, R.; Goedert, M. [alpha]-Synuclein in Lewy bodies. *Nature* **1997,** *388*, 839-840.

25. Goedert, M.; Spillantini, M. G.; Del Tredici, K.; Braak, H. 100 years of Lewy pathology. *Nat. Rev. Neurol.* **2013,** *9*, 13-24.

26. Trojanowski, J. Q.; Lee, V. M. Y. Parkinson's disease and related neurodegenerative synucleinopathies linked to progressive accumulations of synuclein aggregates in brain. *Parkinsonism Relat. Disord.* **2001,** *7*, 247-251.

27. Watson, J. B.; Hatami, A.; David, H.; Masliah, E.; Roberts, K.; Evans, C. E.; Levine, M. S. Alterations in corticostriatal synaptic plasticity in mice overexpressing human α-synuclein. *Neuroscience* **2009,** *159*, 501-513.

28. George, J. M.; Jin, H.; Woods, W. S.; Clayton, D. F. Characterization of a novel protein regulated during the critical period for song learning in the zebra finch. *Neuron* **1995,** *15*, 361-372.

29. Burré, J.; Sharma, M.; Tsetsenis, T.; Buchman, V.; Etherton, M.; Südhof, T. C. α-Synuclein Promotes SNARE-Complex Assembly in vivo and in vitro. *Science* **2010,** *329*, 1663-1667.

30. Chandra, S.; Gallardo, G.; Fernández-Chacón, R.; Schlüter, O. M.; Südhof, T. C. α-Synuclein Cooperates with CSPα in Preventing Neurodegeneration. *Cell* **2005,** *123*, 383-396.

31. Al-Wandi, A.; Ninkina, N.; Millership, S.; Williamson, S. J. M.; Jones, P. A.; Buchman, V. L. Absence of α-synuclein affects dopamine metabolism and synaptic markers in the striatum of aging mice. *Neurobiol. Aging* **2010,** *31*, 796-804.

32. Cooper, A. A.; Gitler, A. D.; Cashikar, A.; Haynes, C. M.; Hill, K. J.; Bhullar, B.; Liu, K.; Xu, K.; Strathearn, K. E.; Liu, F.; Cao, S.; Caldwell, K. A.; Caldwell, G. A.; Marsischky, G.; Kolodner, R. D.; LaBaer, J.; Rochet, J.-C.; Bonini, N. M.; Lindquist, S. α-Synuclein Blocks ER-Golgi Traffic and Rab1 Rescues Neuron Loss in Parkinson's Models. *Science* **2006,** *313*, 324-328.

33. Trexler A. J.; Rhoades E. α-Synuclein binds large unilamellar vesicles as an extended helix. *Biochemistry* **2009,** *48*, 2304-2306.

34. Burré, J.; Sharma, M.; Südhof, T. C. α-Synuclein assembles into higher-order multimers upon membrane binding to promote SNARE complex formation. *Proc. Natl. Acad. Sci.* **2014,** *111*, E4274-E4283.

35. Snead, D.; Eliezer, D. Alpha-Synuclein Function and Dysfunction on Cellular Membranes. *Exp. Neurobiol.* **2014,** *23*, 292-313.

36. Pfefferkorn, C. M.; Jiang, Z.; Lee, J. C. Biophysics of α-synuclein membrane interactions. *Biochim. Biophys. Acta (BBA) - Biomembranes* **2012,** *1818*, 162-171.

37. Dikiy, I.; Eliezer, D. Folding and misfolding of alpha-synuclein on membranes. *Biochim. Biophys. Acta (BBA) - Biomembranes* **2012,** *1818*, 1013-1018.

38. Shi, Z.; Sachs, J. N.; Rhoades, E.; Baumgart, T. Biophysics of [small alpha]-synuclein induced membrane remodelling. *Phys. Chem. Chem. Phys.* **2015,** *17*, 15561-15568.

39. Giasson, B. I.; Murray, I. V. J.; Trojanowski, J. Q.; Lee, V. M. Y. A Hydrophobic Stretch of 12 Amino Acid Residues in the Middle of α-Synuclein Is Essential for Filament Assembly. *J. Biol. Chem.* **2001,** *276*, 2380-2386.

40. Biere, A. L.; Wood, S. J.; Wypych, J.; Steavenson, S.; Jiang, Y.; Anafi, D.; Jacobsen, F. W.; Jarosinski, M. A.; Wu, G.-M.; Louis, J.-C.; Martin, F.; Narhi, L. O.; Citron, M. Parkinson's Disease-associated α-Synuclein Is More Fibrillogenic than β- and γ-Synuclein and Cannot Cross-seed Its Homologs. *J. Biol. Chem.* **2000,** *275*, 34574-34579.

41. Luk, K. C.; Mills, I. P.; Trojanowski, J. Q.; Lee, V. M. Y. Interactions between Hsp70 and the Hydrophobic Core of α-Synuclein Inhibit Fibril Assembly. *Biochemistry* **2008,** *47*, 12614-12625.

42. Ei-Agnaf, O. M. A.; Irvine, G. B. Aggregation and neurotoxicity of α-synuclein and related peptides. *Biochem. Soc. Trans.* **2002,** *30*, 559-565.

43. Rodriguez, J. A.; Ivanova, M. I.; Sawaya, M. R.; Cascio, D.; Reyes, F. E.; Shi, D.; Sangwan, S.; Guenther, E. L.; Johnson, L. M.; Zhang, M.; Jiang, L.; Arbing, M. A.; Nannenga, B. L.; Hattne, J.; Whitelegge, J.; Brewster, A. S.; Messerschmidt, M.; Boutet, S.; Sauter, N. K.; Gonen, T.; Eisenberg, D. S. Structure of the toxic core of alpha-synuclein from invisible crystals. *Nature* **2015,** *525*, 486-490.

44. Nielsen, M. S.; Vorum, H.; Lindersson, E.; Jensen, P. H. Ca2+ binding to alpha-synuclein regulates ligand binding and oligomerization. *J. Biol. Chem.* **2001,** *276*, 22680-22684.

45. Hoyer, W.; Cherny, D.; Subramaniam, V.; Jovin, T. M. Impact of the Acidic C-Terminal Region Comprising Amino Acids 109−140 on α-Synuclein Aggregation in Vitro. *Biochemistry* **2004,** *43*, 16233-16242.

46. Park, S. M.; Jung, H. Y.; Kim, T. D.; Park, J. H.; Yang, C. H.; Kim, J. Distinct roles of the N-terminal-binding domain and the C-terminal-solubilizing domain of alpha-synuclein, a molecular chaperone. *J. Biol. Chem.* **2002,** *277*, 28512-28520.

47. McNulty, B. C.; Young, G. B.; Pielak, G. J. Macromolecular Crowding in the Escherichia coli Periplasm Maintains α-Synuclein Disorder. *J. Mol. Biol.* **2006,** *355*, 893-897.

48. Uversky, V. N.; Li, J.; Fink, A. L. Evidence for a Partially Folded Intermediate in α-Synuclein Fibril Formation. *J. Biol. Chem.* **2001,** *276*, 10737-10744.

49. Weinreb, P. H.; Zhen, W.; Poon, A. W.; Conway, K. A.; Lansbury, P. T. NACP, A Protein Implicated in Alzheimer's Disease and Learning, Is Natively Unfolded. *Biochemistry* **1996,** *35*, 13709-13715.

50. Bernadó, P.; Bertoncini, C. W.; Griesinger, C.; Zweckstetter, M.; Blackledge, M. Defining Long-Range Order and Local Disorder in Native α-Synuclein Using Residual Dipolar Couplings. *J. Am. Chem. Soc.* **2005,** *127*, 17968-17969.

51. Bertoncini, C. W.; Jung, Y. S.; Fernandez, C. O.; Hoyer, W.; Griesinger, C.; Jovin, T. M.; Zweckstetter, M. Release of long-range tertiary interactions potentiates aggregation of natively unstructured alpha-synuclein. *Proc. Natl. Acad. Sci. U. S. A.* **2005,** *102*, 1430-1435.

52. Dedmon, M. M.; Lindorff-Larsen, K.; Christodoulou, J.; Vendruscolo, M.; Dobson, C. M. Mapping Long-Range Interactions in α-Synuclein using Spin-Label NMR and Ensemble Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **2005,** *127*, 476-477.

53. Tran, H. T.; Mao, A.; Pappu, R. V. Role of Backbone−Solvent Interactions in Determining Conformational Equilibria of Intrinsically Disordered Proteins. *J. Am. Chem. Soc.* **2008,** *130*, 7380-7392.

54. Das, R. K.; Ruff, K. M.; Pappu, R. V. Relating sequence encoded information to form and function of intrinsically disordered proteins. *Curr. Opin. Struct. Biol.* **2015,** *32*, 102-112.

55. Schiro, G.; Fichou, Y.; Gallat, F. X.; Wood, K.; Gabel, F.; Moulin, M.; Hartlein, M.; Heyden, M.; Colletier, J. P.; Orecchini, A.; Paciaroni, A.; Wuttke, J.; Tobias, D. J.; Weik, M. Translational diffusion of hydration water correlates with functional motions in folded and intrinsically disordered proteins. *Nature Commun.* **2015,** *6*, 6490.

56. Pavlova, A.; Cheng, C. Y.; Kinnebrew, M.; Lew, J.; Dahlquist, F. W.; Han, S. Protein structural and surface water rearrangement constitute major events in the earliest aggregation stages of tau. *Proc. Natl. Acad. Sci. U. S. A.* **2016,** *113*, E127-E136.

57. Conti Nibali, V.; Havenith, M. New insights into the role of water in biological function: studying solvated biomolecules using terahertz absorption spectroscopy in conjunction with molecular dynamics simulations. *J. Am. Chem. Soc.* **2014,** *136*, 12800-12807.

58. Kim, Y. S.; Liu, L.; Axelsen, P. H.; Hochstrasser, R. M. 2D IR provides evidence for mobile water molecules in β-amyloid fibrils. *Proc. Natl. Acad. Sci. U. S. A.* **2009,** *106*, 17751-17756.

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

60. Jain, N.; Bhasne, K.; Hemaswasthi, M.; Mukhopadhyay, S. Structural and Dynamical Insights into the Membrane-Bound α-Synuclein. *PLoS ONE* **2013,** *8*, e83752.

61. Arya, S.; Mukhopadhyay, S. Ordered water within the collapsed globules of an amyloidogenic intrinsically disordered protein. *J. Phys. Chem. B* **2014,** *118*, 9191-9198.

62. Burai, T. N.; Mukherjee, T. K.; Lahiri, P.; Panda, D.; Datta, A. Early events associated with the excited state proton transfer in 2-(2′-pyridyl)benzimidazole. *J. Chem. Phys.* **2009,** *131*, 034504.

63. Arya, S.; Kumari, A.; Dalal, V.; Bhattacharya, M.; Mukhopadhyay, S. Appearance of annular ring-like intermediates during amyloid fibril formation from human serum albumin. *Phys. Chem. Chem. Phys.* **2015,** *17*, 22862-22871.

64. Jha, A.; Ishii, K.; Udgaonkar, J. B.; Tahara, T.; Krishnamoorthy, G. Exploration of the correlation between solvation dynamics and internal dynamics of a protein. *Biochemistry* **2011,** *50*, 397-408.

65. Horcas, I.; Fernández, R.; Gómez-Rodríguez, J. M.; Colchero, J.; Gómez-Herrero, J.; Baro, A. M. WSXM: A software for scanning probe microscopy and a tool for nanotechnology. *Rev. Sci. Instrum.* **2007,** *78*, 013705.

66. Prendergast, F. G.; Meyer, M.; Carlson, G. L.; Iida, S.; Potter, J. D. Synthesis, spectral properties, and use of 6-acryloyl-2-dimethylaminonaphthalene (Acrylodan). A thiol-selective, polarity-sensitive fluorescent probe. *J. Biol. Chem.* **1983,** *258*, 7541-7544.

67. Changenet-Barret, P.; Choma, C. T.; Gooding, E. F.; DeGrado, W. F.; Hochstrasser, R. M. Ultrafast Dielectric Response of Proteins from Dynamics Stokes Shifting of Coumarin in Calmodulin. *J. Phys. Chem. B* **2000,** *104*, 9322-9329.

68. Pal, S. K.; Peon, J.; Zewail, A. H. Biological water at the protein surface: dynamical solvation probed directly with femtosecond resolution. *Proc. Natl. Acad. Sci. U. S. A.* **2002,** *99*, 1763-1768.

69. Pal, S. K.; Peon, J.; Zewail, A. H. Ultrafast surface hydration dynamics and expression of protein functionality: alpha -Chymotrypsin. *Proc. Natl. Acad. Sci. U. S. A.* **2002,** *99*, 15297-15302.

70. Peon, J.; Pal, S. K.; Zewail, A. H. Hydration at the surface of the protein Monellin: dynamics with femtosecond resolution. *Proc. Natl. Acad. Sci. U. S. A.* **2002,** *99*, 10964-10969.

71. Fee, R. S.; Maroncelli, M. Estimating the time-zero spectrum in time-resolved emmsion measurements of solvation dynamics. *Chem. Phys.* **1994,** *183*, 235-247.

72. Nandi, N.; Bagchi, B. Dielectric Relaxation of Biological Water. *J. Phys. Chem. B* **1997,** *101*, 10954-10961.

73. Jia, M.; Yang, J.; Qin, Y.; Wang, D.; Pan, H.; Wang, L.; Xu, J.; Zhong, D. Determination of Protein Surface Hydration by Systematic Charge Mutations. *J. Phys. Chem. Lett.* **2015,** *6*, 5100-5105.

74. Frank, H. S.; Evans, M. W. Free Volume and Entropy in Condensed Systems III. Entropy in Binary Liquid Mixtures; Partial Molal Entropy in Dilute Solutions; Structure and Thermodynamics in Aqueous Electrolytes. *J. Chem. Phys.* **1945,** *13*, 507-532.

75. Kauzmann, W. Some Factors in the Interpretation of Protein Denaturation1. In *Adv. Protein Chem.*, C.B. Anfinsen, M. L. A. K. B.; John, T. E., Eds. Academic Press: 1959; Volume 14, pp 1-63.

76. Grupi, A.; Haas, E. Segmental Conformational Disorder and Dynamics in the Intrinsically Disordered Protein α-Synuclein and Its Chain Length Dependence. *J. Mol. Biol.* **2011,** *405*, 1267-1283.

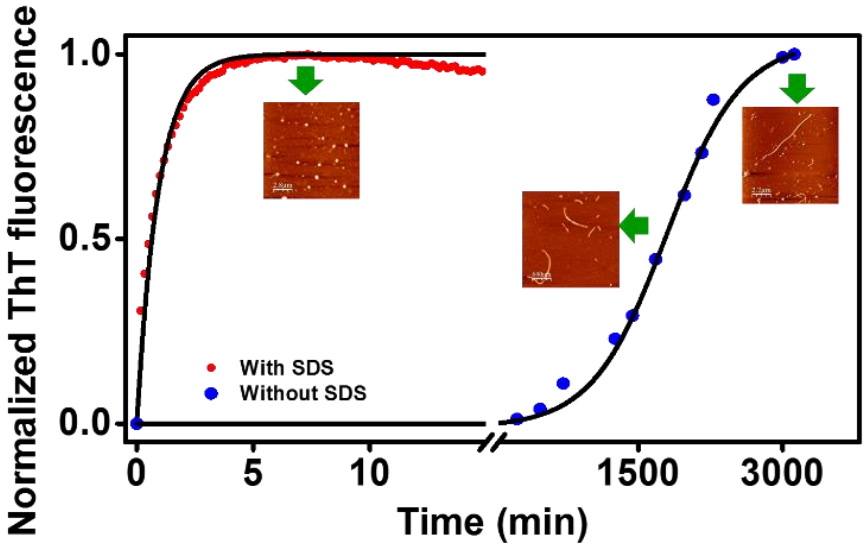
77. Giasson, B. I.; Murray, I. V.; Trojanowski, J. Q.; Lee, V. M. A hydrophobic stretch of 12 amino acid residues in the middle of alpha-synuclein is essential for filament assembly. *J. Biol. Chem.* **2001,** *276*, 2380-2386.

78. Sterpone, F.; Stirnemann, G.; Laage, D. Magnitude and molecular origin of water slowdown next to a protein. *J. Am. Chem. Soc.* **2012,** *134*, 4116-4119.

79. McLain, S. E.; Soper, A. K.; Terry, A. E.; Watts, A. Structure and Hydration of l-Proline in Aqueous Solutions. *J. Phys. Chem. B* **2007,** *111*, 4568-4580.

80. Cheng, C.-Y.; Varkey, J.; Ambroso, M. R.; Langen, R.; Han, S. Hydration dynamics as an intrinsic ruler for refining protein structure at lipid membrane interfaces. *Proc. Natl. Acad. Sci.* **2013,** *110*, 16838-16843.

Mechanism of Aggregation of a Model Amyloidogenic Intrinsically Disordered Protein



5.1 Introduction

Protein aggregation leading to amyloid formation is well-known for its association with a variety of deadly human diseases.[1-7] However, the highly ordered cross β-sheet rich architecture of amyloid fibrils also offers a promising route for the fabrication of robust nano-structures and is being increasingly recognized for a wide variety of applications, such as in bio-nanotechnology[8-11] and food processing industry.[12,13] Typically, protein aggregation is favored under conditions that reduce the stability of the native state of a protein and favor the protein-protein (intermolecular) interactions to form the critical aggregation nuclei.[14] An enormous progress been made in the area of amyloid biology in the past few years, particularly in understanding the mechanism of amyloid formation.[15,16] However, the precise mechanism of amyloid formation still remains elusive due to the complexity associated with the process arising mainly due to the involvement of a large number of factors that influence this process. In order to develop drug targets that can potentially disrupt the interactions that favor amyloid formation and to design novel protein/amyloid based nanomaterials, it is important to understand the underlying mechanism of fibrillation and the factors that predominantly influence this process. One of the important factors that can often trigger the association of proteins is the presence of surfaces with certain physicochemical features that favor aggregation by profoundly influencing the stability of the native state of a protein.[17-18] The recruitment of proteins or peptides by synthetic or biological surfaces can result in the conformational transition of the native (functional) state to the non-native, aggregation prone (non-functional/pathological) state.[17,18] Therefore, there is a growing interest in investigating the physicochemical characteristics of interaction between protein and synthetic or natural surfaces, and aggregation of protein on these surfaces, predominantly lipid bilayers. Lipid bilayer is the fundamental structural constituent of the biological membranes and is considered a two-dimensional liquid that provides a variety of environments, which can significantly influence the structure and dynamics of proteins. Protein-membrane interactions are believed to play a key role in protein misfolding.[19-21] In fact, many of the pathological amyloid deposits have been found to interact with membranes thereby causing membrane disruption.[22-30] Many of these pathological amyloids are often formed by intrinsically disordered proteins (IDPs).[22,24,25] In order to investigate the role of membrane in the amyloid formation by many proteins including IDPs, the amphiphilic molecules such as sodium dodecyl sulfate (SDS), have been extensively used since they provide a membrane mimetic environment.[31,32]

In this study, we have investigated the influence of SDS as an anionic membrane mimetic on the fibrillation of a model amyloidogenic IDP, namely, bovine κ-casein. κ-casein is a milk protein and is a member of a protein family called caseins. Caseins belong to the functional class of IDPs known as scavengers.[33] They are the major protein elements of mammalian milk and play an important role in the cellular uptake of calcium and phosphate ions.[34,35] Additionally, caseins are well recognized for their capability to act as molecular chaperones.[36-39] This chaperone behavior of caseins has significant industrial potential as by making use of this property, the texture and the stability of processed dairy and other high protein foods can be improved.[39,40] The chaperone-like activity of α_{s1}- and β-casein also has considerable physiological advantages in preventing the amyloid formation from two amyloidogenic caseins, α_{s2}- and κ-casein, under physiological conditions.[41-43] The amyloid formation by κ-casein is linked with mammary cancer in bovine.[44,45] Since κ-casein is a milk protein, therefore, in order to understand the process of amyloid formation from κ-casein, it is important to consider the influence of other components of milk on amyloid formation. One of the important components of milk is lipids that comprise approximately 4% w/v of milk.[46] In addition, the membrane of mammary epithelial cells is composed of phospholipid bilayer.[47] Therefore, studying the effect of lipids or lipid mimetics on κ-casein would be of physiological relevance in understanding the mechanism of κ-casein fibrillation. Here, we have used SDS, an anionic detergent, to mimic membrane environment in κ-casein fibrillation. In vitro protein aggregation studies have revealed that the kinetics of fibrillation is often altered in the presence of lipids/detergents.[48,49] In many cases, they accelerate the fibrillation process.[31,32] However, there are cases where they can have inhibitory effect on fibrillation.[50,51] In case of κ-casein, SDS resulted in a drastic alteration in the aggregation mechanism pathway that κ-casein otherwise follows in the absence of any external agent. Additionally, in presence of SDS, mostly large oligomeric species were observed, that did not immediately convert into fibrils.

5.2 Experimental Section

5.2.1 Materials

κ-casein (from bovine milk), dithiothreitol (DTT), iodoacetic acid, guanidinium chloride (GdmCl), Tris HCl, sodium chloride (NaCl), sodium dodecyl sulfate (SDS), 8-anilino-1-naphthalenesulfonic acid ammonium salt (ANS) and sodium hydrogen phosphate (monobasic) were procured from Sigma (St. Louis, MO) and used as received. 5-((((2-iodoacetyl)amino)ethyl)amino) naphthalene-1-sulfonic acid (IAEDANS) was purchased from

Molecular Probes, Invitrogen Inc. All solutions were prepared in Milli-Q water. The pH of the buffers was adjusted using a Metrohm pH meter at ~25 °C.

5.2.2 Circular dichroism (CD) measurements

The CD measurements were carried out on Chirascan Spectrophotometer (Applied Photophysics, UK) using a 1 mm path length quartz cell, and the spectra obtained were corrected for buffer background signal. The buffer corrected spectra were then smoothened using Pro Data software.

5.2.3 Steady-state fluorescence measurements

For all the experiments, except for the data shown in Figure 5.1A, κ-casein was reduced and carboxymethylated (RCM) using DTT and iodoacetic acid using the protocol described previously[52,53] and the stock solution of κ-casein was stored under denatured condition (6 M GdmCl in pH 7, 50 mM phosphate buffer). For Trp steady-state fluorescence and CD measurements, the protein stock was diluted into the native buffer (pH 7.2, 50 mM phosphate buffer) to obtain a final concentration of 20 μM. For SDS binding experiments, a fresh stock of 10 mM SDS was prepared in Milli-Q water and was used for making further dilutions. For AEDANS fluorescence measurements, reduced and denatured κ-casein was labeled with a 50 M excess of IAEDANS under denatured condition (6 M GdmCl in pH 7.6, 100 mM Tris buffer, 2 h at 37 °C). The labeled protein was then passed through a PD-10 column to remove excess dye and was further concentrated using AMICON ultra (3 kDa cutoff; from Millipore). The concentration of the labeled protein was estimated using ε337 = 6100 M^{-1} cm^{-1} for AEDANS.[54] The AEDANS labeled κ-casein was stored under denatured condition (6 M GdmCl in pH 7, 50 mM phosphate buffer).

All the steady-state fluorescence measurements were carried out on Fluoromax-4 (Horiba Jobin Yvon, NJ) at ~25ºC. The samples were excited at 295 nm (Trp), 450 nm (ThT) and 375 nm (ANS and AEDANS). The steady-state fluorescence anisotropies were measured at 350 nm (Trp) and 480 nm (ANS and AEDANS). The steady-state fluorescence anisotropy (r_{ss}) is given by the following relationship:

$$r_{ss} = \frac{I_{\parallel} - I_{\perp} G}{I_{\parallel} + 2 I_{\perp} G} \qquad (1)$$

where $I_{\parallel}$ and $I_{\perp}$ are fluorescence intensities collected using parallel and perpendicular geometry, respectively. For Trp fluorescence kinetics measurements, κ-casein in native buffer

was taken as baseline and then SDS (100 μM) was added to this solution. The SDS solution was manual mixed and the time of mixing was typically 5-10 seconds. For monitoring ThT and ANS fluorescence kinetics upon addition of SDS (100 μM); ThT (20 μM) and ANS (20 μM), respectively were already present in the buffer containing κ-casein (20 μM).

5.2.4 Aggregation reaction

The aggregation reaction was initiated using 200 μM of RCM κ-casein in 50 mM phosphate buffer (pH 7.2). The reaction mixture was continuously stirred at a speed of 300 rpm and the temperature was maintained at 37°C. The aliquots (10 μM) were taken out at several time points and Trp fluorescence was measured. For measuring ThT and ANS fluorescence, 10 μM of ThT and ANS, respectively were added to the aliquots (10 μM) taken out from the aggregation mixture. All the final dilutions were made in 50 mM phosphate buffer (pH 7.2). For carrying out aggregation reaction with AEDANS labeled κ-casein, 20 μM of AEDANS labeled κ-casein was added to 180 μM of RCM κ-casein. The aliquots were taken out at different time points and AEDANS fluorescence (20 μM) was measured without any further dilution. The parameters used for fluorescence measurements are mentioned above. The fluorescence data at different points were then plotted and fitted with the following equation described previously for nucleation-dependent polymerization.[55]

$$y = y0 + \frac{a}{1+\exp(-\frac{x-x0}{b})} \qquad (2)$$

where y is the fluorescence at time x, $y0$ is the initial fluorescence value, $x0$ is the time when fluorescence reaches 50% of its maximum value, and a is the maximum fluorescence at stationary phase. The lag time is given by ($x0-2b$).

5.2.5 Time resolved fluorescence measurements

The fluorescence lifetime data were acquired using a time-correlated single photon counting (TCSPC) setup (Fluorocube, Horiba Jobin Yvon, NJ). The peak count was fixed to 10,000 and the excitation and emission polarizers were oriented at the magic angle (54.7^0). For time-resolved fluorescence anisotropy measurements, the peak difference was 10,000 counts and the orientation of the emission polarizer was 0^0 and 90^0 with respect to the excitation polarizer for parallel fluorescence intensities ($I_{\parallel}$) and perpendicular fluorescence intensities ($I_{\perp}$), respectively. The emission monochromator for ANS was fixed at 480 nm with a bandpass of 12 nm and 375 nm laser diode was used as excitation source. The instrument response function

(IRF) was collected using Ludox (colloidal silica). The width (FWHM) of IRF was ~250 ps. The fluorescence intensity decay curves were deconvoluted taking IRF into account using following equation:

$$I(t) = \Sigma_i \alpha_i e^{-t/\tau_i} \qquad (2)$$

5.2.6 Atomic force microscopy (AFM) imaging

AFM images were acquired on an Innova atomic force microscope (Bruker). The AFM was operated in tapping mode. For imaging, the silicon nitride cantilever probe with radius ~8 nm was used. The samples (typically diluted 1000-2000 fold with aggregation reaction buffer filtered through a 0.22 μm syringe filter) were deposited on a freshly cleaved muscovite mica (Grade V-4 mica from SPI, PA). The mica surface was incubated with the sample for 15-20 min before washing it with filtered Milli-Q water and dried under a gentle stream of nitrogen gas. The images were collected in NanoDrive (v8.03) software at a resolution of 1024 x 1024 pixels. The collected AFM images were further processed and analyzed using WSxM version 4 develop 11.6 software.[56]

5.3 Results and Discussion

5.3.1 κ-casein aggregation under non-reducing and reducing conditions

Under (native) physiological conditions, κ-casein exist as an ensemble of collapsed disordered globules. In order to follow how these collapsed globules of κ-casein coalesce to form amyloids, ThT fluorescence assay was used. Since κ-casein contains two cysteine residues that can potentially form intra- and inter-molecular disulfides, we carried out aggregation reaction using a high concentration of κ-casein (200 μM), under reducing (using DTT as a reducing agent) as well as non-reducing conditions, to decipher the role of disulfides in κ-casein aggregation. A time-dependent increase in ThT fluorescence with a typical initial lag phase, was observed with the reduced form of κ-casein (Figure 5.1A). However, in the absence of reducing agent, no time-dependent change in ThT fluorescence was observed which suggests that intermolecular disulfide bonding prevents or slows down κ-casein fibrillation (Figure 5.1A). This observation is in agreement with previous studies on κ-casein fibrillation.[43] However, the aggregation mechanism observed by us for RCM κ-casein is significantly different from the mechanism proposed previously.[43,57] This sharp contrast could be due to early oligomer formation taking place when reaction is initiated from RCM κ-casein in native

buffer[43,57] which cannot be the case in our studies as we are initiating the aggregation reaction by transfering the protein from the denatured state to the native buffer.

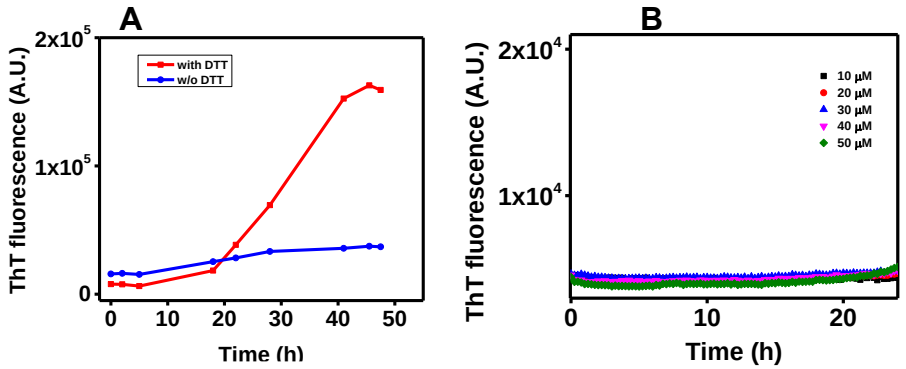


Figure 5.1 Aggregation kinetics monitored using ThT fluorescence for (A) 200 μM of κ-casein, with (red) and without (blue) DTT (1mM). (B) various concentrations of RCM κ-casein: 10 μM (black), 20 μM (red), 30 μM (blue), 40 μM (pink) and 50 μM (olive).

Further, we carried out aggregation reaction with the reduced form of κ-casein (RCM κ-casein), in a concentration-dependent manner, for a range of concentration, varying from 10 μM to 50 μM. No change in ThT fluorescence over a very long period of time was observed, which suggests that κ-casein retains its monomeric state for atleast 20-25 h under normal physiological conditions (Figure 5.1B).

5.3.2 Mechanism of κ-casein aggregation using multiple fluorescence probes

We followed κ-casein aggregation using multiple fluorescent probes to discern the time-dependent progression of the molecular events involved in κ-casein aggregation. Fluorescence spectroscopy being a highly sensitive and multi-parametric technique is quite extensively used for studying protein conformational and dynamical changes by making use of both intrinsic and extrinsic fluorophores.[58] The steady-state fluorescence spectrum provides significant insights into the local environment around the fluorophore and information about an overall size of the protein can be obtained from steady state fluorescence anisotropy measurements.[58] Since κ-casein has two cysteines and a single Trp residue, in order to probe the conformational changes taking place during aggregation, we took advantage of Trp as intrinsic fluorophore and used AEDANS as extrinsic fluorophore by labeling the thiols of Cys residues with IAEDANS. All the probes showed the existence of a nucleation-dependent amyloid assembly pathway with a characteristic lag phase. The lag time and the rate constants for the aggregation reaction were recovered from the time-dependent plots of all the fluorescence readouts using

an equation mentioned previously[55] for fitting a typical nucleation dependent polymerization curve (Figure 5.2). The lag time observed for all the fluorescence readouts was found to be comparable, except for ANS anisotropy. Additionally, the rate of time-dependent change (in the log phase) appeared to be similar for all the fluorescence readouts, except for ANS anisotropy. The similarity in the rates indicate that the structural reorganization, particularly in the region harboring Trp and Cys; and fibrillation occur simultaneously. The lag phase was shorter and the time-dependent increase in ANS fluorescence anisotropy was much slower during amyloid formation. ANS is a highly hydrophobic environment sensitive dye. The free form of ANS is almost non-fluorescent in water with an emission maximum at ~515 nm. However, it becomes fluorescent upon binding to hydrophobic pockets and undergoes a significant blue shift to ~475 nm which is mostly accompanied by a rise in intensity.[59] Due to the ability of ANS to bind to hydrophobic regions/clusters, it is often used in aggregation studies to follow the initial oligomerization and growth steps during fibrillation.

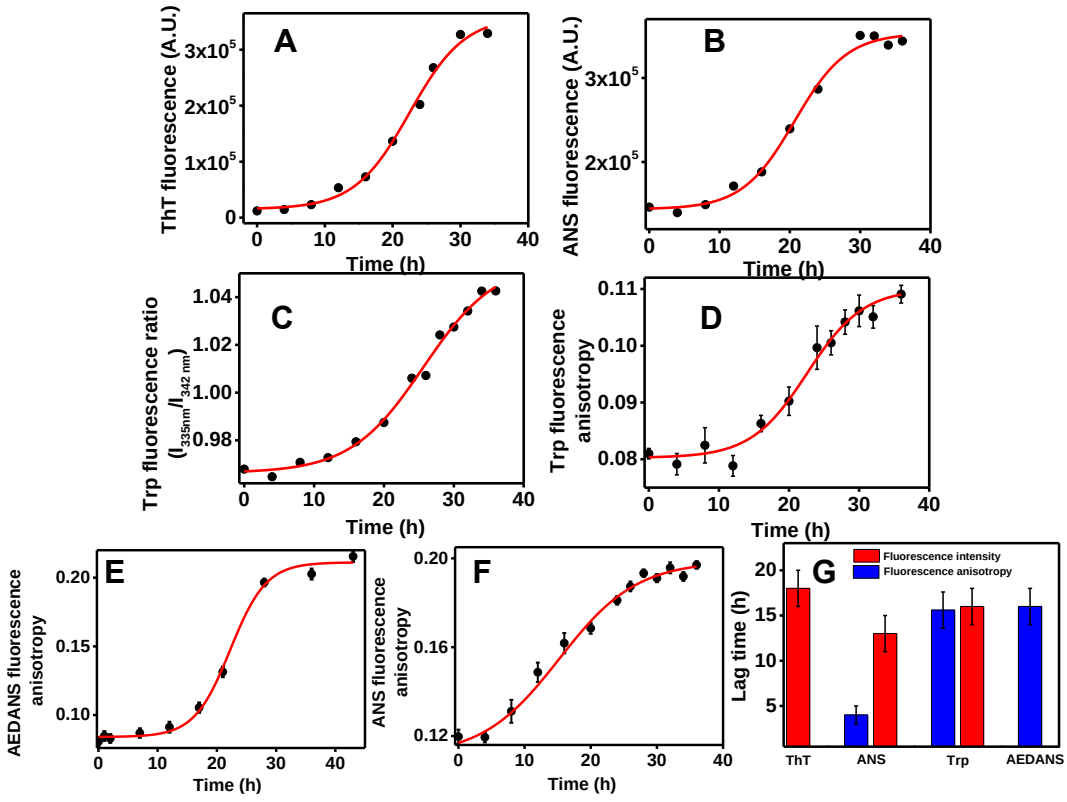


Figure 5.2 Aggregation kinetics monitored using (A) ThT fluorescence, (B) ANS fluorescence (C) Ratio Trp fluorescence intensity at 335 nm and 342 nm, (D) Trp fluorescence anisotropy, (E) AEDANS fluorescence anisotropy, and (F) ANS fluorescence anisotropy. (G) The lag time recovered from fitting the time-dependent fluorescence data shown in (A-F) using equation 2 (For details, refer to Experimental Section).

The monomeric form of κ-casein exists as an ensemble of collapsed disordered globules[53] and hence has a very weak affinity for ANS. However, as these collapsed globules coalesce during aggregation process, an increase in ANS fluorescence anisotropy and fluorescence intensity along with a blue shift is observed. The shorter lag phase observed for ANS anisotropy is suggestive of faster oligomerization, whereas, a slower increase observed in the log phase is indicative of a slower overall growth process leading to fibrillation.[54] After characterizing the mechanism of κ-casein fibrillation under physiological conditions (and in the absence of any external agent), we next embarked upon the studies aimed at understanding the effect of SDS on the fibrillation process.

5.3.3 κ-casein binds to SDS in a concentration-dependent manner

We followed SDS induced conformational changes in κ-casein, using CD and steady-state fluorescence. Upon changing the concentration of SDS, changes in both CD and fluorescence attributes were observed. An increase in the ellipticity at 218 nm with respect to ellipticity at 205 nm was observed with increase in SDS concentration (Figure 5.3A and B).

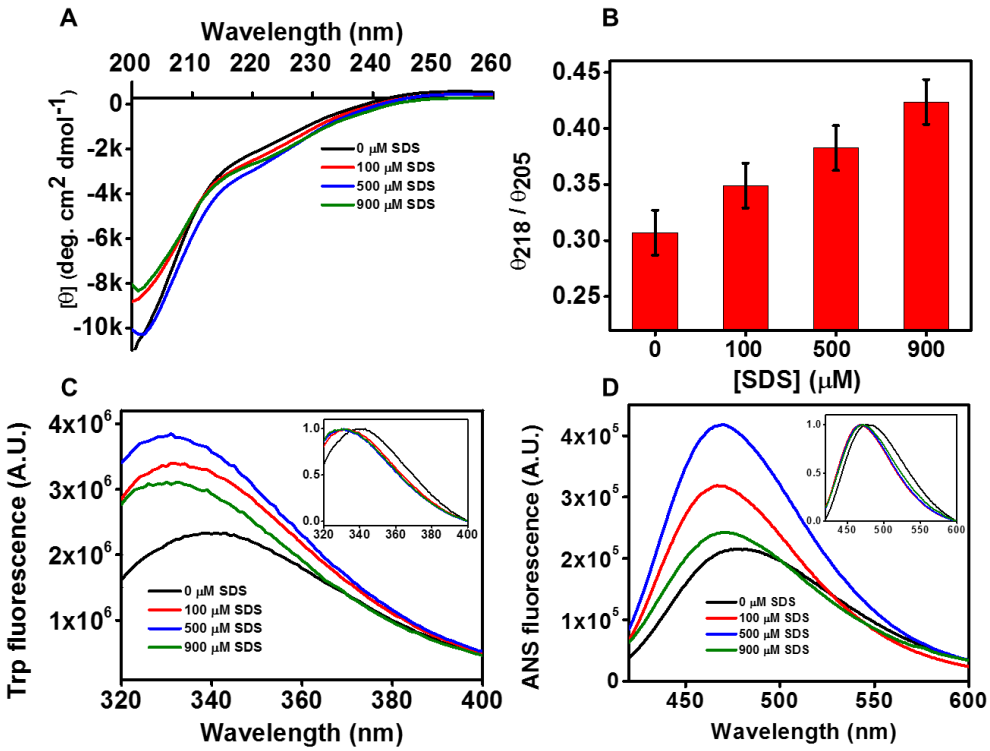


Figure 5.3 (A) CD spectra for κ-casein with various concentration of SDS: monomer (black), 100 μM (red), 500 μM (blue) and 900 μM (olive). (B) The ellipticity ratio from CD data shown in panel A. The error bar is from independent set of experiments. (C) Trp and (D) ANS fluorescence spectra with color scheme same as panel A. The normalized Trp and ANS fluorescence spectra are shown as insets to C and D, respectively.

The CD signal at 205 nm and 218 nm predominantly corresponds to random coil and β-sheets, respectively. Thus, our CD data indicates that in the presence of SDS, α-helix or β-sheet formation takes place at the expense of random coil (Figure 5.3A and B). It is important to emphasize here that the concentrations of SDS that we have used in our experiments are below the critical micellar concentration (cmc) of SDS.

Next, we monitored the changes in the fluorescence attributes of Trp which is present in Q/N-rich positively charged N-terminal domain. The Q/N-rich domain of κ-casein shares some resemblance with the amyloidogenic segment of yeast prion proteins.[60-61] An increase in Trp fluorescence with a concomitant blue shift in the emission maximum was observed with increase in SDS concentration (Figure 5.3C). At higher concentration of SDS (>500 μM), a drop in the fluorescence along with some precipitation was observed. However, this drop in the fluorescence was not associated with a red shift in the emission maximum and thus, it is likely to be either due to the precipitation or fluorescence quenching from the neighboring amino acid residues. We next monitored the changes in the hydrophobicity, taking place in the presence of SDS using ANS, which is a well-known hydrophobic environment reporter. An increase in ANS fluorescence, accompanied by a blue shift in the emission maximum was observed with an increase in SDS concentration and like Trp, at higher concentration of SDS (>500 μM), a drop in ANS fluorescence along with some precipitation was observed. This set of data indicates that SDS-induced conformations have higher affinity towards ANS compared to the monomeric state. Since ANS gives us information about overall hydrophobicity, we speculate that in the presence of SDS, there is formation of some hydrophobic pockets or clusters where ANS is getting encapsulated. ANS molecules can potentially bind to SDS molecules. However, at the concentrations of SDS that we have used in our study, we did not see any ANS binding.

Taken together, our data from CD and fluorescence suggest that κ-casein undergoes appreciable conformational change in the presence of SDS and this change appears to be SDS concentration dependent. Though both CD and fluorescence clearly indicated some conformational change in κ-casein in the presence of SDS, it was not clear whether κ-casein remains monomeric or forms higher order aggregates, akin to other IDPs. Since CD data indicated an increase in $\theta_{218}/\theta_{205}$, we suspected the possibility of aggregation leading to β-sheet rich conformations in the presence of SDS since SDS-induced fibrillation has been previously reported for some IDPs.[31-32, 62] Therefore, we next carried out fluorescence measurements with ThT which is a well-known amyloid marker dye.

5.3.4 κ-casein aggregates without a lag phase in the presence of SDS

A high affinity for ThT was observed for SDS-induced conformations of κ-casein and this binding affinity for ThT was found to be highly concentration dependent (Figure 5.4). The strong dependence on the concentration of κ-casein as well as SDS was observed (Figure 5.4A and B). Thus, it is clear that κ-casein forms ThT active aggregates in the presence of SDS. In order to decipher the mechanism of SDS-induced aggregation, we followed the kinetics of this conformational change using ThT, Trp and ANS. All fluorescent probes indicated that SDS-induced aggregation of κ-casein follows an isodesmic polymerization pathway (Figure 5.4D-F) wherein aggregation occurs rapidly without any lag phase.[63] The absence of a lag phase in SDS-induced aggregation suggests that the nuclei are formed immediately upon addition of SDS. This is clearly in sharp contrast with the nucleation-dependent amyloid assembly pathway observed for κ-casein fibrillation in the absence of SDS (Figure 5.2).

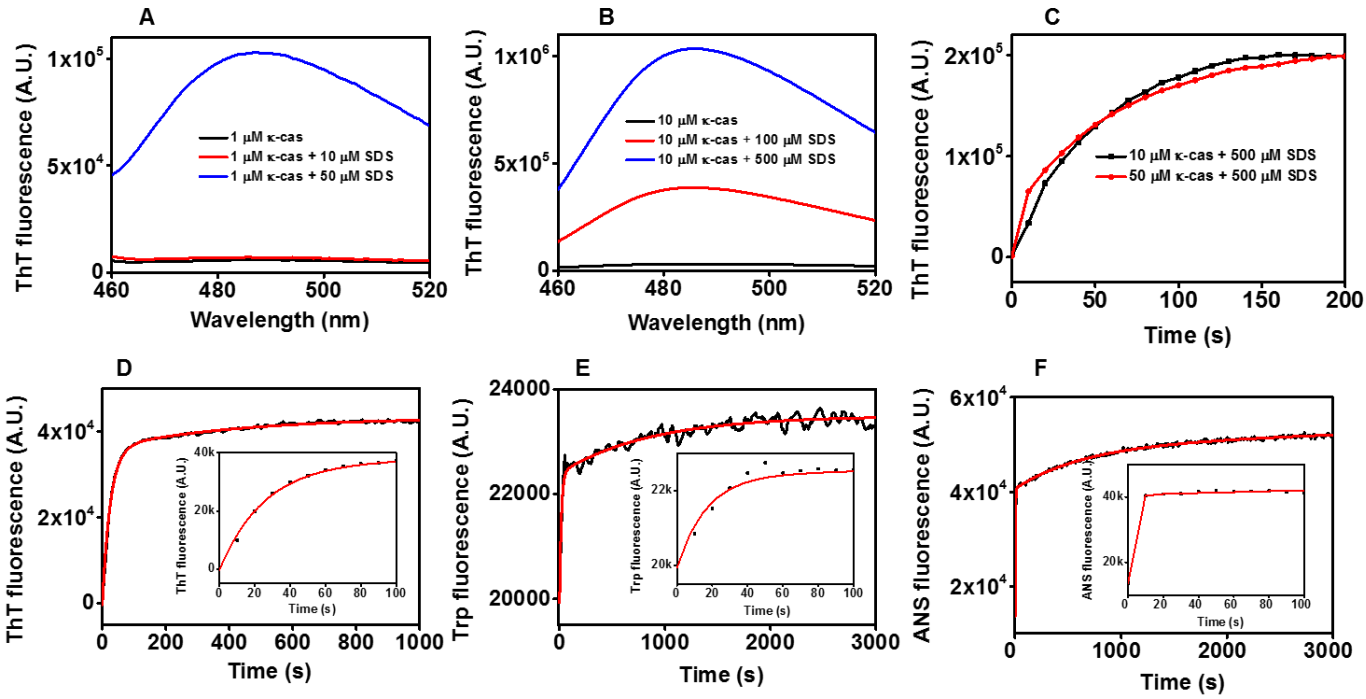


Figure 5.4 ThT fluorescence spectra for (A) 1 μM of κ-casein without (black) and with (i) 10 μM (red) and (ii) 50 μM (blue) of SDS. (B) 10 μM of κ-casein without (black) and with (i) 100 μM (red) and (ii) 500 μM (blue) of SDS. (C) ThT fluorescence kinetics data for (i) 10 μM (black) and (ii) 50 μM (red) of κ-casein with 500 μM of SDS. (D) ThT, (E) Trp and (F) ANS fluorescence kinetics data for 20 μM of κ-casein with 100 μM of SDS.

Thus, our results suggest that SDS acts as a conformational catalyst for κ-casein aggregation and induces a switch in the aggregation mechanism from a typical nucleation dependent polymerization[60,63,64] (that requires a high concentration of κ-casein) to isodesmic polymerization (that does not require a high concentration of κ-casein). Interestingly, ThT and Trp fluorescence reached saturation much faster compared to ANS which suggests a slower growth for these aggregates, akin to the amyloid fibrils formed without SDS (See Section 5.3.2). After confirming that κ-casein undergoes aggregation in the presence of SDS, we next compared the secondary structural and fluorescence attributes of the final conformation attained by the aggregates formed in the absence and in the presence of SDS. We refer to the aggregates formed under physiological conditions, in the absence of SDS, as amyloids as it has

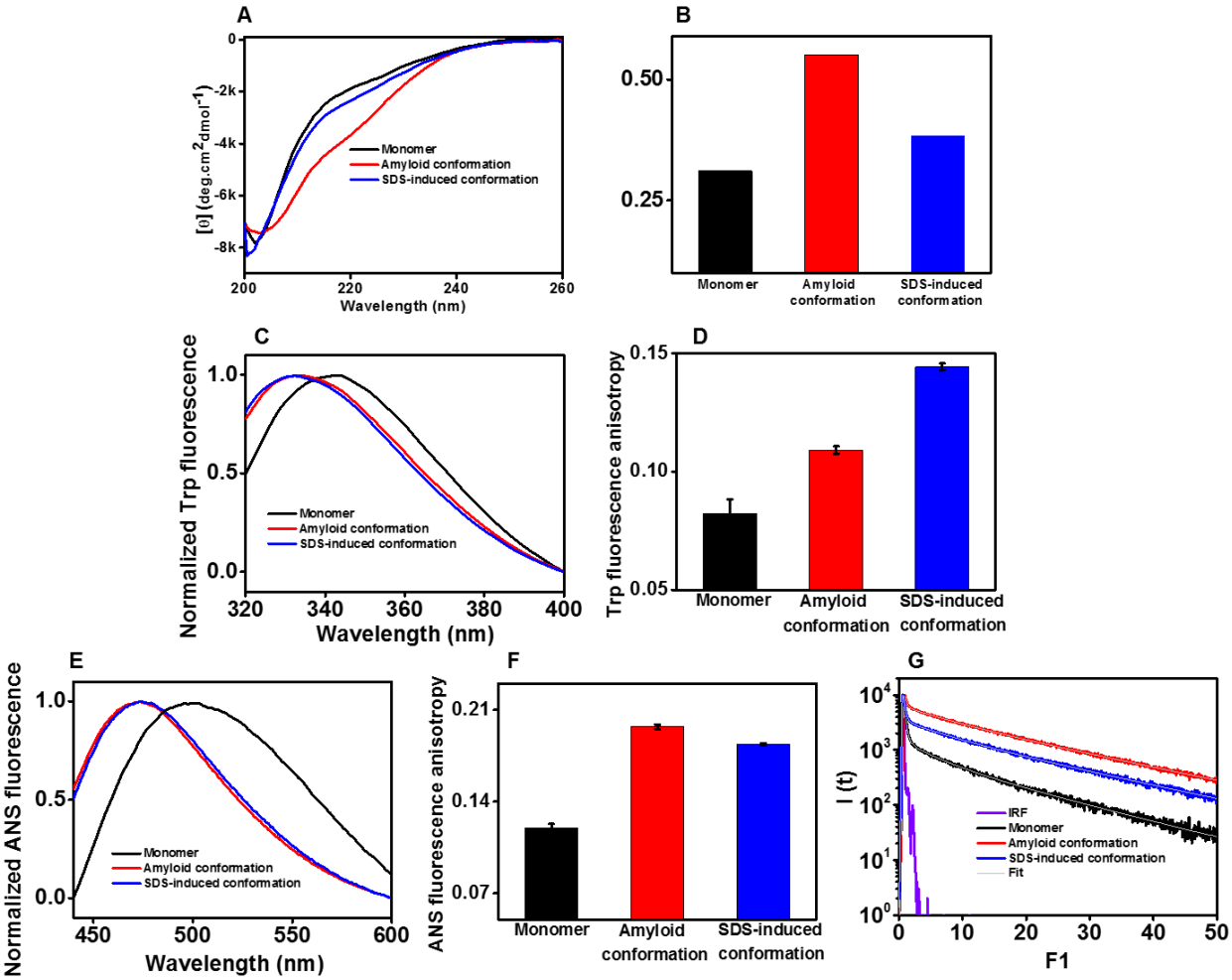


Figure 5.5 (A) CD spectra, (B) the ellipticity ratio obtained from CD data shown in panel A, (C) Normalized Trp fluorescence spectra, (D) Trp fluorescence anisotropy, (E) Normalized ANS fluorescence spectra, (F) ANS fluorescence anisotropy and (G) ANS fluorescence lifetime for monomer (black), Amyloid conformation (red) and SDS (100 μM)-induced conformation (blue).

already been established in Chapter 3 that κ-casein forms amyloid fibrils under this condition and the aggregate conformation obtained in the presence of SDS has been referred to as SDS-induced conformation. For comparison, we are only considering the aggregates formed with 100 μM SDS because the aggregates formed at higher concentration of SDS tend to phase out as a function of time. It is interesting to note that β-sheet content is higher for aggregates formed in the absence of SDS as is clearly visible from $\theta_{218}/\theta_{205}$ (Figure 5.5A and B). In case of Trp fluorescence, though the extent of blue shift is similar for both type of aggregates, anisotropy is much higher for SDS-induced conformation of aggregates (Figure 5.5C and D). This indicates that the structure around Trp is much more rigid in case of aggregates formed in the presence of SDS. This plausibly suggests that SDS molecules neutralize the positive charge in N-terminal domain that harbors Trp and is also rich in aggregation prone Q and N residues. The neutralization of charge in N-terminal domain might lead to favorable chain-chain interactions as a result of which aggregation process gets accelerated. After Trp, we next compared ANS fluorescence characteristics of both the aggregates. The aggregates formed in the presence of SDS showed much lower ANS binding as is clear from the fluorescence anisotropy and lifetime data (Figure 5.5 E-G). The lower ANS binding for the aggregates formed in the presence of SDS is probably associated with the lower content of ordered β-sheet structures, compared to the aggregates formed without SDS. Thus, our ANS corroborates CD data very well. After gaining significant insights into the conformational attributes of the aggregates formed under two different solution conditions, we next performed atomic force microscopy (AFM) imaging to discern the nanoscale morphology of these aggregates.

5.3.5 Insights into the nanoscale morphology of aggregates

Atomic force microscopy (AFM) is quite extensively used for imaging the nanoscale topography of protein aggregates.[54, 65-68] Using AFM, we could clearly see the differences in the aggregates formed in the presence/absence of SDS. In the absence of SDS, we observed long thread-like amyloid fibrils of 6-7 nm in height. The height profiles of these fibrils observed in AFM images collected at different time points in the log phase of aggregation are shown in Figure 5.6(A-C). In the presence of SDS, AFM imaging revealed a heterogeneous co-existence of small and large-sized spherical oligomers (Figure 5.6(D-F)). Mostly, large sized oligomers ranging from 20-60 nm were observed. Though we could not detect the formation of amyloid fibrils from these oligomeric species, we do not completely rule out the possibility of fibril formation upon a much longer incubation of few weeks to few months. Previous studies on κ-

casein aggregation have proposed that the rate determining step in κ-casein fibrillation is the dissociation of larger oligomers into monomeric species.[44] Since in the presence of SDS, mostly large oligomers were observed, it is possible that these oligomers being quite stable require a much longer time to dissociate into smaller amyloidogenic species. As a result, fibril formation does not occur immediately or within few days. These oligomers can be therefore referred to be as off-pathway oligomers. On the contrary, in the absence of SDS, since the oligomers formed are much smaller in size (3-6 nm) and probably much more amyloidogenic than the larger species, the fibril formation is much more facile. Hence, these oligomers can be called as on-pathway oligomers.

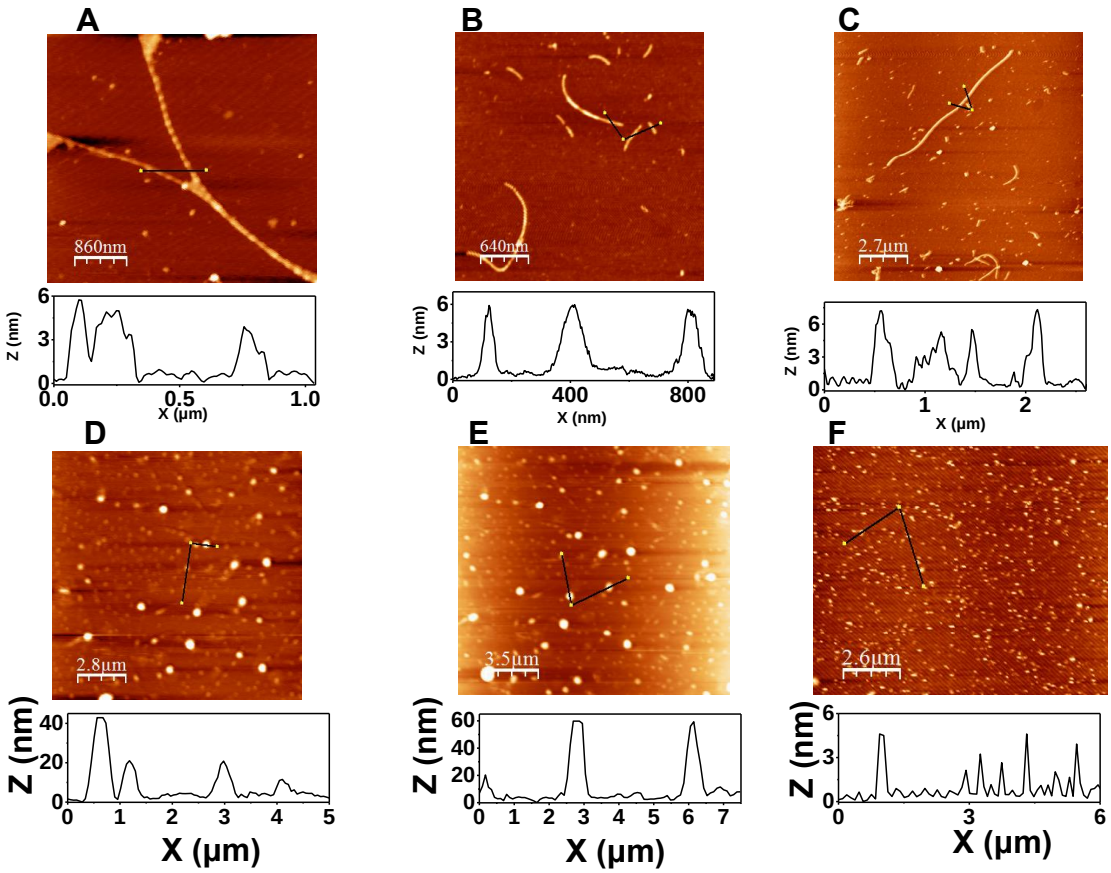


Figure 5.6 AFM images along with the height profiles of κ-casein (A-C) amyloid fibrils formed without SDS and (D-F) spherical oligomeric aggregates formed in the presence of SDS (100 μM).

Collectively, our studies indicate that in the absence of SDS, κ-casein aggregation occurs at a high concentration and follows a nucleation-dependent polymerization pathway. In the presence of SDS, a switch in the aggregation mechanism from nucleation-dependent polymerization to isodesmic polymerization, is observed. The alteration in the aggregation

mechanism also leads to alteration in the morphology of the final aggregate species formed in these two cases, suggesting that the morphology of the protein aggregates might be very strongly coupled to the mechanism of aggregation (Figure 5.7). Since κ-casein is readily available and easy to isolate from milk, therefore, by making use of lipid mimetics/surfactants, we can control the aggregation pathway of κ-casein and modulate the morphology of the aggregates and accordingly, utilize them for various bio-nanomaterials applications. In fact, the fibrillar κ-casein has been used in the past for the encapsulation and controlled delivery of retinoic acid, which was used to promote the differentiation of neuronal cells.[69]

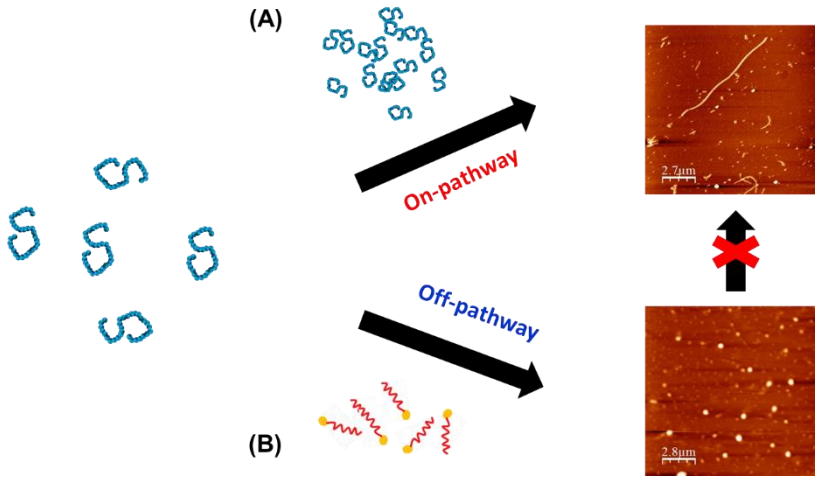


Figure 5.7 Schematic showing two distinct pathways of aggregation followed by κ-casein. (A) Nucleation-dependent polymerization occurring at high concentration leads to fibril formation. (B) Isodesmic polymerization occurring at low protein concentration, in the presence of SDS, leads to the formation of large spherical oligomeric species.

5.4 Summary

Here, we have shown the effect of a well-known lipid mimetic, sodium dodecyl sulfate (SDS) on the mechanism of aggregation of an amyloidogenic IDP, namely, bovine κ-casein using a variety of biophysical tools. The deposition of amyloids formed by κ-casein is linked with mammary cancer. The amyloid formation from κ-casein followed a nucleation dependent polymerization in the absence of SDS, as shown in Chapter 3. In the presence of SDS (sub-micellar concentration), a drastic change in the aggregation kinetics and a switch in the mechanism was observed. Additionally, the morphologies of the aggregates formed in the absence and presence of SDS were found to be different. Without SDS, fibril formation

occurred, whereas, in the presence of SDS, mostly large spherical oligomers were observed. The fibrillar species formed without SDS have higher β-sheet content compared to the spherical oligomeric species formed in the presence of SDS. The formation of large spherical oligomeric species formed in the presence of SDS may have a protective role since these oligomeric species do not readily convert into amyloid fibrils and it has been proposed that in case of κ-casein fibrillation, the dissociation of large oligomeric species is the rate determining step. The interaction of κ-casein with SDS may have physiological relevance as SDS being a membrane mimetic is able to mimic the role of lipids found in milk and milk ducts. Additionally, our studies suggest that the morphology of the protein aggregates can be modulated using surfactants and can be utilized for the designing of novel protein/amyloid based nanomaterials. Thus, the insights gained from our studies on fibrillation mechanism of κ-casein holds promise in the designing of exotic bio-nanomaterials with desired properties. Further studies using other casein proteins, lipids and sugars found in milk would provide a better understanding of the κ-casein fibrillation mechanism, and would be of value to the food industry as well as provide insights into how κ-casein is retained in the monomeric form and fibrillation is prevented inside the mammary glands of bovine.

5.5 References:

1. Nelson, R.; Sawaya, M. R.; Balbirnie, M.; Madsen, A. O.; Riekel, C.; Grothe, R.; Eisenberg, D. Structure of the cross-[beta] spine of amyloid-like fibrils. *Nature* **2005,** *435*, 773-778.

2. Brettschneider, J.; Tredici, K. D.; Lee, V. M. Y.; Trojanowski, J. Q. Spreading of pathology in neurodegenerative diseases: a focus on human studies. *Nat. Rev. Neurosci.* **2015,** *16*, 109-120.

3. Uversky, V. N.; Dave, V.; Iakoucheva, L. M.; Malaney, P.; Metallo, S. J.; Pathak, R. R.; Joerger, A. C. Pathological unfoldomics of uncontrolled chaos: intrinsically disordered proteins and human diseases. *Chem. Rev.* **2014,** *114*, 6844-79.

4. Tipping, K. W.; van Oosten-Hawle, P.; Hewitt, E. W.; Radford, S. E. Amyloid Fibres: Inert End-Stage Aggregates or Key Players in Disease? *Trends Biochem. Sci.* **2015,** *40*, 719-727.

5. Knowles, T. P.; Vendruscolo, M.; Dobson, C. M. The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell Biol.* **2014,** *15*, 384-96.

6. Eisenberg, D.; Jucker, M. The amyloid state of proteins in human diseases. *Cell* **2012,** *148*, 1188-1203.

7. Greenwald, J.; Riek, R. Biology of amyloid: structure, function, and regulation. *Structure* **2010,** *18*, 1244-60.

8. Li, C.; Mezzenga, R. The interplay between carbon nanomaterials and amyloid fibrils in bio-nanotechnology. *Nanoscale* **2013,** *5*, 6207-6218.

9. Bolisetty, S.; Mezzenga, R. Amyloid–carbon hybrid membranes for universal water purification. *Nat. Nanotechnol.* **2016,** *11*, 365-371.

10. Knowles, T. P. J.; Mezzenga, R. Amyloid Fibrils as Building Blocks for Natural and Artificial Functional Materials. *Adv. Mater.* **2016,** *28*, 6546-6561.

11. Cherny, I.; Gazit, E. Amyloids: Not Only Pathological Agents but Also Ordered Nanomaterials. *Angew. Chem. Int. Ed.* **2008,** *47*, 4062-4069.

12. Pearce, F. G.; Mackintosh, S. H.; Gerrard, J. A. Formation of Amyloid-like Fibrils by Ovalbumin and Related Proteins under Conditions Relevant to Food Processing. *J. Agric. Food Chem.* **2007,** *55*, 318-322.

13. Raynes, J. K.; Carver, J. A.; Gras, S. L.; Gerrard, J. A. Protein nanostructures in food – Should we be worried? *Trends Food Sci. Technol.* **2014,** *37*, 42-50.

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

15. Cohen, S. I. A.; Vendruscolo, M.; Dobson, C. M.; Knowles, T. P. J. From Macroscopic Measurements to Microscopic Mechanisms of Protein Aggregation. *J. Mol. Biol.* **2012,** *421*, 160-171.

16. Knowles, T. P. J.; Waudby, C. A.; Devlin, G. L.; Cohen, S. I. A.; Aguzzi, A.; Vendruscolo, M.; Terentjev, E. M.; Welland, M. E.; Dobson, C. M. An Analytical Solution to the Kinetics of Breakable Filament Assembly. *Science* **2009,** *326*, 1533-1537.

17. Bucciantini, M.; Rigacci, S.; Stefani, M. Amyloid Aggregation: Role of Biological Membranes and the Aggregate–Membrane System. *J. Phys. Chem. Lett.* **2014,** *5*, 517-527.

18. Bucciantini, M.; Cecchi, C. Biological Membranes as Protein Aggregation Matrices and Targets of Amyloid Toxicity. In *Protein Misfolding and Cellular Stress in Disease and Aging: Concepts and Protocols*, Bross, P.; Gregersen, N., Eds. Humana Press: Totowa, NJ, 2010; pp 231-243.

19. Sethuraman, A.; Belfort, G. Protein Structural Perturbation and Aggregation on Homogeneous Surfaces. *Biophys. J.* **2005,** *88*, 1322-1333.

20. Zhu, J.; Nathan, C.; Jin, W.; Sim, D.; Ashcroft, G. S.; Wahl, S. M.; Lacomis, L.; Erdjument-Bromage, H.; Tempst, P.; Wright, C. D.; Ding, A. Conversion of Proepithelin to Epithelins: Roles of SLPI and Elastase in Host Defense and Wound Repair. *Cell* **2002,** *111*, 867-878.

21. Gorbenko, G. P.; Kinnunen, P. K. J. The role of lipid–protein interactions in amyloid-type protein fibril formation. *Chem. Phys. Lipids* **2006,** *141*, 72-82.

22. Michikawa, M.; Gong, J.-S.; Fan, Q.-W.; Sawamura, N.; Yanagisawa, K. A Novel Action of Alzheimer's Amyloid β-Protein (Aβ): Oligomeric Aβ Promotes Lipid Release. *J. Neuroscience* **2001,** *21*, 7226.

23. Sparr, E.; Engel, M. F. M.; Sakharov, D. V.; Sprong, M.; Jacobs, J.; de Kruijff, B.; Höppener, J. W. M.; Antoinette Killian, J. Islet amyloid polypeptide-induced membrane leakage involves uptake of lipids by forming amyloid fibers. *FEBS Lett.* **2004,** *577*, 117-120.

24. Zhu, M.; Li, J.; Fink, A. L. The Association of α-Synuclein with Membranes Affects Bilayer Structure, Stability, and Fibril Formation. *J. Biol. Chem.* **2003,** *278*, 40186-40197.

25. Chaudhary, H.; Stefanovic, A. N. D.; Subramaniam, V.; Claessens, M. M. A. E. Membrane interactions and fibrillization of α-synuclein play an essential role in membrane disruption. *FEBS Lett.* **2014,** *588*, 4457-4463.

26. Reynolds, N. P.; Soragni, A.; Rabe, M.; Verdes, D.; Liverani, E.; Handschin, S.; Riek, R.; Seeger, S. Mechanism of Membrane Interaction and Disruption by α-Synuclein. *J. Am. Chem. Soc.* **2011,** *133*, 19366-19375.

27. Lashuel, H. A. Membrane Permeabilization: A Common Mechanism in Protein-Misfolding Diseases. *Sci. Aging Knowledge Environ.* **2005,** *2005*, pe28.

28. Cao, P.; Abedini, A.; Wang, H.; Tu, L.-H.; Zhang, X.; Schmidt, A. M.; Raleigh, D. P. Islet amyloid polypeptide toxicity and membrane interactions. *Proc. Natl. Acad. Sci.* **2013,** *110*, 19279-19284.

29. Goodchild, S. C.; Sheynis, T.; Thompson, R.; Tipping, K. W.; Xue, W.-F.; Ranson, N. A.; Beales, P. A.; Hewitt, E. W.; Radford, S. E. β2-Microglobulin Amyloid Fibril-Induced Membrane Disruption Is Enhanced by Endosomal Lipids and Acidic pH. *PLoS One* **2014,** *9*, e104492.

30. Milanesi, L.; Sheynis, T.; Xue, W.-F.; Orlova, E. V.; Hellewell, A. L.; Jelinek, R.; Hewitt, E. W.; Radford, S. E.; Saibil, H. R. Direct three-dimensional visualization of membrane disruption by amyloid fibrils. *Proc. Natl. Acad. Sci.* **2012,** *109*, 20455-20460.

31. Rivers, R. C.; Kumita, J. R.; Tartaglia, G. G.; Dedmon, M. M.; Pawar, A.; Vendruscolo, M.; Dobson, C. M.; Christodoulou, J. Molecular determinants of the aggregation behavior of α- and β-synuclein. *Protein Sci.* **2008,** *17*, 887-898.

32. Giehm, L.; Oliveira, C. L. P.; Christiansen, G.; Pedersen, J. S.; Otzen, D. E. SDS-Induced Fibrillation of α-Synuclein: An Alternative Fibrillation Pathway. *J. Mol. Biol.* **2010,** *401*, 115-133.

33. Tompa, P.; Fersht, A. *Structure and Function of Intrinsically Disordered Proteins*. Taylor & Francis: 2009.

34. Farrell, H. M.; Kumosinski, T. F.; Malin, E. L.; Brown, E. M. The Caseins of Milk as Calcium-Binding Proteins. In *Calcium-Binding Protein Protocols: Volume 1: Reviews and Case Studies*, Vogel, H. J., Ed. Humana Press: Totowa, NJ, 2002; pp 97-140.

35. De Kruif, C. G.; Holt, C. Casein Micelle Structure, Functions and Interactions. In *Advanced Dairy Chemistry—1 Proteins: Part A / Part B*, Fox, P. F.; McSweeney, P. L. H., Eds. Springer US: Boston, MA, 2003; pp 233-276.

36. Bhattacharyya, J.; Das, K. P. Molecular Chaperone-like Properties of an Unfolded Protein, αs-Casein. *J. Biol. Chem.* **1999,** *274*, 15505-15509.

37. Zhang, X.; Fu, X.; Zhang, H.; Liu, C.; Jiao, W.; Chang, Z. Chaperone-like activity of β-casein. *Int. J. Biochem. Cell Biol.* **2005,** *37*, 1232-1240.

38. Morgan, P. E.; Treweek, T. M.; Lindner, R. A.; Price, W. E.; Carver, J. A. Casein Proteins as Molecular Chaperones. *J. Agric. Food Chem.* **2005,** *53*, 2670-2683.

39. Yong, Y. H.; Foegeding, E. A. Caseins: Utilizing Molecular Chaperone Properties to Control Protein Aggregation in Foods. *J. Agric. Food Chem.* **2010,** *58*, 685-693.

40. Holt, C.; Carver, J. A.; Ecroyd, H.; Thorn, D. C. Invited review: Caseins and the casein micelle: Their biological functions, structures, and behavior in foods1. *J. Dairy Sci.* **2013,** *96*, 6127-6146.

41. Farrell, H. M.; Cooke, P. H.; Wickham, E. D.; Piotrowski, E. G.; Hoagland, P. D. Environmental Influences on Bovine κ-Casein: Reduction and Conversion to Fibrillar (Amyloid) Structures. *J. Protein Chem.* **2003,** *22*, 259-273.

42. Thorn, D. C.; Ecroyd, H.; Sunde, M.; Poon, S.; Carver, J. A. Amyloid Fibril Formation by Bovine Milk αs2-Casein Occurs under Physiological Conditions Yet Is Prevented by Its Natural Counterpart, αs1-Casein. *Biochemistry* **2008,** *47*, 3926-3936.

43. Thorn, D. C.; Meehan, S.; Sunde, M.; Rekas, A.; Gras, S. L.; MacPhee, C. E.; Dobson, C. M.; Wilson, M. R.; Carver, J. A. Amyloid Fibril Formation by Bovine Milk κ-Casein and Its Inhibition by the Molecular Chaperones αS- and β-Casein. *Biochemistry* **2005,** *44*, 17027-17036.

44. Ecroyd, H.; Thorn, David C.; Liu, Y.; Carver, John A. The dissociated form of κ-casein is the precursor to its amyloid fibril formation. *Biochem. J.* **2010,** *429*, 251-260.

45. Sokolovski, M.; Sheynis, T.; Kolusheva, S.; Jelinek, R. Membrane interactions and lipid binding of casein oligomers and early aggregates. *Biochim. Biophys. Acta* **2008,** *1778*, 2341-2349.

46. MacGibbon, A. K. H.; Taylor, M. W. Composition and Structure of Bovine Milk Lipids. In *Advanced Dairy Chemistry Volume 2 Lipids*, Fox, P. F.; McSweeney, P. L. H., Eds. Springer US: Boston, MA, 2006; pp 1-42.

47. Liu, J.; Dehle, F. C.; Liu, Y.; Bahraminejad, E.; Ecroyd, H.; Thorn, D. C.; Carver, J. A. The Effect of Milk Constituents and Crowding Agents on Amyloid Fibril Formation by κ-Casein. *J. Agric. Food Chem.* **2016,** *64*, 1335-1343.

48. Sacchettini, J. C.; Kelly, J. W. Therapeutic strategies for human amyloid diseases. *Nat. Rev. Drug Discov.* **2002,** *1*, 267-275.

49. Rhie, A.; Kirby, L.; Sayer, N.; Wellesley, R.; Disterer, P.; Sylvester, I.; Gill, A.; Hope, J.; James, W.; Tahiri-Alaoui, A. Characterization of 2′-Fluoro-RNA Aptamers That Bind Preferentially to Disease-associated Conformations of Prion Protein and Inhibit Conversion. *J. Biol. Chem.* **2003,** *278*, 39697-39705.

50. Ryan, T. M.; Griffin, M. D. W.; Teoh, C. L.; Ooi, J.; Howlett, G. J. High-affinity amphipathic modulators of amyloid fibril nucleation and elongation. *J. Mol. Biol.* **2011,** *406*, 416-429.

51. Ryan, T. M.; Friedhuber, A.; Lind, M.; Howlett, G. J.; Masters, C.; Roberts, B. R. Small Amphipathic Molecules Modulate Secondary Structure and Amyloid Fibril-forming

Kinetics of Alzheimer Disease Peptide Aβ1–42. *J. Biol. Chem.* **2012,** *287*, 16947-16954.

52. Arya, S.; Mukhopadhyay, S. Ordered water within the collapsed globules of an amyloidogenic intrinsically disordered protein. *J. Phys. Chem. B* **2014,** *118*, 9191-9198.

53. Jain, N.; Bhattacharya, M.; Mukhopadhyay, S. Chain Collapse of an Amyloidogenic Intrinsically Disordered Protein. *Biophys. J.* **2011,** *101*, 1720-1729.

54. Arya, S.; Kumari, A.; Dalal, V.; Bhattacharya, M.; Mukhopadhyay, S. Appearance of annular ring-like intermediates during amyloid fibril formation from human serum albumin. *Phys. Chem. Chem. Phys.* **2015,** *17*, 22862-22871.

55. Arora, A.; Ha, C.; Park, C. B. Inhibition of insulin amyloid formation by small stress molecules. *FEBS Lett.* **2004,** *564*, 121-125.

56. Horcas, I.; Fernández, R.; Gómez-Rodríguez, J. M.; Colchero, J.; Gómez-Herrero, J.; Baro, A. M. WSXM: A software for scanning probe microscopy and a tool for nanotechnology. *Rev. Sci. Instrum.* **2007,** *78*, 013705.

57. Ecroyd, H.; Thorn, David C.; Liu, Y.; Carver, John A. The dissociated form of κ-casein is the precursor to its amyloid fibril formation. *Biochem. J.* **2010,** *429*, 251.

58. Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*. Springer US: 2007.

59. Daniel, E.; Weber, G. Cooperative Effects in Binding by Bovine Serum Albumin. I. The Binding of 1-Anilino-8-naphthalenesulfonate. Fluorimetric Titrations*. *Biochemistry* **1966,** *5*, 1893-1900.

60. Krishnan, R.; Lindquist, S. L. Structural insights into a yeast prion illuminate nucleation and strain diversity. *Nature* **2005,** *435*, 765-772.

61. Mukhopadhyay, S.; Krishnan, R.; Lemke, E. A.; Lindquist, S.; Deniz, A. A. A natively unfolded yeast prion monomer adopts an ensemble of collapsed and rapidly fluctuating structures. *Proc. Natl. Acad. Sci. U. S. A.* **2007,** *104*, 2649-2654.

62. Abelein, A.; Kaspersen, J. D.; Nielsen, S. B.; Jensen, G. V.; Christiansen, G.; Pedersen, J. S.; Danielsson, J.; Otzen, D. E.; Gräslund, A. Formation of Dynamic Soluble Surfactant-induced Amyloid β Peptide Aggregation Intermediates. *J. Biol. Chem.* **2013,** *288*, 23518-23528.

63. Frieden, C. Protein aggregation processes: In search of the mechanism. *Protein Sci.* **2007,** *16* (11), 2334-2344.

64. Lee, J.; Culyba, E. K.; Powers, E. T.; Kelly, J. W. Amyloid-beta forms fibrils by nucleated conformational conversion of oligomers. *Nat. Chem. Biol.* **2011,** *7*, 602-609.

65. Bhattacharya, M.; Jain, N.; Dogra, P.; Samai, S.; Mukhopadhyay, S. Nanoscopic Amyloid Pores Formed via Stepwise Protein Assembly. *J. Phys. Chem. Lett.* **2013,** *4*, 480-485.

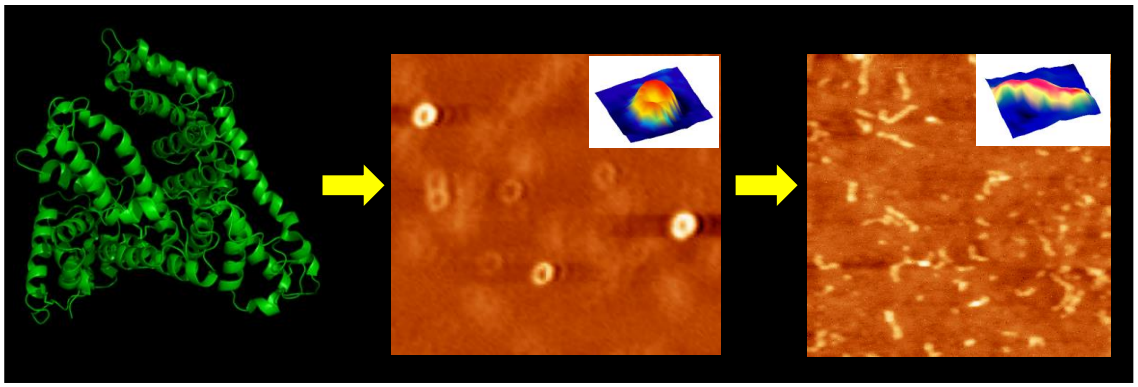
66. Usov, I.; Adamcik, J.; Mezzenga, R. Polymorphism in bovine serum albumin fibrils: morphology and statistical analysis. *Faraday Discuss.* **2013,** *166*, 151-162.

67. Volpatti, L. R.; Vendruscolo, M.; Dobson, C. M.; Knowles, T. P. J. A Clear View of Polymorphism, Twist, and Chirality in Amyloid Fibril Formation. *ACS Nano* **2013,** *7*, 10443-10448.

68. Sweers, K. K. M.; van der Werf, K. O.; Bennink, M. L.; Subramaniam, V. Atomic Force Microscopy under Controlled Conditions Reveals Structure of C-Terminal Region of α-Synuclein in Amyloid Fibrils. *ACS Nano* **2012,** *6*, 5952-5960.

69. Chun, J.; Bhak, G.; Lee, S.-G.; Lee, J.-H.; Lee, D.; Char, K.; Paik, S. R. κ-Casein-Based Hierarchical Suprastructures and Their Use for Selective Temporal and Spatial Control over Neuronal Differentiation. *Biomacromolecules* **2012,** *13*, 2731-2738.

Appearance of Annular Ring-like Intermediates during Amyloid Fibril Formation from Human Serum Albumin



The work described in this chapter has been published in *Physical Chemistry Chemical Physics*.

Reference: Arya et al. *Phys. Chem. Chem. Phys.* **2015,** *17*, 22862--22871.

6.1 Introduction

Amyloid fibrils are ordered protein aggregates and are characterized by a cross-β motif which is comprised of β-strands that run perpendicular to the fibril axis and are arranged in hydrogen-bonded ribbon-like β-sheets.[1–7] The conversion of soluble monomeric protein into intractable amyloid aggregates is implicated in a multitude of human disorders that include Alzheimer's disease, Parkinson's disease, type II diabetes, and the spongiform encephalopathies.[8–16] Owing to the extraordinary supramolecular architectural characteristics possessed by these amyloids, they are not just limited to diseases but have also been discovered in living organisms to perform an array of important biological functions.[17–19] In addition, the robust mechanical strength possessed by amyloids makes them attractive candidates for nanotechnological applications as bionanomaterials.[20–23] On the pathway of amyloid formation, various intermediates with different morphologies have been observed. An increasing body of evidence suggests that these oligomeric intermediates might be more cytotoxic than the matured fibrils.[15,24] The morphology of these intermediates dictates the toxicity of the amyloid species and various morphologies have been observed for oligomeric as well as fibrillar species. The annular pore-like nanostructures are one of the most commonly observed morphologies that have been shown to exert cytotoxic effects in neurodegenerative diseases by virtue of disrupting the biological membranes.[25–31] Thus, the morphologies of these oligomeric intermediates might play a very crucial role in the pathogenesis of amyloid diseases. Therefore, understanding the morphological as well as conformational characteristics of the oligomeric intermediates and the underlying molecular mechanism of amyloid formation is important not only for designing better therapeutics to target amyloid disorders but also for developing novel bionanomaterials. Most proteins, if not all, are capable of undergoing structural transition from their native state to amyloid fibrils under certain experimental conditions in vitro, suggesting that the amyloid formation might be a generic property of polypeptides.[16] The mechanism of amyloid formation has been investigated in a variety of model proteins such as lysozyme, β-lactoglobulin, β_2-microglobulin, insulin, islet amyloid polypeptide and serum albumins.[31–40] Amongst these proteins, human serum albumin (HSA) is an all α-helical protein and has been considered a good model system for protein aggregation studies due to its propensity to easily aggregate under in vitro conditions.

HSA is a single polypeptide chain consisting of 585 amino acids and is composed of three homologous domains I–III and each domain has two sub-domains, A and B (Figure 6.1A).[41,42] It has physiological importance as a carrier protein in transporting fatty acids, drugs and other

organic compounds in the circulatory system.[41,43] Several studies on HSA have shown that varying certain conditions like pH, ionic strength, and temperature can have a drastic effect on the native structure of HSA, for example, low pH and high temperature lead to the destabilization of its native structure.[44–47] The higher temperature mimics the heat stress condition that leads to heat shock or unfolded protein response which results in protein aggregation and imbalance of protein homeostasis.[48] Also, the high temperature provides protein molecules with a sufficient amount of energy to overcome the high activation barrier thereby accelerating the aggregation reaction and making it more accessible to the laboratory timescale. In the case of a multi-domain protein like HSA, the increase in temperature might lead to partial destabilization of its well-organized native structure. The resultant partially destabilized state might have resemblance with the disordered state of proteins having a higher propensity to self-assemble than the native form that otherwise lacks any structural attributes that suggest predilection towards amyloid formation.[9,49–51] HSA aggregation has been investigated under different pH, temperature and solution conditions using a variety of tools.[52–56] However, an in-depth molecular insight into the structural transition from an α-helical native to cross β sheet-rich amyloid state via oligomeric intermediates still remains elusive. In this work, we have utilized steady-state as well as time-resolved fluorescence spectroscopy, circular dichroism and atomic force microscopy to unravel the mechanism of HSA fibrillation. Efforts were directed towards elucidating the protein conformational - and size changes as a function of time with an emphasis on delineating the key structural transitions in a domain-specific manner. Additionally, Raman spectroscopic measurements at various aggregation stages provided a wealth of structural insights into alterations in the protein backbone amides and side chain residues as a function of amyloid assembly.

6.2 Experimental Section

6.2.1 Materials

Albumin from human serum (HSA), 8-anilino-1-naphthalenesulfonic acid ammonium salt (ANS), thioflavin-T (ThT), and sodium hydrogen phosphate (monobasic) were obtained from Sigma (St. Louis, MO). 6-Acryloyl-2 (dimethylamino) naphthalene (acrylodan) and 5-((((2-iodoacetyl)amino)ethyl)amino) naphthalene-1-sulfonic acid (IAEDANS) were procured from Molecular Probes, Invitrogen Inc. Milli-Q water was used for the preparation of all solutions. The pH of the buffers was adjusted to ~7.4 using a Metrohm pH meter at ~25 ^{0}C.

6.2.2 Aggregation reaction

The HSA stock solution was prepared in 10 mM phosphate buffer (pH 7.4). The accurate protein concentration was estimated by measuring the absorbance at 280 nm using a molar absorption coefficient (ε) of 35219 M^{-1} cm^{-1}.[52] For the aggregation reaction, the stock solution was diluted using the same buffer containing 50 mM NaCl to a final protein concentration of 100 µM. The resultant solution was heated at 65 ± 2 ^{0}C in a heating block, preset at the required temperature under quiescent conditions. Aliquots were taken out at regular intervals for the kinetics measurements. For aggregation studies using labeled HSA, the labeled protein concentration was 5% of the total protein concentration (100 µM) in the aggregation reaction mixture.

6.2.3 Arrhenius plot

In order to monitor the temperature dependence of the aggregation reaction using ThT fluorescence, the reaction was carried out at 55 ^{0}C, 65 ^{0}C, 70 ^{0}C and 75 ^{0}C. ThT fluorescence observed under different temperature conditions could be fitted to single exponential. The activation energy was obtained using the Arrhenius equation,

$$k = Aexp(\frac{-E_a}{RT}) \qquad (1)$$

where A is the pre-exponential factor, T is the absolute temperature, R is the universal gas constant and E_a is the activation energy. The slope of ln(k) versus 1/T plot yielded the activation energy (E_a) for the reaction.

6.2.4 Circular dichroism (CD) measurements

The CD spectra were acquired using a Chirascan Spectrophotometer (Applied Photophysics, UK) in a 1 mm path length quartz cell with a scan range of 200-260 nm and 1 nm as step size. The spectra were corrected for the buffer signal and were plotted in Origin 8.5 software.

6.2.5 Atomic force microscopy (AFM) imaging

AFM images of oligomers and amyloid fibrils were acquired on MultiView 2000^{TM} instrument (Nanonics Imaging Ltd, Jerusalem, Israel) operating in intermittent contact mode based on the phase feedback technique. For AFM imaging, the aliquots were withdrawn after 10 min and 2 h. The sample withdrawn after 10 min was diluted 4000-fold and the one which was withdrawn after 2 h was diluted 2000-fold using an aggregation reaction buffer filtered through a 0.22 µm syringe filter. The sample (10 µL in volume) was deposited on freshly cleaved and buffer-

washed muscovite mica (Grade V-4 mica from SPI, PA). The mica surface was incubated with the sample for 20 min before washing it with filtered Milli-Q water and dried under a gentle stream of nitrogen gas. AFM scanning was performed under ambient environmental conditions using a Cr-coated cantilevered glass probe of diameter ~10 nm oscillating at a resonance frequency of ~32 kHz. AFM height images of 400 × 400 pixels in resolution, covering areas of 10 µm × 10 µm were collected with a sample delay of 4 ms. Smaller areas (2.5 µm × 2.5 µm or 1.45 µm × 1.45 µm) were scanned to image finer morphological features. The AFM images were acquired using the Quartz software (provided with MultiView) and were processed using the WSxM software provided with our AFM instrument.

6.2.6 Fluorescence labeling of HSA with acrylodan and IAEDANS

The labeling of the free cysteine in domain I with acrylodan and IAEDANS was carried out separately in 10 mM phosphate buffer (pH 7.4) according to the protocol described previously.[57] Briefly, acrylodan/IAEDANS and HSA were mixed in a 10 : 1 molar ratio and allowed to stir in the dark for 24 h at room temperature. Stock solutions of acrylodan (in acetonitrile) and IAEDANS (in dimethyl sulfoxide (DMSO)) were prepared freshly. The labeled protein was then passed through a PD-10 column to remove any unreacted dye and was further concentrated using AMICON ultra (30 kDa cutoff; from Millipore). Finally, the reaction mixture was dialyzed against 10 mM phosphate buffer (pH 7.4) over a 4 day period at 4 ^{0}C; with buffer change after every 12 h. The concentration of the labeled protein was estimated using ε365 = 12800 M^{-1} cm^{-1}, for acrylodan[58] and ε337 = 6100 M^{-1} cm^{-1} for AEDANS.[59] Steady-state fluorescence measurements. All the steady-state fluorescence measurements were performed on a FluoroMax-4 spectrofluorometer (Horiba Jobin Yvon, NJ) at ~25 ^{0}C. The aliquots of the protein sample were taken out at different time intervals and were diluted 5-fold in 10 mM phosphate buffer (pH 7.4) for all the steady-state fluorescence measurements, except for the AEDANS-labeled protein. The AEDANS samples were not diluted before the fluorescence measurements. The samples were excited at 295 nm (Trp), 350 nm (ANS), 360 nm (acrylodan), 337 nm (AEDANS) and 450 nm (ThT). The concentration of ANS and ThT used for ANS and ThT binding experiments, respectively, was 10 µM. The steady-state fluorescence anisotropies were measured at 350 nm (Trp), 470 nm (ANS), 500 nm (acrylodan) and 475 nm (AEDANS). The steady-state fluorescence anisotropy (r_{ss}) is given by the following relationship:[60,61]

$$r_{ss} = \frac{I_{\parallel} - I_{\perp} G}{I_{\parallel} + 2 I_{\perp} G} \qquad (2)$$

where $I_{\parallel}$ and $I_{\perp}$ are fluorescence intensities collected using parallel and perpendicular geometry, respectively. The perpendicular components were always corrected using a G-factor.

6.2.7 Time-resolved fluorescence measurements

All the time-resolved fluorescence data were acquired using a time-correlated single photon counting (TCSPC) setup (Fluorocube, Horiba Jobin Yvon, NJ). For fluorescence lifetime measurements, the peak count was fixed to 10,000 and the excitation and emission polarizers were oriented at the magic angle (54.7^0). For time-resolved fluorescence anisotropy measurements, the peak difference was 10,000 counts and the orientation of the emission polarizer was 0^0 and 90^0 with respect to the excitation polarizer for parallel fluorescence intensities ($I_{\parallel}$) and perpendicular fluorescence intensities ($I_{\perp}$), respectively. The emission monochromator for AEDANS was fixed at 500 nm with a bandpass of 4 nm. The 375 nm laser diode was used as a light source for the excitation of AEDANS. The instrument response function (IRF) was collected using Ludox (colloidal silica). The width (FWHM) of IRF was ~250 ps. The anisotropy decays were analyzed by globally fitting $I_{\parallel}(t)$ and $I_{\perp}(t)$.[60] Briefly, the fluorescence intensity decay curves were deconvoluted taking IRF into account using

$$I(t) = \Sigma_i \alpha_i e^{-t/\tau_i} \qquad (3)$$

where α_i and τ_i represent the contributions and lifetime of the different lifetime components, respectively. The fitted parameters were then used as input for the anisotropy decay analysis and $I_{\parallel}(t)$ and $I_{\perp}(t)$ were fitted using:

$$I_{\parallel}(t) = \frac{I(t)[1+ 2r(t)}{3} \qquad (4)$$

$$I_{\perp}(t) = \frac{I(t)[1- r(t)}{3} \qquad (5)$$

where $I_{\parallel}(t)$ and $I_{\perp}(t)$ represent the parallel and perpendicular fluorescence intensities, respectively. I(t) represents the fluorescence intensity at magic angle (54.7^0) and r(t) is the anisotropy decay function. The perpendicular fluorescence intensities were corrected for G-factor which was obtained for AEDANS using ANS in 100% ethanol. The anisotropy decays were then analyzed using a biexponential anisotropy decay function that describes the fast (local) and slow (global) motion of a fluorophore.[60,62]

$$r(t) = r_0[\beta_{fast} \exp\left(-\frac{t}{\Phi_{fast}}\right) + \beta_{slow} \exp\left(-\frac{t}{\Phi_{slow}}\right)] \qquad (6)$$

where r(0) represents the intrinsic anisotropy of a fluorophore; Φ_{fast} and Φ_{slow} represent the fast and slow rotational correlation times, respectively, with β_{fast} and β_{slow} as their respective amplitudes. For estimating r(0) for all the fluorophores, the fluorophores were taken in 100% glycerol and the decays were collected. The r(0) so obtained was used for analyzing the anisotropy data with a window of ± 0.005. The slow correlation time represents the global motion of a fluorophore and hence can be used for estimating the size of a molecule using the Stokes–Einstein relationship given by

$$\Phi_{slow} = \frac{\eta V}{k_B T} \qquad (7)$$

where η is the viscosity, k_B is the Boltzmann constant, T is the temperature which was kept as 25 ^{0}C for all the experiments and V is the molecular volume given by $V = (4/3)\pi R_h^3$ where R_h is the hydrodynamic radius of the molecule.

6.2.8 Raman spectroscopy

Raman spectra of the samples were collected on an inVia Raman microscope with 180^0 scattering geometry (Renishaw, UK). The samples were prepared freshly. The aliquots were withdrawn at regular time intervals and the samples were precipitated with acetone (100% v/v). The precipitates were then deposited on a glass slide covered with aluminium foil. A HPNIR laser with 785 nm wavelength was used as an excitation source and focused into the sample spot using a 50 × objective lens (Nikon, Japan). The scattered light was collected through the same objective and was allowed to pass through an edge filter (785 nm) in order to remove the Rayleigh scattering. A 1200 l per mm grating was used for the dispersion of the scattered light and the signal was finally detected using an air-cooled CCD detector. The data acquisition was done using Wire 3.1 software provided with our Raman spectrometer. All the spectra were averaged over 50 scans with an exposure time of 10 s for the spectral range 400-2500 cm^{-1}, 1600-1700 cm^{-1} and 1200-1300 cm^{-1}. All the data were corrected for tilt in the baseline using the cubic spline interpolation method and smoothened in Wire 3.1 software. The baseline corrected and smoothened Raman spectra were plotted in Origin 8.5 software. The deconvolution of Raman spectra was done in Origin 8.5 software as described in our previous publication.[63] The spectral range for the estimation of the secondary structure content was 1620-1700 cm^{-1} (amide I) and 1220-1300 cm^{-1} (amide III). The peaks were first analyzed by the peak analyzer option in Origin 8.5 software. Mainly three bands were observed at 1640-1700 cm^{-1} and 1220-1300 cm^{-1} that correspond to α-helix (1645-1654 cm^{-1}, 1264-1272 cm^{-1}), β-sheet (1665-1674 cm^{-1}, 1230-1240 cm^{-1}) and coils, turns or extended conformation (1680-

1690 cm^{-1}, 1245-1255 cm^{-1}). These bands were then fitted using multiple peak-fitting with the Gaussian function.

6.3 Results and Discussion

6.3.1 Kinetics of amyloid fibril formation and secondary structural changes

The progression of HSA fibrillation at high temperature was monitored by thioflavin-T (ThT) which is a well-known amyloid marker.[64] A sharp increase in ThT fluorescence, devoid of any lag phase, was observed within 1 h of incubation indicating a rapid structural transition from the native α-helical state to the amyloidogenic state (Figure 6.1B). Typically, amyloid formation via the nucleation-dependent polymerization mechanism exhibits a lag phase followed by an assembly phase.[14,50] The absence of a lag phase in HSA fibrillation suggests that the nuclei are formed immediately upon incubation at high temperature which is in contrast to that observed during a nucleation-dependent amyloid assembly. This observation might indicate a ‘‘downhill polymerization’’ mechanism[50] and is in agreement with the previous reports on HSA fibrillation.[52] However, the reaction rate observed by us was faster. A single exponential fit of the time-course of ThT fluorescence was adequate to describe the fibril formation kinetics suggesting a quasi-two-state transition. Additionally the observed rate of fibrillation, monitored by ThT fluorescence, increased with an increase in temperature and a linear Arrhenius plot was obtained (Figure 6.1C). The activation energy estimated (eqn (1) in Experimental Section) from the slope of the Arrhenius plot was found to be ~157 kJ mol^{-1} (~37.5 kcal mol^{-1}) which is higher than that of bovine serum albumin.[65]

The CD spectrum of native HSA shows two minima at 208 and 222 nm indicating the characteristics of an α-helical conformation and is consistent with its crystal structure.[41,42] As the aggregation reaction progressed, a significant decrease in the ellipticity at 222 nm (θ_{222}) was observed (Figure 6.1D).We also observed an increase in $\theta_{215}/\theta_{222}$ and $\theta_{205}/\theta_{222}$ as a function of time that signified the emergence of β-sheet and disordered structures, respectively, at the expense of the α-helices during amyloid formation. The CD signature of the amyloid fibrils remained unchanged even after prolonged incubation at room temperature. Next, AFM imaging was performed to gain insights into the nanoscale morphology of early oligomeric intermediates and amyloid fibrils.

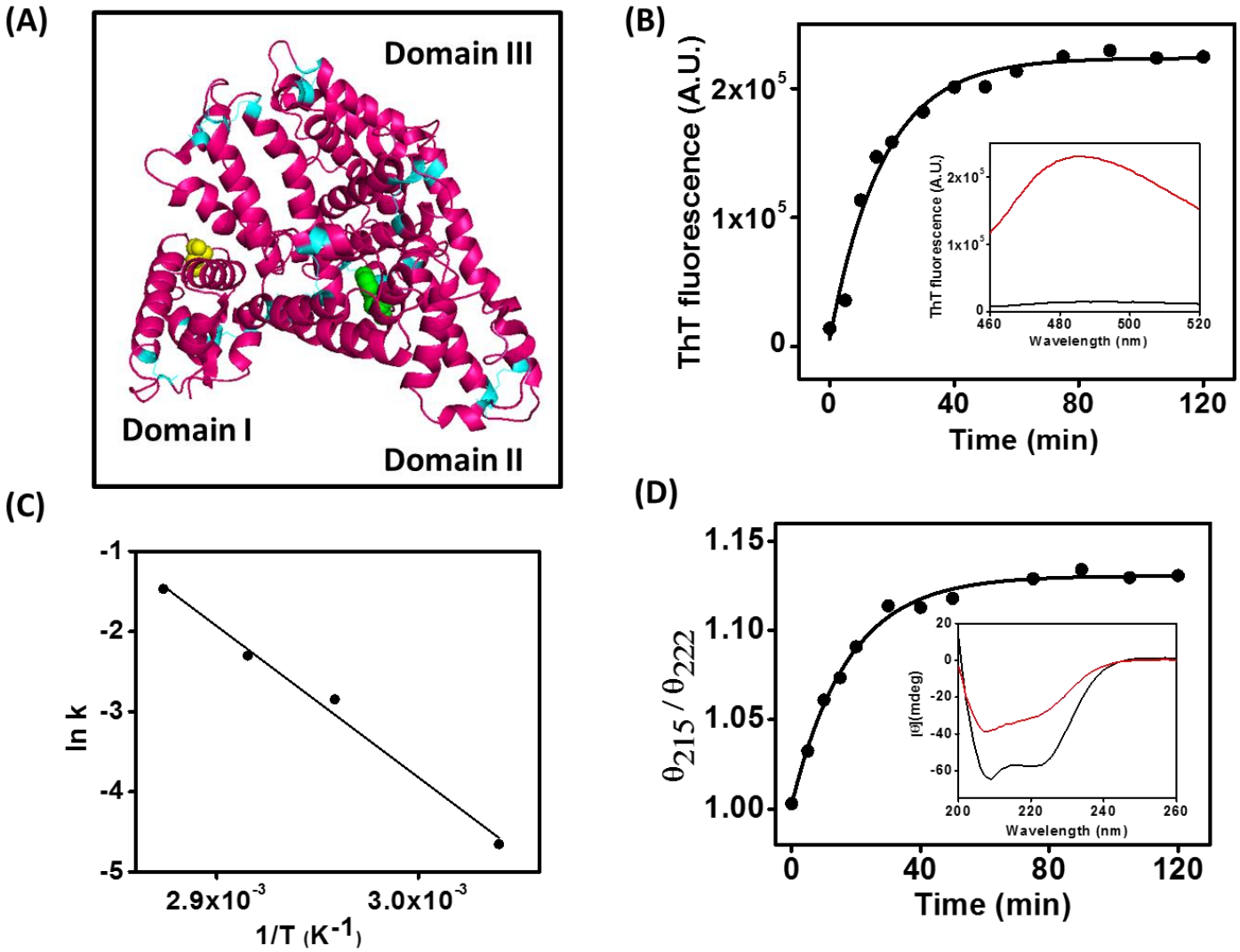


Figure 6.1 (A) The crystal structure of human serum albumin (PDB ID: 1UOR) generated using PyMol (DeLano Scientific, CA). The disulfides, tryptophan (in domain II) and the free cysteine (in domain I) are shown in cyan, green and yellow, respectively. The amyloid formation kinetics monitored by (B) ThT fluorescence and (D) CD. The continuous black lines through the data points represent the fits obtained using a single-exponential kinetic model (the inset in both the figures shows the spectra for zero time point (black) and final time point (red) of fibrillation). (C) Arrhenius plot (ln k versus 1/T) showing temperature dependence of HSA aggregation. The activation energy calculated from the slope of the plot was ~157 kJ/mol (~37.5 kcal/mol).

6.3.2 Nanoscale morphological transitions

Atomic force microscopy (AFM) is a non-invasive and chemical free imaging technique which has emerged as a promising tool for investigating nanoscale topographical features of protein amyloids.[66–68]

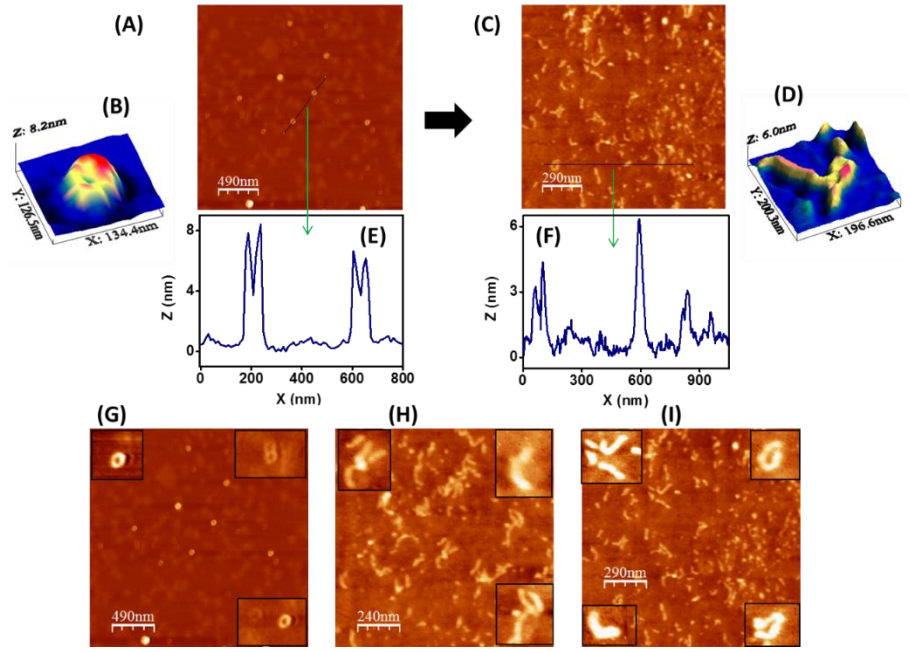


Figure 6.2 AFM images and the height profiles of (A) oligomeric ring-like structures formed after 10 min (2.5 mm × 2.5 mm area) and (C) worm-like fibrils formed after 2 h incubation at 65 ^{0}C (1.45 mm × 1.45 mm area). The 3D images of (B) a single ring-like structure and (D) fibrils show the topography of the observed morphologies. (E and F) The AFM line-profiles of the images are shown in height plots. (G–I) The AFM images showing transition of oligomeric ring-like structures (with figure G insets showing the individual as well as annealed ring-like structures) to worm-like fibrils (with figure H and I insets showing the transition from ring-like intermediates to fibrils).

In our study, we utilized AFM in a time-dependent manner to visualize the morphologies of early oligomeric species and their subsequent transition to matured fibrillar species (Figure 6.2). Therefore, prior to incubating the sample at elevated temperature, AFM imaging was performed. We observed small spherical particles of heights ranging from 2-4 nm and upon statistical analysis, a unimodal height distribution centered at ~2.5 nm was obtained (Figure 6.3A and C) indicating that even at 100 µM concentration, HSA remained predominantly monomeric. Upon incubation at higher temperature, AFM imaging revealed a heterogeneous co-existence of nanoscopic ring-like aggregates and protofibrils (Figure 6.2A and Figure 6.3B).

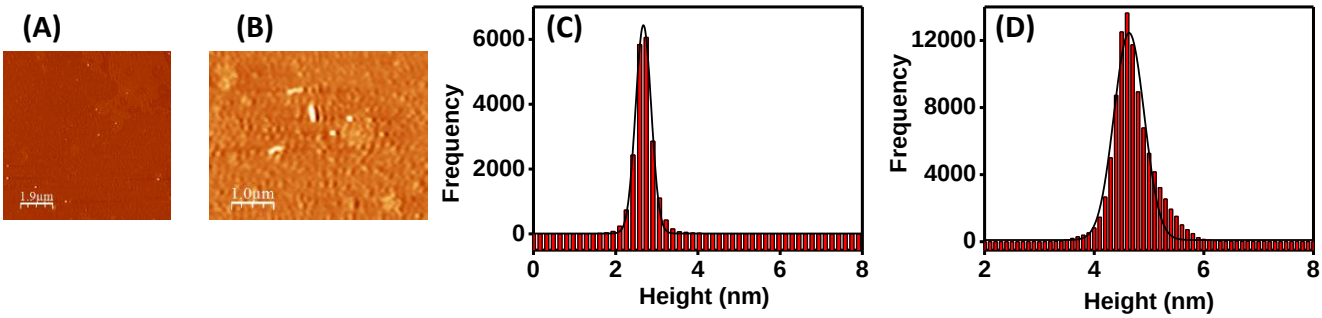


Figure 6.3 AFM images of (A) zero time point (0 minute) sample and (B) early oligomers/protofibrils (10 minute). The statistical height distributions of (C) zero time point (0 minute) sample and (D) oligomers (10 minute).

The statistical analysis of the topography of ring-like species revealed a height range of 2-8 nm (Figure 6.3D), similar to that of protofibrils (Figure 6.4), and an inner diameter of 30-50 nm (Figure 6.4). The 3D image of a single ring has been shown for clarity (Figure 6.2B). It is possible that the smaller rings anneal to form protofibrils (Figure 6.2G inset) that subsequently anneal to form bigger rings. At later stages of aggregation (≥ 2 h), a mixed population of crescent-shaped intermediates, sliced rings and worm-like fibrils was observed (Figure 6.2C, H and I and insets). Such morphological diversity among HSA aggregates in addition to a predominant population of worm-like fibrils suggested that probably, the early ring-like intermediates unravel to form curly worm like fibrils (Figure 6.2H and I and insets). The height of these fibrils ranged between 3 and 6 nm and the average diameter for the sliced rings was 40-50 nm. These fibrils do not undergo any further growth and sliced rings were evident even after prolonged incubation (Figure 6.4A). Therefore, a careful investigation of HSA aggregation using AFM indicated a complex process comprising a time-dependent, stepwise evolution of morphologically-diverse nanoscopic species on the pathway to amyloid fibrils. At first, the predominantly monomeric HSA molecules associate and transform into ring-like oligomeric intermediates that are apparently metastable. Subsequently, these ring-like species undergo further transition to worm-like fibrils presumably, by a ring-opening mechanism. The proposed mechanism might hold true for many disease-related proteins that form ring-like nanostructures in the intermediate stages of amyloid aggregation.[27–31]

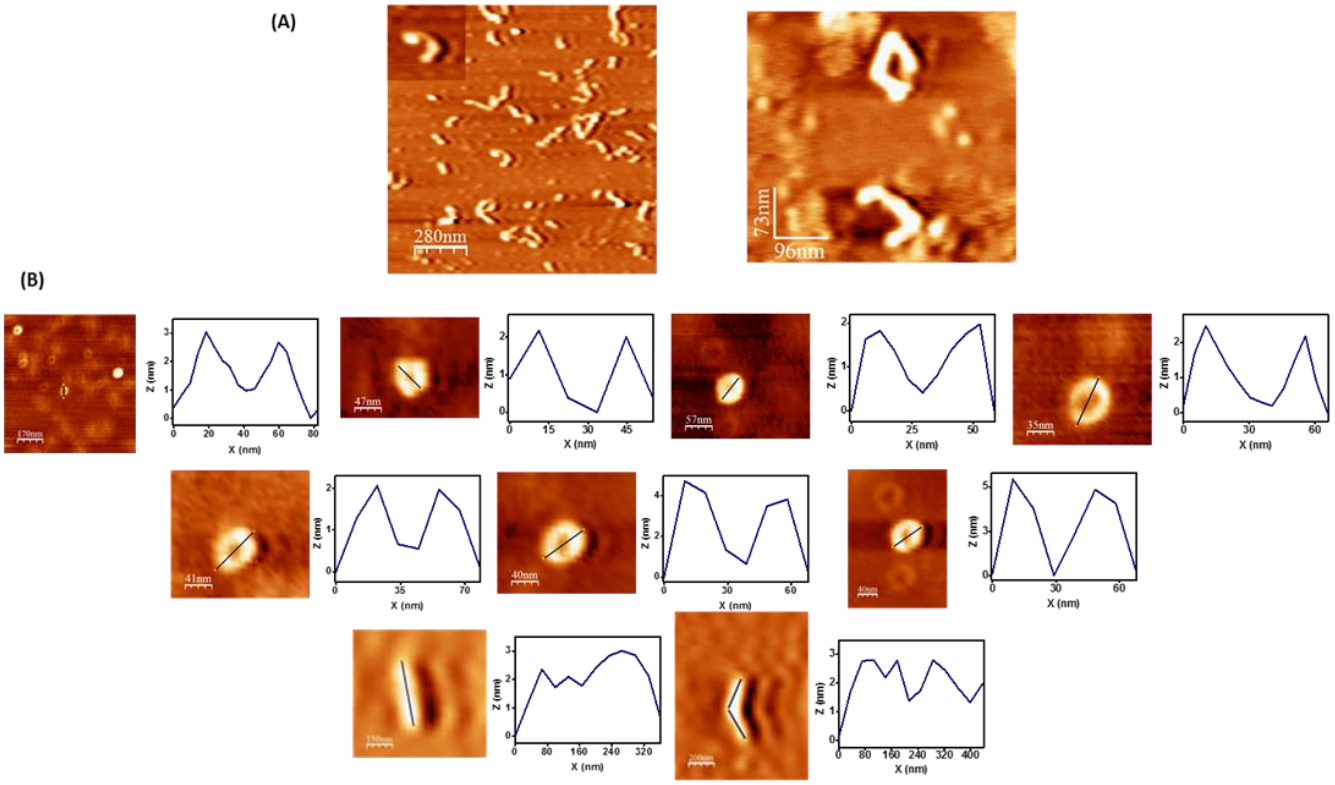


Figure 6.4 AFM images of (A) matured HSA fibrils formed upon prolonged incubation (for one month at room temperature) and (B) individual rings (along with their height profiles). The height profiles of the individual rings were analyzed using WSxM software.

6.3.3 Conformational and dynamical changes during amyloid formation

After establishing the morphological characteristics of various nanoscopic species *en-route* to HSA amyloid fibrillation, next we embarked upon probing the protein structural changes using fluorescence spectroscopy with an emphasis on delineating the conformational changes in various domains. Fluorescence spectroscopy is a highly sensitive and multi-parametric technique for studying protein conformational and dynamical changes in a site-specific manner whereby intrinsic and/or extrinsic fluorophores can be employed. The steady-state fluorescence spectrum provides insights into the local environment around the fluorophore whereas the steady state fluorescence anisotropy provides information about an overall size of the protein with a presumption that the overall size changes predominate over the local dynamics.[60,61] HSA contains a free cysteine (Cys 34) in domain I and a single tryptophan (Trp 214) in domain II (Figure 6.1A).[41–43] In order to probe the conformational changes around domain I during HSA fibrillation, the free thiol group of Cys 34 was covalently modified with two distinct extrinsic fluorophores namely, acrylodan and IAEDANS whereas the conformational changes around domain II were probed by Trp which is an intrinsic fluorophore. Additionally, time-resolved

fluorescence decay measurements provided quantitative estimates about the average sizes of the HSA aggregates.

6.3.4 Structural changes in domain I

The local conformational changes in domain I were monitored by acrylodan (covalently attached to Cys 34) due to its high sensitivity towards the micro-environment.[69] An increase in acrylodan fluorescence with a concomitant blue shift in emission maxima of ~22 nm (from 499 nm to 477 nm; Figure 6.5A and B) suggested that the region near Cys 34 gets buried into the hydrophobic environment during aggregation. The emergence of more hydrophobic regions in HSA aggregates could be due to non-specific association between thermally-unfolded, disordered structures. The sequestration of Cys 34 into hydrophobic regions is also accompanied by an enhancement in the aggregate size since an increase in the acrylodan fluorescence anisotropy was observed during HSA fibrillation (Figure 6.5C). We also monitored the steady-state fluorescence anisotropy kinetics of AEDANS (covalently attached to Cys 34 in domain I) since it has a longer fluorescence lifetime (≥ 10 ns) and therefore, is more sensitive to the overall size-growth of aggregates.[70] As aggregation progressed, the AEDANS fluorescence anisotropy showed a time-dependent increase indicating the formation of large-sized aggregates and hence, corroborated the acrylodan anisotropy data (Figure 6.6A).

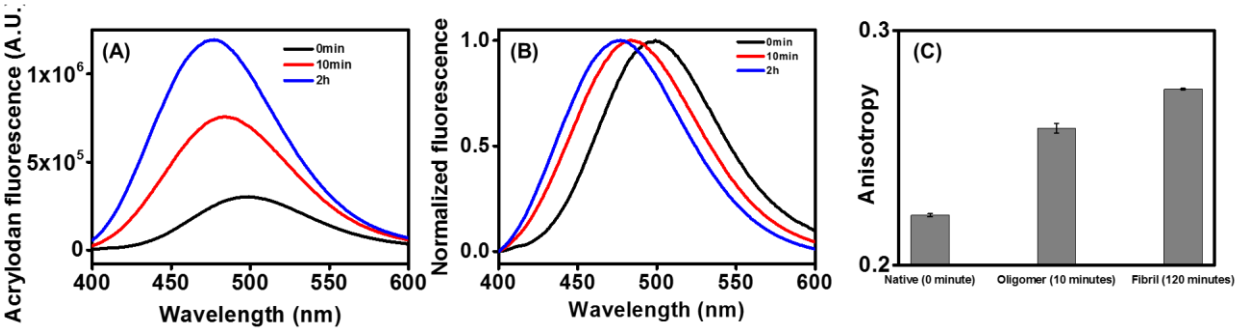


Figure 6.5 (A) Acrylodan (covalently attached to Cys 34) fluorescence at three different time points of aggregation. (B) Normalized fluorescence of acrylodan shows blue shift of 22 nm in the emission maximum from zero time point (0 minute) to final time point (120 minute). (C) The steady-state fluorescence anisotropy of acrylodan at three different time points.

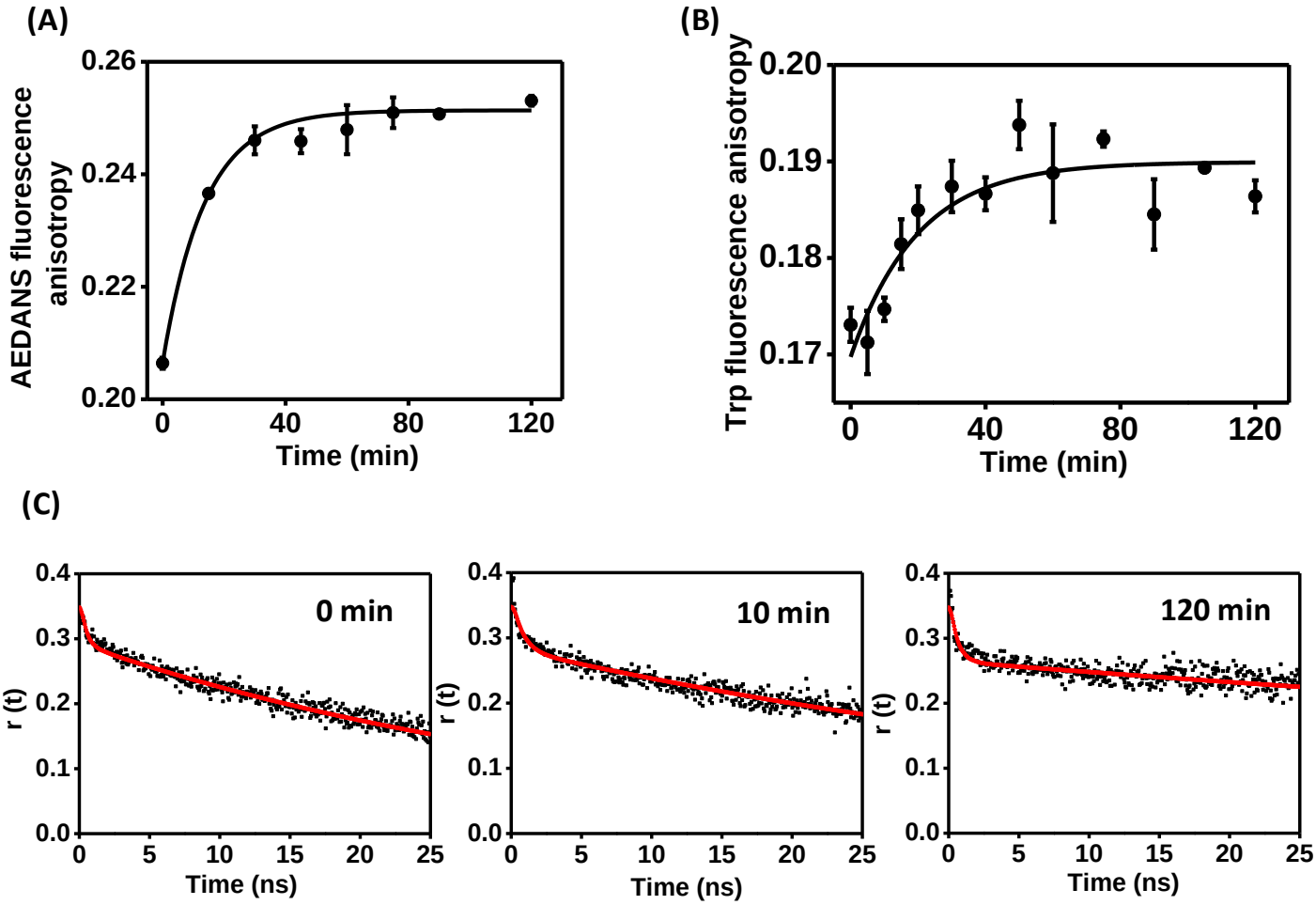


Figure 6.6 Kinetics of amyloid formation monitored by (A) AEDANS covalently attached to the free cysteine and (B) Trp steady-state fluorescence anisotropy. The continuous black lines through the data points represent the fits obtained using a single-exponential kinetic model. (C) Time-resolved fluorescence anisotropy decay for AEDANS at three different stages of aggregation. The red lines represent the bi-exponential anisotropy decay fits using eqn (6) (see Table 6.1 for the recovered parameters associated with depolarization kinetics).

6.3.5 Structural changes in domain II

Investigations into the structural changes in domain II utilizing Trp 214 fluorescence revealed a drop in the fluorescence intensity along with a blue shift of ~10 nm (from 340 nm to 330 nm) in the emission maxima (Figure 6.7A). The blue shift indicates that during aggregation Trp 214 gets relocated into a more hydrophobic environment and the nearby residues such as histidines and cystines probably quench its fluorescence resulting in a drop in the intensity. Additionally, changes in Trp fluorescence indicate changes in the tertiary structure during aggregation. Next, we monitored the steady-state fluorescence anisotropy of Trp 214 which increased as a function

of time, suggestive of the formation of larger aggregates and hence, reaffirmed the acrylodan and AEDANS anisotropy results (Figure 6.6B). Next, in order to investigate the changes in hydrophobicity of HSA during amyloid aggregation, alterations in ANS (non-covalently bound, extrinsic probe) fluorescence intensity were monitored as a function of time since ANS is a well-known hydrophobic environment reporter.[71,72]

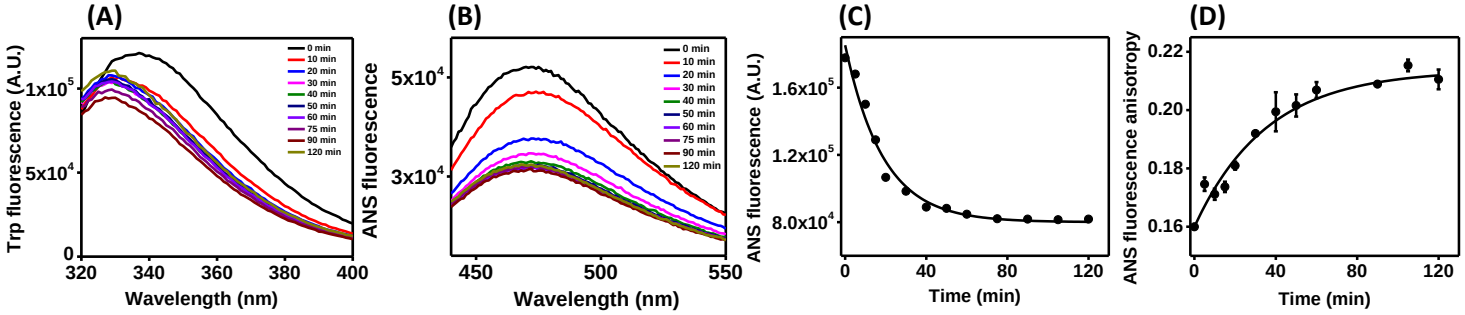


Figure 6.7 (A) Trp and (B) ANS fluorescence spectra at different time points of aggregation. The kinetics of amyloid formation monitored by (C) ANS fluorescence intensity, (D) ANS steady-state fluorescence anisotropy. The continuous black lines through the data points represent the fit obtained using a single-exponential equation.

6.3.6 Changes in overall hydrophobicity

Free ANS is scarcely fluorescent in water with an emission maximum at ~515 nm that undergoes a significant blue shift to ~475 nm with a simultaneous rise in intensity upon binding to hydrophobic pockets.[73] Additionally, it has been demonstrated that the cytotoxicity of amyloidogenic entities could be correlated with the changes in ANS fluorescence intensity.[74] Monomeric native HSA contains at least three hydrophobic, ANS binding sites (in domains II and III) that primarily differ by higher and lower binding affinities.[75,76] In our study, initially we observed a high ANS fluorescence intensity (0 min sample; prior to heating) indicating that ANS was bound to the hydrophobic pockets of the predominantly monomeric protein at room temperature. As aggregation was triggered upon raising the temperature, we observed a progressive drop in ANS fluorescence at ~475 nm which saturated to ~50% of the initial intensity within one hour of amyloid fibrillation (Figure 6.7B and C). Since ANS fluorescence provides information about the overall hydrophobicity, we speculate that at elevated temperature, the hydrophobic sites specifically harboring the ANS get deformed due to the thermal unfolding of domains II and III[77] and consequently, ANS molecules can no longer bind to the hydrophobic pockets. However, the fact that the ANS emission does not undergo any red-shift from ~475 nm (Figure 6.7B) suggests that probably, ANS molecules still remain

bound to the hydrophobic clusters of HSA aggregates. Additionally, a rise in the ANS fluorescence anisotropy (Figure 6.7D) during amyloid aggregation indicated the formation of larger aggregates which is in accordance with our tryptophan anisotropy data. Taken together, our results suggest that both domains I and II undergo conformational changes during HSA fibrillation.

6.3.7 Dynamical signature and average sizes of the aggregates

In order to gather insights into the dynamical changes and the size-growth associated with HSA amyloid formation, time-resolved fluorescence measurements were performed on Cys 34 (domain I) covalently linked with AEDANS, which is considered a suitable fluorophore for studying the depolarization measurements due to the long fluorescence lifetime.[70] The fluorescence anisotropy decay of AEDANS in the native form (0 min), ring-like intermediates (10 min) and worm-like fibrils (120 min) showed bi-exponential depolarization kinetics comprising a fast and a slow rotational correlation times that represent the local motion of the probe and the global motion (tumbling) of the protein molecules/aggregates, respectively (Figure 6.6C and eqn (6) in Experimental methods). The results from the detailed analyses of the time-resolved fluorescence data are summarized in Table 6.1. The longer component (global) was used for estimating the average hydrodynamic sizes of the predominantly monomeric, oligomeric and fibrillary HSA (eqn (7) in Experimental Section) that were found to be 3.4 ± 0.1 nm, 3.9 ± 0.1 nm and 5.2 ± 0.7 nm, respectively. The observed increase in the longer rotational correlation time for oligomers and fibrils is suggestive of the dampening of the global motion resulting from the association of protein molecules to form oligomeric species that subsequently form large-sized amyloid fibrils. Therefore, taken together, all of the fluorescence measurements revealed that both domains I and II get progressively buried into hydrophobic regions which evolve due to non-specific association between temperature-induced disordered segments that are generated during aggregation. The emergence of hydrophobic clusters is also supported by ANS fluorescence at ~475 nm although the hydrophobic pockets in native HSA probably get deformed upon increase in temperature. The progressive increase in the average aggregate sizes, as indicated by steady-state anisotropy, was further confirmed by the time-resolved fluorescence data. Next, we embarked upon delineating finer structural changes at the molecular level during various stages of the assembly process using Raman spectroscopy.

Table 6.1

Typical parameters (rotational correlation times, their respective amplitudes, mean fluorescence lifetimes) associated with the time-resolved fluorescence anisotropy decay of AEDANS in native (0 minute), oligomeric (10 minute) and fibrillar (120 minute) states of human serum albumin (HSA). R_h denotes the average hydrodynamic radii of HSA aggregates, extracted using Equation 7 (see Experimental Methods), at various time points.

Time point (minutes)	ϕ_{fast} (ns) (β_{fast})	ϕ_{slow} (ns) (β_{slow})	τ_{mean}(ns)	R_h (nm)
0	0.7 (0.2)	38.5 (0.8)	15.3	3.5
10	0.7 (0.2)	58.2 (0.8)	14	4.0
120	0.6 (0.3)	> 150 (0.7)	12.2	> 6.0

6.3.8 Structural insights into amyloid assembly using Raman spectroscopy

Raman spectroscopy being a very powerful technique allows us to pinpoint the fine differences among various protein conformations on the basis of mainly backbone amides (amide I: 1620-1700 cm^{-1} and amide III: 1220-1300 cm^{-1}) and several characteristic side-chain vibrations at a residue-specific level such as Cys, Trp, and Tyr.[63,78–80] In our study, we monitored the Raman spectra of HSA aggregates in a time dependent manner, especially the amide I and amide III bands (Figure 6.8) since the deconvolution of the peaks in these regions gave an estimate of secondary structural elements such as α-helices, β-sheets and random coil/extended conformations at different aggregation stages (Tables 6.2 and 6.3). Additionally, detailed analyses of amide III region provided information about the Ramachandran ψ dihedral angle which allowed us to predict the β-sheet packing in amyloid nanostructures.[81] Moreover, the tyrosines and disulfides provided some key insights into the protein conformational transitions during amyloid aggregation.

6.3.9 Insights into the backbone amide I and III regions

The Raman spectrum obtained for HSA prior to incubation at 65 ^{0}C (0 min) showed a major amide I peak at 1658 cm^{-1}, indicative of an α-helical structure along with two minor peaks at 1679 cm^{-1} and 1690 cm^{-1} which correspond to coils/extended conformation (Figure 6.8B). As aggregation progressed upon raising the temperature, the amide I band showed a transition from the initially α-helix rich state (1658 cm^{-1}) to the β-sheet rich state (1670 cm^{-1}) via

oligomeric intermediates rich in extended conformation/random coil content (1660 cm^{-1}) (Figure 6.8B). For instance, the 10 min sample comprising oligomeric species and the ring-like structures retained a high content of random coil and extended conformation along with α-helices and β-sheets (Table 6.2). For subsequent time points, the contents of α-helices, random coils and extended conformations dropped significantly with an appreciable increase in the β-

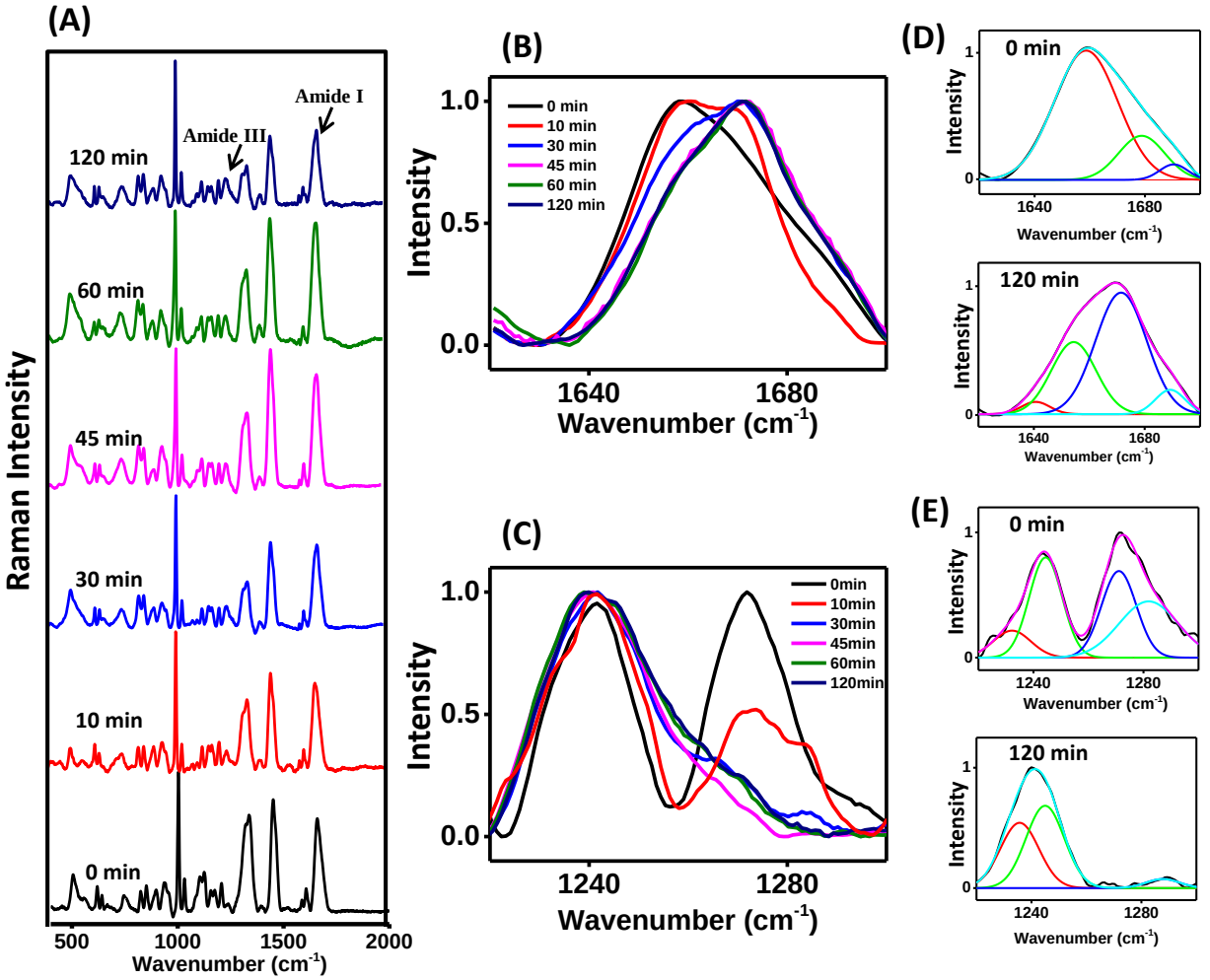


Figure 6.8 (A) Raman spectra (400-2000 cm^{-1}) at different time points of HSA aggregation with black arrows showing amide I and amide III regions. (B) The amide I (1620-1700 cm^{-1}) and (C) amide III region (1220-1300 cm^{-1}) at different time points of aggregation showing transition from a predominantly α-helical to β-sheet-rich structure. The deconvolution of (D) amide I region for zero (0 min) and final time point (120 min) along with (E) amide III region for the same time points is shown. (The individual peaks obtained upon fitting are shown in blue, red, green and magenta. The cumulative peaks are shown in cyan.)

sheet content as is evident from the shift in amide I peak at 1658 cm^{-1} to 1670 cm^{-1} (Figure 6.8B) which is a characteristic of the amyloid architecture. The kinetics of the growth of 1670 cm^{-1} band at the expense of 1658 cm^{-1} band was similar to that of CD and ThT fluorescence (Figure 6.8 and 6.9).

Table 6.2

The estimated content of secondary structural elements (%) from Amide I (1620-1700 cm^{-1}) analysis using Raman spectroscopy:

Time point (minutes)	**α-helix & coils (1645-1665 cm^{-1})**	**β-sheet (1665-1674 cm^{-1})**	**Extended conformation (1680-1690 cm^{-1})**
0	77	18	3
10	74	10	11
30	48	49	10
45	37	57	5
60	39	55	5
120	37	60	8

The amide III peaks of the native HSA (0 min sample) were centered at 1245 cm^{-1} and 1270 cm^{-1} (Figure 6.8C), which correspond to coils/turns and a-helices, respectively. The amide III peak at 1270 cm^{-1} almost completely disappeared with time suggesting an attenuation in the α-helical content (Figure 6.8C) whereas the peak at 1243 cm^{-1} gained prominence. This observation suggested that the substantial drop in α-helical content was concomitant with an increase in the β-sheet content, thus corroborating the results obtained from the amide I region. The peak observed at 1243 cm^{-1} is in agreement with the results obtained for insulin filaments and ovalbumin amyloid pores.[63,82] Further analysis of the amide III peak (at 1243 cm^{-1}), using the approach reported previously,[81] enabled us to estimate the Ramachandran ψ dihedral angle which was found to be $\approx+143^0$; suggestive of antiparallel β-sheet conformation. Interestingly, the analysis of amide III band for early ring-like nanoscopic species formed after 10 min also showed the presence of a higher amount of random coil and extended conformation (Table 6.3) compared to the later times that showed more ordered β-sheet rich structures. Therefore, based on the results obtained from amide I and amide III regions, we suggest that HSA fibrillation proceeds through a loss of α-helical content in conjunction with an increase in the random coil content followed by reorganization into antiparallel β-sheet rich amyloid fibrils.

Table 6.3

The estimated content of secondary structural elements (%) from Amide III (1220-1300 cm^{-1}) analysis using Raman spectroscopy:

Time point (minutes)	**α-helix (1264-1300 cm^{-1})**	**β-sheet, coils & turns (1230-1255 cm^{-1})**
0	59	41
10	43	57
30	35	65
45	18	82
60	19	81
120	5	95

6.3.10 Insights into residue-specific changes during aggregation

After gaining information about protein backbone conformational changes, we investigated the side chain vibrations at the residue-specific level. The peaks observed at 830 and 850 cm^{-1} represent the Fermi doublet for tyrosine. The intensity ratio of 850 to 830 is an indicator of the hydrogen bonding strength between the hydroxyl group of tyrosine and the surrounding solvent molecules.[78,79] The I_{850}/I_{830} ratio for the native conformation was ~1.3 which decreased to ~1.0 during aggregation. This indicates weakening of the H-bonding between the tyrosine moiety and the neighboring water molecules in β-sheet rich conformations which in turn implies that the tyrosines get progressively buried during the aggregation. However, the weakening of the hydrogen bonding is an average assessment since HSA contains 18 tyrosines. We next monitored the changes in the disulfide Raman intensities since HSA contains 17 disulfide bridges that contribute towards the overall structural rigidity but allow sufficient conformational flexibility to HSA upon changes in solution conditions.[83] Typically, the disulfide S–S stretching frequency appears at ~510 ± 5 cm^{-1} as observed during insulin[82,84] and lysozyme aggregation.[85] In our study, we observed multiple peaks in 500-600 cm^{-1} for the disulfide bonds wherein a major peak at ~507 cm^{-1} remained conserved throughout the aggregation. Some new peaks (~520 cm^{-1}, ~550 cm^{-1}) emerged during aggregation and the peak at 507 cm^{-1} became broader (Figure 6.9C). Analysis of all the Raman disulfide peaks revealed a strong dependence of S–S wavenumber position on the torsional angle χ(C–S–S–C) which provides information about the internal rotation about C–S and C–C bonds in C_α–C_β–S–S–$C_{\beta'}$–$C_{\alpha'}$ conformations.[86,87] The peak at 510 ± 5 cm^{-1} (~507 cm^{-1}, in our case) is suggestive of the

gauche–gauche–gauche conformation (χ: ±60^0) that was confirmed by careful analysis of the disulfides in the crystal structure using PyMol (Figure 6.1A). The additional peaks observed at 525 ± 5 cm^{-1} and at 540 ± 5 cm^{-1} are attributed to *gauche–gauche–trans* (χ: ±80-90^0) and *trans–gauche–trans* conformations, respectively. Our results indicate the emergence of multiple, heterogeneous disulfide conformers during the transformation from the α-helical state to cross-β sheet-rich amyloid fibrils comprising antiparallel β-sheets. Additionally, the broadening of peak position during the transition from the native to the amyloid state is suggestive of structural rearrangement as a consequence of intermolecular non-covalent interactions and ‘‘perturbations in local environment’’[85] that emerge during HSA amyloid assembly.

6.3.11 Mechanism of HSA amyloid fibrillation using multiple structural probes

In this study, multiple structural probes were employed to discern the time-dependent evolution of molecular events underlying HSA amyloid formation. We followed the aggregation kinetics by CD, steady-state as well as time-resolved fluorescence, and Raman spectroscopy that revealed a lag phase independent assembly (Figure 6.9). The time-dependence of various spectroscopic readouts could be satisfactorily described using a single exponential kinetics and an apparent first order rate constant of (5 ± 0.8) × 10^{-2} min^{-1} was obtained for all the probes (Figure 6.7, 6.9A and 6.9B,), except for ANS anisotropy. The similarity in the rates suggested that the structural reorganization and amyloid formation occur simultaneously.

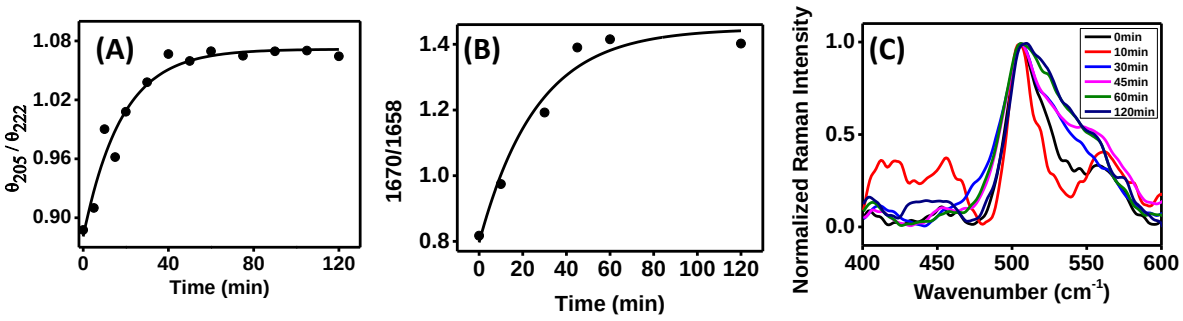


Figure 6.9 The kinetics of amyloid formation monitored by (A) circular dichroism (CD) and (B) Raman spectroscopy (ratio of Raman intensity at 1670 cm^{-1} to 1658 cm^{-1}). (C) Raman spectra of disulfides (400-600 cm^{-1}) at different time points of HSA aggregation.

However, the aggregate growth rate, monitored by ANS fluorescence anisotropy, was found to be somewhat slower compared to the structural changes (Figure 6.10 and Figure 6.7B). These findings suggest that the amyloid formation involves a major structural change that triggers the self-assembly process followed by a slow growth phase.

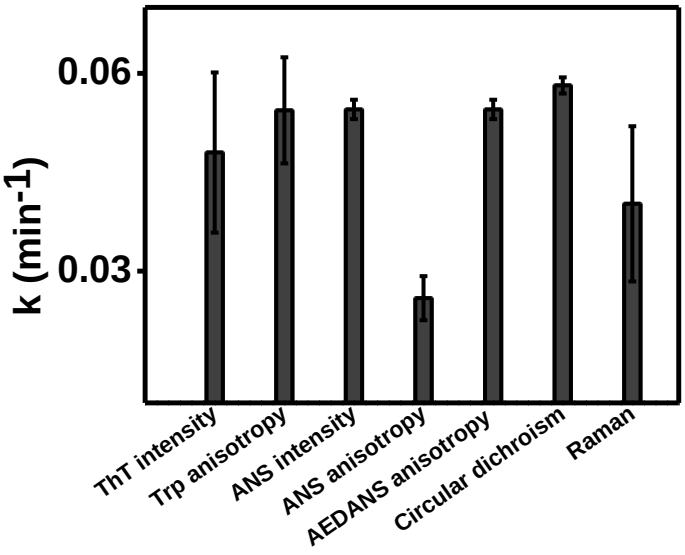


Figure 6.10 The plot comparing the apparent first order rate constants recovered from the amyloid formation kinetics monitored using multiple structural probes viz. ThT and ANS fluorescence intensity; Trp, ANS and AEDANS fluorescence anisotropy; circular dichroism (ratio of ellipticity at 205 nm to 222 nm); amide I vibration from Raman spectroscopy (ratio of 1670 cm^{-1} to 1658 cm^{-1} bands). See also Figure 6.7 and for the kinetics data.

6.4 Summary

Here, we have investigated the mechanism of amyloid fibril formation from human serum albumin using a host of biophysical tools. The kinetic analyses of multiple spectroscopic readouts indicated that the amyloid formation involves a major conformational change from a predominantly monomeric all α-helical to a cross-β sheet rich structure in a lag phase independent manner. Raising the temperature triggers the formation and accumulation of flexible, non-native unfolded or disordered conformers that facilitate the protein self-assembly mediated by various non-covalent including hydrophobic interactions. Also, almost all the domains participate in antiparallel β sheet-rich amyloid assembly whereby tryptophan, tyrosines, and the lone cysteine get sequestered into the hydrophobic interior. Additionally, AFM imaging revealed the formation of a multitude of morphologically-diverse nanoscopic intermediates during amyloid aggregation. We speculate that the nanoscopic ring-like intermediates might be metastable structures that subsequently mature into worm-like fibrils

through a unique mechanism that involves both annealing and unraveling of rings. The antiparallel β-sheet structures have been linked with amyloid-induced toxicity and have been observed for various toxic, annular pore-like oligomers of disease-related proteins or peptides. The ring-like intermediates observed by us might have resemblance with the annular pores that can interact with the membrane and disrupt the membrane by pore formation. The elucidation of the cascade of molecular events and morphological transition is important for designing small molecule inhibitors for anti-amyloid therapeutics. We believe that such mechanistic studies will open new avenues for the controlled fabrication of exotic bionanomaterials having a wide array of nano-biotechnological applications.

6.5 References:

1. Lednev, I. K. Amyloid fibrils: the eighth wonder of the world in protein folding and aggregation. *Biophys. J.* **2014,** *106*, 1433-1435.

2. Pappu, R. V. Frozen in beta. *Biophys. J.* **2014,** *107*, 795-797.

3. Tycko, R.; Wickner, R. B. Molecular Structures of Amyloid and Prion Fibrils: Consensus versus Controversy. *Acc. Chem. Res.* **2013,** *46*, 1487-1496.

4. Sarell, C. J.; Stockley, P. G.; Radford, S. E. Assessing the causes and consequences of co-polymerization in amyloid formation. *Prion* **2013,** *7*, 359-368.

5. Greenwald, J.; Riek, R. Biology of amyloid: structure, function, and regulation. *Structure* **2010,** *18*, 1244-1260.

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

7. Nelson, R.; Sawaya, M. R.; Balbirnie, M.; Madsen, A. O.; Riekel, C.; Grothe, R.; Eisenberg, D. Structure of the cross-beta spine of amyloid-like fibrils. *Nature* **2005,** *435*, 773-778.

8. Brettschneider, J.; Del Tredici, K.; Lee, V. M.; Trojanowski, J. Q. Spreading of pathology in neurodegenerative diseases: a focus on human studies. *Nat. Rev. Neurosci.* **2015,** *16*, 109-120.

9. Uversky, V. N.; Davé, V.; Iakoucheva, L. M.; Malaney, P.; Metallo, S. J.; Pathak, R. R.; Joerger, A. C. Pathological Unfoldomics of Uncontrolled Chaos: Intrinsically Disordered Proteins and Human Diseases. *Chem. Rev.* **2014,** *114*, 6844-6879.

10. Knowles, T. P.; Vendruscolo, M.; Dobson, C. M. The amyloid state and its association with protein misfolding diseases. *Nat Rev Mol Cell Biol* **2014,** *15* (6), 384-396.

11. Meier, B. H.; Bockmann, A. The structure of fibrils from 'misfolded' proteins. *Curr. Opin. Struct. Biol.* **2014,** *30C*, 43-49.

12. Eisenberg, D.; Jucker, M. The amyloid state of proteins in human diseases. *Cell* **2012,** *148*, 1188-1203.

13. Moreno-Gonzalez, I.; Soto, C. Misfolded Protein Aggregates: Mechanisms, Structures and Potential for Disease Transmission. *Semin. Cell Dev. Biol.* **2011,** *22*, 482-487.

14. Lee, J.; Culyba, E. K.; Powers, E. T.; Kelly, J. W. Amyloid-β forms fibrils by nucleated conformational conversion of oligomers. *Nat. Chem. Biol.* **2011,** *7*, 602-609.

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

16. Chiti, F.; Webster, P.; Taddei, N.; Clark, A.; Stefani, M.; Ramponi, G.; Dobson, C. M. Designing conditions for in vitro formation of amyloid protofilaments and fibrils. *Proc. Natl. Acad. Sci. U. S. A.* **1999,** *96*, 3590-3594.

17. Fowler, D. M.; Koulov, A. V.; Balch, W. E.; Kelly, J. W. Functional amyloid – from bacteria to humans. *Trends Biochem. Sci.* **2007,** *32*, 217-224.

18. Barnhart, M. M.; Chapman, M. R. Curli Biogenesis and Function. *Annu. Rev. Microbiol.* **2006,** *60*, 131-147.

19. Wu, H. Higher-Order Assemblies in a New Paradigm of Signal Transduction. *Cell* **2013,** *153* (2), 287-292.

20. Cherny, I.; Gazit, E. Amyloids: Not Only Pathological Agents but Also Ordered Nanomaterials. *Angew. Chem. Int. Ed.* **2008,** *47*, 4062-4069.

21. Scheibel, T.; Parthasarathy, R.; Sawicki, G.; Lin, X.-M.; Jaeger, H.; Lindquist, S. L. Conducting nanowires built by controlled self-assembly of amyloid fibers and selective metal deposition. *Proc. Natl. Acad. Sci. U. S. A.* **2003,** *100*, 4527-4532.

22. Li, C.; Mezzenga, R. The interplay between carbon nanomaterials and amyloid fibrils in bio-nanotechnology. *Nanoscale* **2013,** *5*, 6207-6218.

23. Shimanovich, U.; Efimov, I.; Mason, T. O.; Flagmeier, P.; Buell, A. K.; Gedanken, A.; Linse, S.; Åkerfeldt, K. S.; Dobson, C. M.; Weitz, D. A.; Knowles, T. P. J. Protein Microgels from Amyloid Fibril Networks. *ACS Nano* **2015,** *9* (1), 43-51.

24. Stefani, M. Structural features and cytotoxicity of amyloid oligomers: Implications in Alzheimer's disease and other diseases with amyloid deposits. *Prog. Neurobiol.* **2012,** *99*, 226-245.

25. Bucciantini, M.; Giannoni, E.; Chiti, F.; Baroni, F.; Formigli, L.; Zurdo, J.; Taddei, N.; Ramponi, G.; Dobson, C. M.; Stefani, M. Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature* **2002,** *416,* 507-511.

26. Butterfield, S. M.; Lashuel, H. A. Amyloidogenic Protein–Membrane Interactions: Mechanistic Insight from Model Systems. *Angew. Chem. Int. Ed.* **2010,** *49,* 5628-5654.

27. Dahlgren, K. N.; Manelli, A. M.; Stine, W. B.; Baker, L. K.; Krafft, G. A.; LaDu, M. J. Oligomeric and Fibrillar Species of Amyloid-β Peptides Differentially Affect Neuronal Viability. *J. Biol. Chem.* **2002,** *277,* 32046-32053.

28. Lashuel, H. A.; Lansbury, P. T. Are amyloid diseases caused by protein aggregates that mimic bacterial pore-forming toxins? *Q. Rev. Biophys.* **2006,** *39,* 167-201.

29. Stöckl, M.; Zijlstra, N.; Subramaniam, V. α-Synuclein Oligomers: an Amyloid Pore? *Mol. Neurobiol.* **2013,** *47,* 613-621.

30. Lin, H. A. I.; Bhatia, R.; Lal, R. Amyloid β protein forms ion channels: implications for Alzheimer's disease pathophysiology. *The FASEB Journal* **2001,** *15,* 2433-2444.

31. Nath, A.; Miranker, A. D.; Rhoades, E. A membrane-bound antiparallel dimer of rat islet amyloid polypeptide. *Angew. Chem. Int. Ed. Engl.* **2011,** *50,* 10859-10862.

32. Cao, A.; Hu, D.; Lai, L. Formation of amyloid fibrils from fully reduced hen egg white lysozyme. *Protein Sci.* **2004,** *13* (2), 319-324.

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

34. Jiménez, J. L.; Nettleton, E. J.; Bouchard, M.; Robinson, C. V.; Dobson, C. M.; Saibil, H. R. The protofilament structure of insulin amyloid fibrils. *Proc. Natl. Acad. Sci. U. S. A.* **2002,** *99,* 9196-9201.

35. Taboada, P.; Barbosa, S.; Castro, E.; Mosquera, V. Amyloid Fibril Formation and Other Aggregate Species Formed by Human Serum Albumin Association. *J. Phys. Chem. B* **2006,** *110,* 20733-20736.

36. Bhattacharya, M.; Jain, N.; Mukhopadhyay, S. Insights into the Mechanism of Aggregation and Fibril Formation from Bovine Serum Albumin. *J. Phys. Chem. B* **2011,** *115,* 4195-4205.

37. Mishra, R.; Sörgjerd, K.; Nyström, S.; Nordigården, A.; Yu, Y.-C.; Hammarström, P. Lysozyme Amyloidogenesis Is Accelerated by Specific Nicking and Fragmentation but Decelerated by Intact Protein Binding and Conversion. *J. Mol. Biol.* **2007,** *366*, 1029-1044.

38. Eichner, T.; Kalverda, A. P.; Thompson, G. S.; Homans, S. W.; Radford, S. E. Conformational Conversion during Amyloid Formation at Atomic Resolution. *Mol. Cell* **2011,** *41*, 161-172.

39. Hill, S. E.; Robinson, J.; Matthews, G.; Muschol, M. Amyloid protofibrils of lysozyme nucleate and grow via oligomer fusion. *Biophys. J.* **2009,** *96*, 3781-3790.

40. Schlamadinger, Diana E.; Miranker, Andrew D. Fiber-Dependent and -Independent Toxicity of Islet Amyloid Polypeptide. *Biophys. J.* **2014,** *107*, 2559-2566.

41. Curry, S.; Mandelkow, H.; Brick, P.; Franks, N. Crystal structure of human serum albumin complexed with fatty acid reveals an asymmetric distribution of binding sites. *Nat. Struct. Mol. Biol.* **1998,** *5*, 827-835.

42. Sugio, S.; Kashima, A.; Mochizuki, S.; Noda, M.; Kobayashi, K. Crystal structure of human serum albumin at 2.5 Å resolution. *Protein Eng.* **1999,** *12*, 439-446.

43. He, X. M.; Carter, D. C. Atomic structure and chemistry of human serum albumin. *Nature* **1992,** *358*, 209-215.

44. Dockal, M. Conformational Transitions of the Three Recombinant Domains of Human Serum Albumin Depending on pH. *J. Biol. Chem.* **2000,** *275*, 3042-3050.

45. Das, N. K.; Ghosh, N.; Kale, A. P.; Mondal, R.; Anand, U.; Ghosh, S.; Tiwari, V. K.; Kapur, M.; Mukherjee, S. Temperature induced morphological transitions from native to unfolded aggregated States of human serum albumin. *J. Phys. Chem. B* **2014,** *118*, 7267-7276.

46. Shaw, A. K.; Pal, S. K. Spectroscopic studies on the effect of temperature on pH-induced folded states of human serum albumin. *J. Photochem. Photobiol. B* **2008,** *90*, 69-77.

47. Qiu, W.; Zhang, L.; Okobiah, O.; Yang, Y.; Wang, L.; Zhong, D.; Zewail, A. H. Ultrafast Solvation Dynamics of Human Serum Albumin: Correlations with

Conformational Transitions and Site-Selected Recognition. *J. Phys. Chem. B* **2006,** *110*, 10540-10549.

48. Richter, K.; Haslbeck, M.; Buchner, J. The heat shock response: life on the verge of death. *Mol. Cell* **2010,** *40*, 253-266.

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

50. Pappu, R. V.; Wang, X.; Vitalis, A.; Crick, S. L. A polymer physics perspective on driving forces and mechanisms for protein aggregation. *Arch. Biochem. Biophys.* **2008,** *469*, 132-141.

51. Qiao, Q.; Bowman, G. R.; Huang, X. Dynamics of an Intrinsically Disordered Protein Reveal Metastable Conformations That Potentially Seed Aggregation. *J. Am. Chem. Soc.* **2013,** *135*, 16092-16101.

52. Juárez, J.; Taboada, P.; Mosquera, V. Existence of Different Structural Intermediates on the Fibrillation Pathway of Human Serum Albumin. *Biophys. J.* **2009,** *96*, 2353-2370.

53. Juárez, J.; Taboada, P.; Goy-López, S.; Cambón, A.; Madec, M.-B.; Yeates, S. G.; Mosquera, V. Additional Supra-Self-Assembly of Human Serum Albumin under Amyloid-Like-Forming Solution Conditions. *J. Phys. Chem. B* **2009,** *113*, 12391-12399.

54. Stirpe, A.; Pantusa, M.; Rizzuti, B.; Sportelli, L.; Bartucci, R.; Guzzi, R. Early stage aggregation of human serum albumin in the presence of metal ions. *Int. J. Biol. Macromol.* **2011,** *49*, 337-42.

55. Juárez, J.; Alatorre-Meda, M.; Cambón, A.; Topete, A.; Barbosa, S.; Taboada, P.; Mosquera, V. Hydration effects on the fibrillation process of a globular protein: the case of human serum albumin. *Soft Matter* **2012,** *8*, 3608.

56. Movaghati, S.; Moosavi-Movahedi, A. A.; Khodagholi, F.; Digaleh, H.; Kachooei, E.; Sheibani, N. Sodium dodecyl sulphate modulates the fibrillation of human serum albumin in a dose-dependent manner and impacts the PC12 cells retraction. *Colloids Surf. B. Biointerfaces* **2014,** *122*, 341-349.

57. Wang, R.; Sun, S.; Bekos, E. J.; Bright, F. V. Dynamics Surrounding Cys-34 in Native, Chemically Denatured, and Silica-Adsorbed Bovine Serum Albumin. *Anal. Chem.* **1995,** *67*, 149-159.

58. Flora, K.; Brennan, J. D.; Baker, G. A.; Doody, M. A.; Bright, F. V. Unfolding of acrylodan-labeled human serum albumin probed by steady-state and time-resolved fluorescence methods. *Biophys. J.* **1998,** *75*, 1084-1096.

59. Hudson, E. N.; Weber, G. Synthesis and characterization of two fluorescent sulfhydryl reagents. *Biochemistry* **1973,** *12*, 4154-4161.

60. Lakowicz, J. R. *Principles of fluorescence spectroscopy*. Springer: New York, 2006.

61. Jain, N.; Mukhopadhyay, S. Applications of Fluorescence Anisotropy in Understanding Protein Conformational Disorder and Aggregation. In *Applied Spectroscopy and the Science of Nanomaterials*, Misra, P., Ed. Springer Singapore: 2015; Vol. 2, pp 41-57.

62. Arya, S.; Mukhopadhyay, S. Ordered Water within the Collapsed Globules of an Amyloidogenic Intrinsically Disordered Protein. *J. Phys. Chem. B* **2014,** *118*, 9191-9198.

63. Bhattacharya, M.; Jain, N.; Dogra, P.; Samai, S.; Mukhopadhyay, S. Nanoscopic Amyloid Pores Formed via Stepwise Protein Assembly. *J. Phys. Chem. Lett.* **2013,** *4*, 480-485.

64. Levine, H. Thioflavine T interaction with synthetic Alzheimer's disease β-amyloid peptides: Detection of amyloid aggregation in solution. *Protein Sci.* **1993,** *2*, 404-410.

65. Holm, N. K.; Jespersen, S. K.; Thomassen, L. V.; Wolff, T. Y.; Sehgal, P.; Thomsen, L. A.; Christiansen, G.; Andersen, C. B.; Knudsen, A. D.; Otzen, D. E. Aggregation and fibrillation of bovine serum albumin. *Biochim. Biophys. Acta* **2007,** *1774*, 1128-1138.

66. Usov, I.; Adamcik, J.; Mezzenga, R. Polymorphism in bovine serum albumin fibrils: morphology and statistical analysis. *Faraday Discuss.* **2013,** *166*, 151.

67. Volpatti, L. R.; Vendruscolo, M.; Dobson, C. M.; Knowles, T. P. J. A Clear View of Polymorphism, Twist, and Chirality in Amyloid Fibril Formation. *ACS Nano* **2013,** *7*, 10443-10448.

68. Sweers, K. K. M.; van der Werf, K. O.; Bennink, M. L.; Subramaniam, V. Atomic Force Microscopy under Controlled Conditions Reveals Structure of C-Terminal Region of α-Synuclein in Amyloid Fibrils. *ACS Nano* **2012,** *6*, 5952-5960.

69. Hibbs, R. E.; Talley, T. T.; Taylor, P. Acrylodan-conjugated Cysteine Side Chains Reveal Conformational State and Ligand Site Locations of the Acetylcholine-binding Protein. *J. Biol. Chem.* **2004,** *279*, 28483-28491.

70. Mukhopadhyay, S.; Nayak, P. K.; Udgaonkar, J. B.; Krishnamoorthy, G. Characterization of the Formation of Amyloid Protofibrils from Barstar by Mapping Residue-specific Fluorescence Dynamics. *J. Mol. Biol.* **2006,** *358*, 935-942.

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

72. Marshall, K. E.; Morris, K. L.; Charlton, D.; O'Reilly, N.; Lewis, L.; Walden, H.; Serpell, L. C. Hydrophobic, aromatic, and electrostatic interactions play a central role in amyloid fibril formation and stability. *Biochemistry* **2011,** *50*, 2061-2071.

73. Daniel, E.; Weber, G. Cooperative Effects in Binding by Bovine Serum Albumin. I. The Binding of 1-Anilino-8-naphthalenesulfonate. Fluorimetric Titrations. *Biochemistry* **1966,** *5* (6), 1893-1900.

74. Bolognesi, B.; Kumita, J. R.; Barros, T. P.; Esbjorner, E. K.; Luheshi, L. M.; Crowther, D. C.; Wilson, M. R.; Dobson, C. M.; Favrin, G.; Yerbury, J. J. ANS Binding Reveals Common Features of Cytotoxic Amyloid Species. *ACS Chem. Biol.* **2010,** *5*, 735-740.

75. Takehara, K.; ocirc; Yuki, K.; Shirasawa, M.; Yamasaki, S.; Yamada, S. Binding Properties of Hydrophobic Molecules to Human Serum Albumin Studied by Fluorescence Titration. *Anal. Sci.* **2009,** *25*, 115-120.

76. Kuznetsova, I. M.; Sulatskaya, A. I.; Povarova, O. I.; Turoverov, K. K. Reevaluation of ANS binding to human and bovine serum albumins: key role of equilibrium microdialysis in ligand - receptor binding characterization. *PLoS One* **2012,** *7*, e40845.

77. Farruggia, B.; Rodriguez, F.; Rigatuso, R.; Fidelio, G.; Picó, G. The Participation of Human Serum Albumin Domains in Chemical and Thermal Unfolding. *J. Protein Chem.* **2001,** *20*, 81-89.

78. Nemecek, D.; Stepanek, J.; Thomas, G. J. Raman Spectroscopy of Proteins and Nucleoproteins. In *Current Protocols in Protein Science*, John Wiley & Sons, Inc.: 2001.

79. Tuma, R. Raman spectroscopy of proteins: from peptides to large assemblies. *J. Raman Spectrosc.* **2005,** *36*, 307-319.

80. Shashilov, V. A.; Sikirzhytski, V.; Popova, L. A.; Lednev, I. K. Quantitative methods for structural characterization of proteins based on deep UV resonance Raman spectroscopy. *Methods* **2010,** *52*, 23-37.

81. Mikhonin, A. V.; Bykov, S. V.; Myshakina, N. S.; Asher, S. A. Peptide Secondary Structure Folding Reaction Coordinate: Correlation between UV Raman Amide III Frequency, Ψ Ramachandran Angle, and Hydrogen Bonding. *J. Phys. Chem. B* **2006,** *110*, 1928-1943.

82. Zako, T.; Sakono, M.; Hashimoto, N.; Ihara, M.; Maeda, M. Bovine Insulin Filaments Induced by Reducing Disulfide Bonds Show a Different Morphology, Secondary Structure, and Cell Toxicity from Intact Insulin Amyloid Fibrils. *Biophys. J.* **2009,** *96*, 3331-3340.

83. Bhattacharya, M.; Jain, N.; Bhasne, K.; Kumari, V.; Mukhopadhyay, S. pH-induced Conformational Isomerization of Bovine Serum Albumin Studied by Extrinsic and Intrinsic Protein Fluorescence. *J. Fluoresc.* **2011,** *21*, 1083-1090.

84. Huang, K.; Maiti, N. C.; Phillips, N. B.; Carey, P. R.; Weiss, M. A. Structure-Specific Effects of Protein Topology on Cross-β Assembly: Studies of Insulin Fibrillation. *Biochemistry* **2006,** *45*, 10278-10293.

85. Remmele, R., Jr.; McMillan, P.; Bieber, A. Raman spectroscopic studies of hen egg-white lysozyme at high temperatures and pressures. *J. Protein Chem.* **1990,** *9*, 475-486.

86. Biswas, N.; Waring, A. J.; Walther, F. J.; Dluhy, R. A. Structure and conformation of the disulfide bond in dimeric lung surfactant peptides SP-B1-25 and SP-B8-25. *Biochim. Biophys. Acta* **2007,** *1768*, 1070-1082.

87. Van Wart, H. E.; Scheraga, H. A. Raman spectra of strained disulfides. Effect of rotation about sulfur-sulfur bonds on sulfur-sulfur stretching frequencies. *J. Phys. Chem.* **1976,** *80*, 1823-1832.

Conclusions and future directions

7.1 Role of chain-chain and chain-water interactions in amyloid aggregation

The intricate balance between chain-chain and chain-water interaction underlies the major physical driving force for protein aggregation (discussed in detail in Chapter 1). A wealth of information has been obtained on the mechanism of coupled conformational switch and aggregation from the protein perspective. However, the precise mechanism of amyloid aggregation, the behavior of water molecules in and around IDPs, and the role of water in amyloid formation remains elusive. The work described in this thesis addresses the important role of both chain-chain and chain-water interaction in amyloid aggregation. Below, I briefly summarize:

(i) How the water arrangement in the native state of a collapsed IDP (Chapter 2 and 3) differs from the water structure in an expanded IDP (Chapter 4),

(ii) How water rearranges itself upon disorder-to-amyloid transition in a collapsed (Chapter 3) as well as an expanded IDP (Chapter 4), and

(iii) The mechanistic insights obtained from the amyloid formation studies carried out using a model amyloidogenic IDP (Chapter 5) and a partially disordered form of a globular protein (Chapter 6).

The conformational changes (chain-chain interactions) during the amyloid formation were mainly studied using circular dichroism (CD), atomic force microscopy (AFM) and Raman spectroscopy, and the behavior of water molecules in IDPs and amyloids (chain-solvent) was monitored using femtosecond- and picosecond-resolved fluorescence spectroscopy.

7.1.1 Interplay of conformational preference and water mobility in IDPs

IDPs lack a well-defined 3D structure and exist as dynamic ensembles of rapidly fluctuating interconverting conformations. The amino acid sequence composition dictates the conformational ensemble most accessible to a particular IDP, in a given solvent. The work presented in this thesis has focused on two IDPs, namely, κ-casein and α-synuclein. κ-casein exists as an ensemble of collapsed globules and α-synuclein belongs to the expanded class of IDPs. Since the behavior of water in and around IDPs is only beginning to be understood, I wanted to discern the differences in the water arrangement around these two IDPs that fall into two different conformational class of IDPs (These different conformational classes of IDPs have been discussed in Section 1.3.3 of Chapter 1) and establish a relation between the conformational preference of an IDP and the mobility of water molecules around it. As described in Chapter 2 and 3, a small fraction of water molecules present within the collapsed

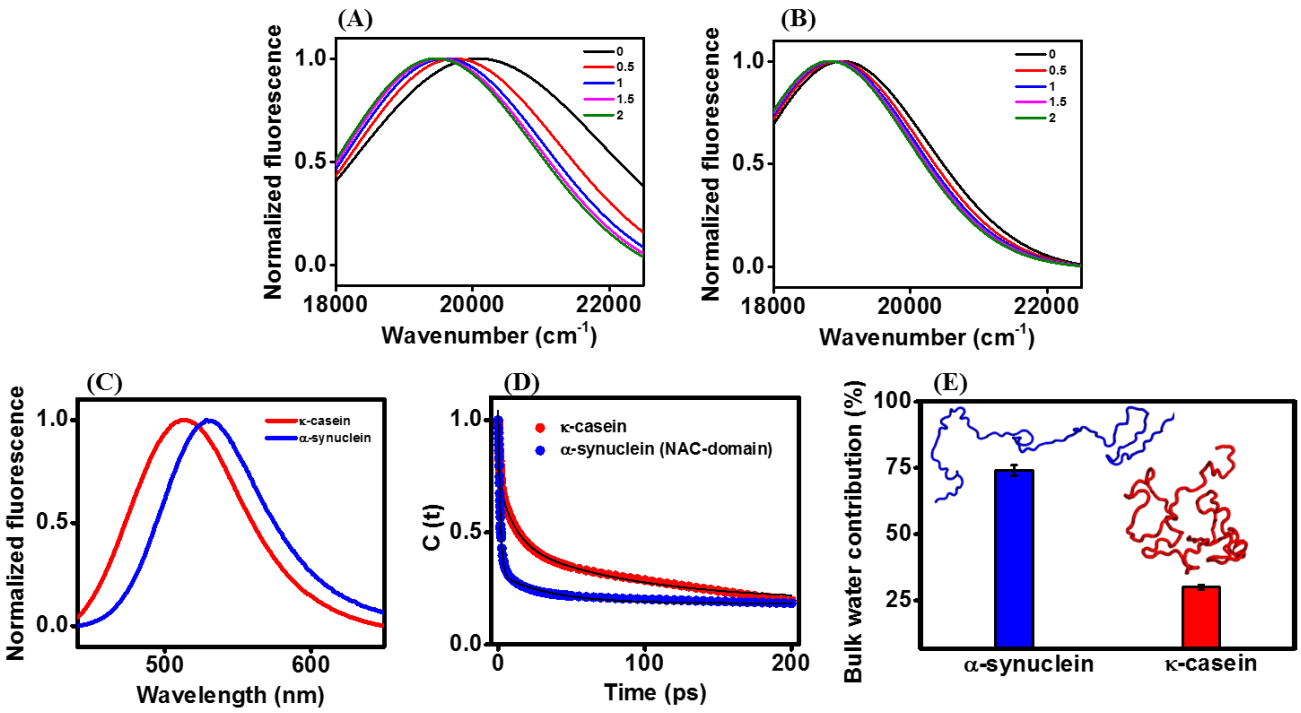


Figure 7.1 A comparison between a collapsed and an expanded IDP, κ-casein and α-synuclein, respectively. TCSPC TRES for probing nanosecond component of solvation in (A) κ-casein and (B) α-synuclein (residue 78 in NAC domain). (C) The steady-state fluorescence spectra for acrylodan in κ-casein (red) and α-synuclein (residue 78 in NAC domain) (blue). (D) The solvation correlation function [C(t)] plots from femtosecond upconversion, for κ-casein (red) and α-synuclein (residue 78 in NAC domain) (blue).

globules of κ-casein showed much retarded (nanosecond) dynamics, which was 3 orders of magnitude slower than the bulk water and atleast an order of magnitude slower than the biological water (Figure 7.1A). The expanded IDP α-synuclein (Chapter 4) did not exhibit the presence of this nanosecond solvation component (Figure 7.1B). Furthermore, when we compared the solvation correlation function [C(t)] plots generated from the femtosecond upconversion data for these two IDPs, it was clear that the surface water mobility gets retarded in case of κ-casein (~80 ps) compared to α-synuclein (~30 ps). Additionally, the bulk water contribution was much lower in case of κ-casein (~25%) compared to α-synuclein (~75%) (Figure 7.1E). Thus, it is clear that there are some significant differences in the water arrangement around these two conformationally distinct IDPs. These differences are likely to have a strong correlation with the sequence composition as these two IDPs have very different amino acid compositions. For example, in case of κ-casein, the prevalence of polar residues like Gln and Asn that can form extended hydrogen bonded network with water molecules, is much more compared to α-synuclein. Taken together, these results suggest that this water

mobility is likely to be strongly coupled to the conformational preference of IDPs. By carrying out similar hydration dynamics studies on other IDPs with different conformational preferences, we can set up a paradigm for characterization of IDPs based on their hydration profiles.

7.1.2 Commonalities and discrepancies in the water arrangement within amyloid fibrils formed by two different IDPs

In Chapter 3 and Chapter 4, I have discussed the intriguing aspects of water rearrangement upon disorder-to-order amyloid transition in κ-casein and α-synuclein, respectively. When the hydration profiles of the amyloid fibrils were compared, it was interesting to note that in both cases, the ns component of solvation (Figure 7.2) of ~2 ns was observed which suggests that the slow solvation component (strongly held water molecules) might be a common feature for most amyloid fibrils. These highly ordered water molecules might have an important role in the maintenance of the robust architecture of amyloid fibrils. Though both fibrils showed the presence of ns component, the contribution of this component was more in case of α-synuclein amyloid fibrils. Moreover, the loss in solvation (with respect to time-zero recovered from our femtosecond experiments) within the time resolution of our TCSPC setup (~250 ps) was more in case of κ-casein amyloid fibrils (Figure 7.2) which reaffirms the dominance of ns component of solvation in case of α-synuclein amyloid fibrils compared to κ-casein. The differences in the water arrangement might account for the differences in the solubility of amyloid fibrils as κ-casein amyloid fibrils are soluble in nature whereas α-synuclein forms insoluble amyloid

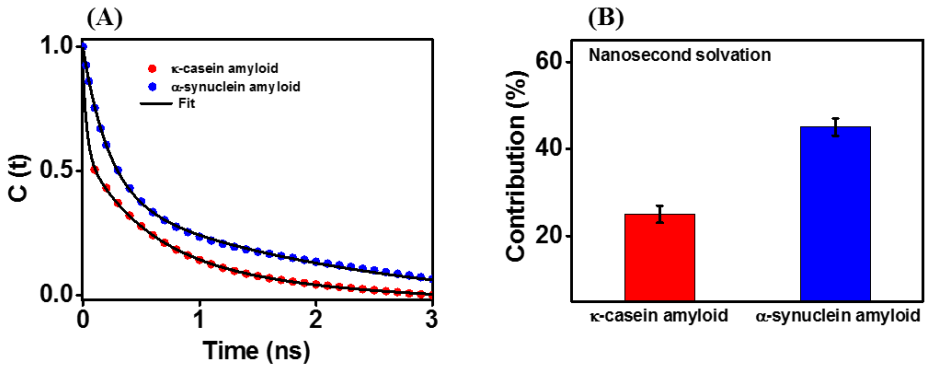


Figure 7.2 A comparison between the amyloid fibrils formed from a collapsed and an expanded IDP, κ-casein and α-synuclein, respectively. (A) TCSPC TRES for probing nanosecond component of solvation in and (B) The percentage contribution for the nanosecond component of solvation for amyloid fibrils formed by κ-casein (red) and α-synuclein (residue 78 in NAC domain) (blue).

fibrils. The higher contribution of water molecules that can potentially exchange with bulk water (relaxing on the timescale of 10-100 ps) in case of κ-casein fibrils (see Chapter 3) might be responsible for keeping them in soluble form and the lower contribution of such water molecules in case of α-synuclein fibrils might explain why they tend to phase out or precipitate. The case of these two fibrils, one of which is soluble and the other is insoluble, provides some cues into the important role of water in protein aggregation. This approach can be further extended to other aggregation-prone IDPs. I speculate that oligomeric species in case of many pathologically relevant IDPs such as α-synuclein and Tau, which are soluble in nature (like κ-casein fibrils studied in this thesis) and often highly cytotoxic, would show significant differences in the water arrangement compared to fibrillar species. Thus, this approach might prove very promising for the designing of amyloid detection tools.

7.1.3 Mechanistic insights into the amyloid formation using a model globular protein and IDP

In Chapter 5 and Chapter 6, I have discussed about the mechanism of amyloid formation from a model amyloidogenic IDP, κ-casein and a model globular protein, human serum albumin (HSA), respectively. κ-casein followed a nucleation dependent polymerization pathway under normal physiological conditions and lead to the formation of amyloid fibrils. However, in the presence of an anionic surfactant, SDS which is a well-known lipid mimetic, a switch in the aggregation mechanism from nucleation-dependent polymerization to lag-phase independent isodesmic polymerization, was observed. Additionally, in the presence of SDS, I observed the formation of mostly large spherical oligomeric species, which did not convert into fibrils. In case of HSA, upon inducing partial denaturation, partially disordered monomers converted into amyloid fibrils via isodesmic polymerization pathway. During the process of amyloid formation, a significant population of annular ring-like oligomeric species was observed. The morphology of the oligomeric species as well as amyloid fibrils formed by HSA was very different from those formed by κ-casein, which suggests a strong link between the mechanism and morphology. Thus, by controlling the mechanism, the nanoscale morphology of the fibrils can be modulated. This particular aspect would be of benefit for the designing of amyloid-based novel nanomaterials.

7.2 Future outlook: p53 and cancer

Human p53 is a key regulatory protein that participates in cellular processes such as apoptosis, DNA repair, and cell cycle control.[1] It functions as a homotetrameric nuclear phosphoprotein, and its function is lost in more than 50% of human cancers.[2] It is comprised of 393 amino acid residues and can be divided into three domains: the N-terminal transactivation domain, the core domain (p53C) and the tetramerization or the C-terminal domain.[1] The core domain (p53C) comprises the residues 94-312, that constitute the DNA binding domain (DBD). Most mutations in p53 that are related to cancer development are found in the DBD.[3] More than 90% of the point mutations that are related to malignancy are found in this segment. Moreover, amyloid aggregates of both mutant and WT (wild-type) forms of p53 have been detected in tumor tissues.[4-5] Therefore, aggregation might be a crucial aspect of cancer development, as p53 would lose its functions in an aggregated state. Mutant p53 can also exert a dominant-negative regulatory effect on WT protein, as mutant p53 can convert WT protein into more aggregation-prone species, leading into gain of function in addition to the loss of tumor suppressor function.[2] Some of the hotspot mutations include R175H, R248Q, R273H, R248W and R273C. I have created R248Q mutation in p53c using site-directed mutagenesis.* Since R248Q is a structural mutant, therefore, studying the aggregation behavior of WT and the mutant protein would be of significance for understanding the role of amyloid formation in cancer. Furthermore, the N-terminal transactivation domain (NTD) of p53 is intrinsically disordered in nature. In order to carry out studies on NTD of p53, I created a construct for expressing NTD by introducing a stop codon in the full length construct of p53, using site-directed mutagenesis.* Therefore, one can study the role of this intrinsically NTD in p53 aggregation. For carrying out these studies, I have created cysteine mutants in NTD at 9, 33 and 90 position.** These mutants can be labeled with an environment sensitive fluorophore to carry out aggregation studies.

*The constructs for full length p53 (24859: human p53-(1-393) and p53c (24866: human p53-(94-312)) were purchased from Addgene.

Primers for **1-93** construct:

Forward: 5' C TCC TGG CCC CTG TGA TCT TCT GTC CCT TC 3'

Reverse: 5' GA AGG GAC AGA AGA TCA CAG GGG CCA GGA G 3'

Primers for R**248**Q p53c construct:

Forward: 5' ATG GGC GGC ATG AAC CAG AGG CCC ATC CTC ACC 3'

Reverse: 5' GGT GAG GAT GGG CCT CTG GTT CAT GCC GCC CAT 3'

** Primers for (a) S**9**C:

Forward: 5' CAG TCA GAT CCT TGC GTC GAG CCC CCT CTG 3'
Reverse: 5' CAG AGG GGG CTC GAC GCA AGG ATC TGA CTG 3'
(b) S**33**C:
Forward: 5' AAC AAC GTT CTG TGC CCC TTG CCG TCC 3'
Reverse: 5' GGA CGG CAA GGG GCA CAG AAC GTT GTT 3'
(c) S**90**C:
Forward: 5' GCC CCC TGC TGG CCC CTG TGA TCT TCT 3'
Reverse: 5' AGA AGA TCA CAG GGG CCA GCA GGG GGC 3'

References:

1. Joerger, A. C.; Fersht, A. R. Structural Biology of the Tumor Suppressor p53. *Annu. Rev. Biochem.* **2008,** *77*, 557-582.

2. Silva, J. L.; Gallo, C. V. D. M.; Costa, D. C. F.; Rangel, L. P. Prion-like aggregation of mutant p53 in cancer. *Trends Biochem. Sci.* **2014,** *39*, 260-267.

3. Muller, P. A. J.; Vousden, K. H. p53 mutations in cancer. *Nat. Cell Biol.* **2013,** *15*, 2-8.

4. Levy, C. B.; Stumbo, A. C.; Ano Bom, A. P. D.; Portari, E. A.; Carneiro, Y.; Silva, J. L.; De Moura-Gallo, C. V. Co-localization of mutant p53 and amyloid-like protein aggregates in breast tumors. *Int. J. Biochem. Cell Biol.* **2011,** *43*, 60-64.

5. Ano Bom, A. P. D.; Rangel, L. P.; Costa, D. C. F.; de Oliveira, G. A. P.; Sanches, D.; Braga, C. A.; Gava, L. M.; Ramos, C. H. I.; Cepeda, A. O. T.; Stumbo, A. C.; De Moura Gallo, C. V.; Cordeiro, Y.; Silva, J. L. Mutant p53 Aggregates into Prion-like Amyloid Oligomers and Fibrils: Implications For Cancer. *J. Biol. Chem.* **2012,** *287*, 28152-28162.