**Michigan
Technological
University**

Michigan Technological University
Digital Commons @ Michigan Tech

Dissertations, Master's Theses and Master's Reports

2016

Effect of disulfide bond scrambling on protein stability, aggregation, and cytotoxicity

Colina Dutta

Michigan Technological University, ctdutta@mtu.edu

Copyright 2016 Colina Dutta

Recommended Citation

Dutta, Colina, "Effect of disulfide bond scrambling on protein stability, aggregation, and cytotoxicity", Open Access Dissertation, Michigan Technological University, 2016.
<https://doi.org/10.37099/mtu.dc.etdr/129>

Follow this and additional works at: <https://digitalcommons.mtu.edu/etdr>

Part of the [Biochemistry, Biophysics, and Structural Biology Commons](#), and the [Chemistry Commons](#)

**EFFECT OF DISULFIDE BOND SCRAMBLING ON PROTEIN STABILITY,
AGGREGATION, AND CYTOTOXICITY**

**By
Colina Dutta**

A DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

In Chemistry

MICHIGAN TECHNOLOGICAL UNIVERSITY

2016

©2016 Colina Dutta

This dissertation has been approved in partial fulfillment of the requirements for the Degree of DOCTOR OF PHILOSOPHY in Chemistry.

Department of Chemistry

Dissertation Advisor: *Dr. Ashutosh Tiwari*

Committee Member: *Dr. Haiying Liu*

Committee Member: *Dr. Lanrong Bi*

Committee Member: *Dr. Chandrashekhhar Joshi*

Department Chair: *Dr. Cary Chabalowski*

This dissertation is dedicated to

“My Mother

Table of Contents

Preface.....	viii
Acknowledgements.....	ix
Abstract.....	xi
Chapter 1. Introduction	1
1.1 Proteins.....	1
1.2 Structure of Proteins.....	2
1.3 The Protein Folding Problem	5
<i>1.3.1 Energetics of Protein Folding</i>	6
<i>1.3.2 The Misfolded Scenario</i>	9
1.4 Protein Aggregation	11
<i>1.4.1 Types of Aggregates</i>	13
<i>1.4.2 Protein Aggregation and Neurodegenerative Diseases</i>	15
1.5 The Model Proteins	19
<i>1.5.1 Human Insulin</i>	20
<i>1.5.2 Lysozyme</i>	22
<i>1.5.3 Serum Albumin</i>	23
1.6 Aims of Research	25
References	28

Chapter 2. Methods	33
2.1 Spectroscopy	33
2.1.1 Ultraviolet (UV)-Visible Spectroscopy	34
2.1.2 Fluorescence Spectroscopy	37
2.1.3 Fourier-Transform Infrared Spectroscopy	45
2.2 Microscopy	49
2.2.1 Scanning Electron Microscopy	49
2.2.2 Atomic Force Microscopy	54
2.3 Cell Based Assays	56
2.4 Polyacrylamide Gel Electrophoresis	58
2.5 Mass Assisted Laser Desorption/Ionization (Time-of-Flight) Mass Spectrometry	62
References	65
Chapter 3. Preformed Seeds Modulate Native Insulin Aggregation Kinetics	68
Abstract	68
Introduction	70
Materials and Methods	74
Results	78
Discussion	111

Conclusion.....	121
References	122
Chapter 4. Disulfide Bond Scrambling at Acidic pH Promotes Insulin Aggregate	
Toxicity	132
Abstract	133
Introduction	134
Materials and Methods.....	138
Results & Discussion	153
Conclusion.....	164
References	165
Chapter 5. Protein aggregation in a crowded environment.....	
Abstract	174
Introduction	175
Materials and Methods.....	177
Results & Discussion	180
Conclusion.....	189
References	190
Chapter 6. Conclusion and Future Directions.....	
Appendix.....	196

Permission for Figures 1.1, 1.2 and 2.12.....	196
Permission for Figure 1.3	197
Permission for Figure 1.4.....	198
Permission for Figure 1.5	199
Permission for Chapter 3.....	200

Preface

Contents of Chapter 1, 2, and 6 are written by the author of this dissertation. Chapters 3, 4 and 5 are edited and organized by the author based on previously published paper (Chapter 3) and a manuscript in preparation (Chapter 5), details of which are described below.

Chapter 3 is based on a five-author article that was published in *The Journal of Physical Chemistry* **2015**, 119:15089-99. The author of this dissertation performed majority of the experiments, analyzed the data and wrote the paper. Ms. Mu Yang performed the ANS and Bis-ANS fluorescence experiments and Dr. Fei Long and Dr. Reza Shahbazian-Yassar helped with AFM imaging and analysis, respectively. Dr. Ashutosh Tiwari is the corresponding author and he revised the manuscript.

Chapter 4 is based on a two-author article that is ready to be submitted in *Chemical Sciences (RSC Publications)*. The author of this dissertation performed all of the experiments, analyzed the data and wrote the paper. Dr. Ashutosh Tiwari will be the corresponding author and will revise the manuscript.

Chapter 5 is a three-author article for which the manuscript is under preparation. The author of this dissertation performed majority of the experiments, analyzed all the data and wrote the paper. Ms. Rashmi Adhikari performed the electrophoresis experiments. Dr. Ashutosh Tiwari will be the corresponding author and will revise the manuscript.

Acknowledgements

Pursuing doctoral study is a task that can never be possible without the help and support of so many people who have touched my life in their own ways. I am indebted to all!

First and foremost, I would like to express my sincere gratitude towards my advisor Dr. Ashutosh Tiwari for his continuous support, inspiration, and motivation for my Ph.D. study and related research. His unflinching guidance and impeccable patience succored me all the way through writing my dissertation.

Besides my advisor, I am thankful to the rest of my committee members: Dr. Haiying Liu, Dr. Lanrong Bi and Dr. Chandrashekhar Joshi for their precious time, insightful comments and encouragement.

I also owe my deepest gratitude towards Dr. Ravindra Pandey, without whom my Ph.D. journey would not have been possible. I am also thankful to Dr. Ranjit Pati for his encouragement, suggestions and support both on personal and professional front.

I would like to convey my sincere thanks to our former and current Department Chair Dr. Sarah Green and Dr. Cary Chabalowski, respectively. I am also grateful to the Department of Chemistry for the continuous funding support that made my journey easier and less stressful. Million thanks to Celine Grace, Denise Laux, Charlene Page and Margaret Dunsten for their assistance with a smile!

I would also like to thank my colleagues Dr. Jagadeesh Janjanam, Nethaniah Dorh, Mu Yang and Rashmi Adhikari for all the scientific and non-scientific discussions and for being there. Special thanks to Nethaniah for being extremely supportive during

desperate times. I can't thank Jagadeesh enough for patiently teaching me everything I wanted to learn. He has been a continuous support as a postdoctoral researcher and a friend in the Tiwari Lab and even after that.

Teaching has been an integral part of my graduate studies. I would like to thank my teaching supervisors Sarah Hill, Laleh Vahdat, Aparna Pandey, Kelley Smith, and Lorry Reilly, who taught me the 'art of teaching'.

I would also like to thank some very special friends (Neil (Rahul) Bose, Noopur Sharma, and Yogita Uplane) for their unique contributions towards my journey.

I am deeply indebted to my parents and my in-laws for their love and faith in me.

Last, but not the least, I owe everything to my husband Partha, without whom I could not have reached here. His immense faith, sacrifice, dedication and encouragement towards me, helped me sail through my doctoral studies.

Abstract

Proteins are nano-machines that carry out majority of the cellular functions. Thermodynamically they are functional and stable within a very narrow range (1 kcal/mol). External perturbations in the form of pH change, thermal, or oxidative/reducing stress can destabilize the protein resulting in misfolding and aggregation. Prolonged environmental stress can affect the cells adaptive response resulting in loss of ability to refold or recycle proteins. This can lead to accumulation of misfolded or aggregated proteins within the cell. Such accumulation of aggregated proteins have been associated with neurodegenerative disorders such as Amyotrophic Lateral Sclerosis (ALS), Parkinson's, Huntington's, and Alzheimer's disease. There is a general consensus among scientists that aggregated proteins cause disease by a '*toxic gain of function*'. However, there is a huge debate among scientists on what are the toxic forms of protein aggregates. This is largely due to lack of research that have looked at the relationship between morphology of aggregates and their toxicity. Therefore, such studies that can help clarify the relationship between morphologically different forms of aggregates and their associated toxicity are needed.

In this dissertation, we study how disulfide reducing environment can impact protein stability, aggregation, and cytotoxicity. We also study effect of molecular crowding agent polyethylene glycol (PEG) (600 and 2000 g/mol) on providing protection against reducing and thermal stress. Disulfide bonds are covalent interactions that provide major stability to the protein in conjunction with several non-covalent interactions such as hydrophobic interactions, hydrogen bonding, Van der Waal and electrostatic

interactions. These interactions help the proteins to fold into its native three-dimensional fold. However, the cytoplasmic environment in cells is highly reducing and can compromise the disulfide bonds leading to protein aggregation. To better understand how disulfide bond *scrambling* can affect protein aggregation, we used insulin as a model protein. We made *seeds* of insulin by incubating the protein in presence of disulfide reducing agent for varying lengths of time at 37 °C. These seeds were labeled as ‘*nascent*’, ‘*intermediate*’, or ‘*mature*’ based on their ability to induce and promote aggregation of native protein with different kinetics. *Nascent seeds* promoted fastest insulin aggregation and formed amorphous aggregates. In another related study we used a combination of pH (acidic to basic) and temperature (37 and 65 °C) to generate morphologically different types of insulin aggregates under disulfide reducing/non-reducing conditions. These aggregates were characterized by different techniques and tested for their toxicity on SH-SY5Y cells. Cytotoxicity studies of insulin aggregates on neuroblastoma cells showed that aggregates formed from disulfide reduced proteins at acidic pH were more toxic compared to the aggregates formed at neutral or basic pH. Lastly, we wanted to study how these properties could be impacted by molecular crowding. We mimicked the intracellular crowded milieu *in vitro* by using PEG to investigate the effect of crowding on lysozyme stability and aggregation under thermal and reducing stress. We observed that PEG-2000 stabilized the molten globule intermediate of lysozyme in the presence of a non-thiol based reducing agent. Overall, the results indicates an intricate relationship of pH, temperature, and reducing environment impacting proteins aggregation and its associated toxicity.

Chapter 1. Introduction

1.1 Proteins

Proteins form the center of any biological process. They essentially are the building blocks of life that are composed mainly of 20 standard amino acids. The sequence of amino acid in a protein, encodes for its function; that in turn dictates its interaction with other biological molecules. The function of proteins range from maintaining cell shape (structural proteins) to catalyzing biochemical reactions within the cell (enzymes). Proteins also function as hormones that act as chemical messengers and direct processes within the living system.

The amino acids within a protein, are covalently bonded together by the peptide bonds between the alpha carbon ($C\alpha$) atoms of amino (N-terminal) and carboxyl (C-terminal) groups of adjacent amino acids. Linkage between amino acids occurs by a condensation reaction in which a water molecule is expelled leading to the formation of the peptide bond. Studies by Pauling and Corey revealed that the alpha carbon atoms in the peptide bond are separated by three covalent bonds $C\alpha-C-N-C\alpha$. X-ray diffraction studies revealed that the C-N bond in a co-planar peptide bond is slightly shorter than that present in an amine. This suggests the presence of a partial negative charge on oxygen and a positive charge on nitrogen. All these observations lead to the conclusion that the C-N bond within a peptide bond is rigid and free rotation is restricted due to partial double bond character. However, free rotation is possible around N- $C\alpha$ and C- $C\alpha$ bonds. The rigidity of the peptide bond in a protein provides

the structural basis in proteins and defines the secondary structure and tertiary fold in the protein.

1.2 Structure of Proteins

The primary structure of a protein is formed by its amino acid sequence. Amino acids have a positively charged amino group and a negatively charged carboxyl group around neutral pH. They differ by the alkyl (R) group that can be broadly divided into hydrophobic and hydrophilic. Uncharged as well as aromatic alkyl groups are hydrophobic whereas charged ones are hydrophilic in nature. The folding of a protein is determined by the amino acid sequence since the hydrophobic residues form the core of the protein while the charged hydrophilic residues interact with the solvent in the cellular milieu. This 'preference' of interaction requires the protein backbone (peptide bond) to fold into the hydrophobic core, although it is highly polar and hydrophilic. This leads to the formation of secondary structures: α -helix and β -sheet that help in exposing the peptide bond to the solvent while the hydrophobic amino acids form the core of the folded protein.

The α -helix is a common element of secondary structure in proteins and is predominantly right-handed (Figure 1.1). The helical structure forms in a way such that the N-H group of the n th residue of the protein is hydrogen bonded to the C=O group of the $(n-4)$ residue. This ensures tight packing and Van der Waal contact throughout the helix. The alkyl groups point outward from the helix to minimize steric effects.

Figure 1.1 Right handed helix with the planes of the peptide bond parallel along the long axis of helix (shown as the rod)¹. From: Lehninger Principles of Biochemistry 6E, by David L. Nelson and Michael M. Cox, Copyright 2013, by W.H. Freeman and Company. Used by permission of the publisher.

Another element of the secondary structure in protein is the β -sheet (Figure 1.2). The protein backbone is fully extended in this pleated structure and hydrogen bonding exists between the NH and C=O groups of adjacent polypeptide chains. β -strands arrange themselves in two ways in order to form β -sheet, namely parallel and anti-parallel.

Hydrogen bonded polypeptide chains extend in the same and opposite directions for parallel and anti-parallel strands, respectively. Hydrogen bonding in parallel β -sheet is distorted, making it less stable compared to the anti-parallel arrangement.

Figure 1.2. Top view of (A) antiparallel and (B) parallel β -sheet. R-groups extend out from the β -sheet (purple spheres) and hydrogen bonding between adjacent chains are shown as blue lines ¹. From: Lehninger Principles of Biochemistry 6E, by David L. Nelson and Michael M. Cox, Copyright 2013, by W.H. Freeman and Company. Used by permission of the publisher.

The third category of secondary structure is called ‘turns’ or ‘loops’ that connect the α -helix and β -sheets in a protein. These are flexible and irregular bends containing at least four amino acid residues hydrogen bonded to each other.

The elements of secondary structure of a protein combine together to form the tertiary structure that is the three-dimensional folded protein. The tertiary structure ensures that amino acids that are far apart in the sequence interact with each other via weak interactions such as hydrogen bonding, electrostatic interactions and/or hydrophobic interactions. Additionally, covalent bond formation is also observed in the form of

disulfide bonds between cysteine residues. The native folding of the protein establishes a stable conformation with minimum free energy. The stable conformation of a folded protein is attained when the hydrophobic amino acids form the interior of the protein whereas the charged residues interact with the aqueous milieu. Another dimension of stability is added by the formation of disulfide bonds. These covalent bonds are formed by oxidation of sulfhydryl groups of cysteine residues and confer stability to the protein in a relatively hostile extracellular environment.

Many proteins have multiple polypeptide chains (subunits) that interact with each other via hydrogen bonding, electrostatic interactions and/or disulfide bonds. The subunits associate themselves in a geometric arrangement and this spatial configuration forms the quaternary structure of the protein. Proteins with multiple subunits are prevalent mostly in metal binding proteins and enzymes. This provides a structural basis as a binding site for metal or receptors for proper functioning of the protein/enzyme. The quaternary structure of proteins forms the top of the protein-folding pyramid.

1.3 The Protein Folding Problem

Proteins need to fold into a three dimensional structure in order to function. The folding process was initially postulated to be a template driven process. However the limiting factor was the understanding as to how these “templates” formed. Anfinsen, in his experiment with Ribonuclease A, showed that the protein can be denatured and refolded and the process is reversible. In 1973, this became known as Anfinsen’s hypothesis that the amino acid sequence of a protein determines its three dimensional folding². Although this hypothesis explained the thermodynamics of protein folding to

a great extent, the kinetics of protein folding was still unclear. In an attempt to understand the protein folding process, the folding pathways need to be elucidated. If proteins were to fold by randomly exploring all possible conformations, it would take approximately billions of years to fold. However, protein-folding process is in the order of microseconds. This inference, known as Levinthal paradox, lead to the hypothesis that protein folding is an ordered process in which the formation of the native fold is accompanied by the increase in conformational stability and decrease in free energy³.

1.3.1 Energetics of Protein Folding

The problem of protein folding is two-fold. The polypeptide chain needs to fold in its lowest energy conformation and also in doing so, must fold rapidly. Hence, the protein folding phenomena entails both thermodynamics and kinetics. The native state of the protein is favored by only 5-15 kcal/mol free energy. This means that the free energy of the native state of the protein is marginally lower than its denatured state. Protein folding is an entropically unfavorable process that is compensated by the decrease in enthalpy according to the second law of thermodynamics.

$$\Delta G = \Delta H - T\Delta S \quad (1.1)$$

Van der Waal forces and electrostatic interactions stabilize the native fold and are unique for every protein. Proteins with conserved amino acid residues fold similarly, and it arises from the fact that interactions that stabilize a particular fold become destabilizing if the protein deviates from the native fold at physiological conditions. Hence the free energy of a folded protein is at a 'deep' energy minimum that is significantly distant from any destabilized protein folds. This gives rise to the classic

view that protein folding occurs via well-defined intermediates. Wolynes, Baldwin and Dill postulated the landscape theory⁴⁻⁶ where the energy surface of a folding protein is funnel shaped with the bottom of the funnel representing the native folded protein with minimum global free energy (Figure 1.3⁵). The total binding energy required for a protein to fold is around $100 K_B T$, significantly larger than the barrier energy (approximately $10 K_B T$)⁵. Proteins fold into many native-like conformations that are close to the energy minimum before reaching the native fold. This is accounted for by the rugged topography of the funnel landscape of protein folding. Hence, a protein might also get trapped in a local minimum until it acquires the energy to overcome the barrier energy and continue the folding process.

Figure 1.3. Schematic of funnel-like landscape of protein folding. Native fold signifies lowest free energy⁵.

The kinetics of protein folding has been explained by the advent of experimental techniques such as NMR, mass spectroscopy, and hydrogen exchange studies as well as statistical mechanics theoretical studies. The progression of protein folding is monitored using the variable – fraction of native contacts that goes from 0 in the case of unfolded protein to 1 for the native state. For an unfolded protein, the number of conformational ensembles is at a maximum that reduces throughout the folding process

and defines the shape of the folding funnel⁶. The normalized time taken (τ_f) for a protein to fold and reach the global minima in the funnel is given by

$$\tau_f = \int_{n_u}^{n_f} \tau(n) dn \quad (1.2)$$

where the depth of the energy landscape funnel (n) varies from the unfolded state of the protein (n_u) to the folded state (n_f).

1.3.2 The Misfolded Scenario

The non-covalent interactions that stabilize protein structure are weak in nature. Hence marginal deviation beyond physiological conditions leads to protein instability. In the presence of destabilizing conditions such as removal of water and/or change in pH, temperature, salt concentration etc, can lead to the inter-conversion of the protein from its native state to unfolded state. The stability of a protein is expressed as the difference between the free energy between the unfolded (U) state and the native fold (N).

$$N \rightleftharpoons U \quad (1.3)$$

The unfolded and native states have been studied experimentally using various spectroscopic and calorimetric techniques. It has been shown that native protein stability is not only due to the weak intermolecular interactions but also due to the protein-solvent interactions. However, experimental and theoretical studies have confirmed that protein structure is stabilized largely due to its hydrophobic interactions. Calculations using lysozyme have shown that hydrophobic and electrostatic interactions contribute 450 and 87 Kcal/mol respectively towards the free energy of

denaturation at room temperature⁷. These values are similar for globular proteins. At room temperature, hydrophobic free energy is mainly entropic that can be compensated by arranging the solvent molecule around the protein. Hence, hydrophobic interaction favors burial of the solvent exposed protein surface.

Amidst all the favorable energy situations, the equilibrium of the protein towards the native state shifts towards unfolding due to stresses that lead to misfolding of the protein. Most globular proteins misfold by exposing the hydrophobic residues to the aqueous milieu. These processes are irreversible for most cases. This irreversible shifting of the equilibrium to the unfolded state leads to the hypothesis that misfolded proteins are also thermodynamically stable. Misfolded proteins come together in highly ordered forms to form protein aggregates. The aggregation pathway, stability of aggregates and associated energetics have been explained by the Lumry-Eyring framework as shown in scheme 1.4, where the native protein (N) undergoes reversible conformational change to form transition ensemble (TS*) that forms an intermediate (A_I) after irreversible re-arrangement⁸. The misfolded intermediate A_I associates with an aggregated protein molecule A_m, thereby continuing the aggregation process.

This framework is true for all first order reactions and can be extended for bimolecular aggregation reactions as well. This stability of the aggregate form is attained by the interactions between the aggregated protein and the destabilizing factors that help in

stabilizing the aggregate form. This is explained using the free energy diagram in Figure 1.4

Figure 1.4. Free energy diagram of the Lumry-Eyring framework explaining the energetics of protein aggregation and stability of aggregates. Kinetic energy barriers are shown using curved lines⁸.

1.4 Protein Aggregation

The process by which proteins unfold and associate themselves in *clusters* is called protein aggregation. The striking resemblance between different aggregated protein forms in spite of disparate amino acid residues suggests a common underlying mechanism of protein aggregation. This has led to the hypothesis of a generic

mechanism of protein aggregation. Destabilizing factors such as pH, temperature, solvent etc. disrupt the native interactions within the protein as well as between the protein and aqueous media. This exposes the hydrophobic amino acid residues leading to aggregate formation. This mechanism of protein aggregation is termed as *off-pathway*. Depending on the pathway followed, distinct aggregate forms are generated. Energetically speaking, protein folding via the *off-pathway* route leads to aggregates with an altered global minima (Figure 1.5)

Figure 1.5. Energy landscape in case of protein folding and aggregation. Smooth funnel is shown in light grey whereas the ruggedness increases with the complexity of a protein leading to alternate folding forms (aggregates) shown in dark grey funnel⁹.

1.4.1 Types of Aggregates

Protein aggregates vary in morphology depending on the perturbation that leads to aggregation¹⁰. Highly ordered aggregate forms, known as amyloid fibrils, have been the focus of study for a long time due to their association with neurodegenerative

diseases such as Alzheimer's and Parkinson's disease¹¹. Amyloid deposits are found within the brain and also major organs such as the liver and the spleen. These aggregates exhibit characteristic properties such as birefringence under plane polarized optical microscope, cross-beta X-ray diffraction pattern and tight binding with amyloid specific dyes such as thioflavin T¹². Information gathered using these characterization methods reveal that amyloids are rich in β -sheet structure that runs perpendicular to the fibril axis¹³. Amyloid fibrils have been observed using many unrelated proteins¹⁴ and polymers suggesting a generic mechanism of aggregation¹⁵, although the propensity of fibril formation varies with proteins and is believed to be sequence-dependent¹⁶. The mechanism of fibril formation highlights the formation of intermediates and oligomers. They accelerate fibril growth both by nucleation-dependent and chain polymerization methods¹⁷.

However apart from fibrils, other disordered aggregate forms (loosely termed as amorphous aggregates) have been found both in *in vitro* and *in vivo* studies¹⁸. These disordered aggregates have been broadly divided into particulates and spherulites¹⁹. Particulates²⁰ are bead-like amorphous protein aggregates that have been found predominantly in β -lactoglobulin under thermal and pH stress²¹. This form of aggregate has not yet been associated with any medical relevance, however protein particulate formation is of great concern in the food industry. When heated, the protein forms clear fibrillar gels away from its isoelectric point whereas at the isoelectric point, it forms particulates. These monodisperse particulates are a result of partial denaturation of β -lactoglobulin with average sizes varying between 1 to 1.5 μm . The size of these

particulates changes with heating rate and temperature. Rapidly heating the protein to 100 °C results in smaller particulates with average size of 500 nm, whereas slow heating (1 °C/min) to 80 °C leads to bigger size distribution²². The mechanism of protein aggregation to form particulates has been elucidated in some studies. One model suggests that it is similar to a free radical polymerization method with initiation, propagation and termination steps. In this scenario, cysteine residues have been hypothesized to play the role of free radicals in disulfide-rich proteins. Another model explains protein particulate formation using nucleation and growth mechanism. Unfolded proteins form the nucleus that stimulates the formation of aggregates and increases the kinetics of aggregation. This method was used to explain the varying particle size of the particulates with respect to change in pH and temperature.

Spherulites are another form of disordered aggregates that exhibit a characteristic maltese-cross pattern when viewed under the cross polarizers of an optical microscope²³. These are radial arrangements of amyloid fibrils with a central core. However, the reason behind the equilibrium shift from amyloid fibrils to spherulites and its energetics is still unknown. Spherulitic growth has been observed in lysozyme, β -lactoglobulin, A β -42 and insulin among other proteins²⁴. They also have been observed within amyloid plaques of Creutzfeld-Jacob disease, Down's syndrome and rat model Alzheimer's disease²⁵.

1.4.2 Protein Aggregation and Neurodegenerative Diseases

Neurodegenerative diseases such as Alzheimer's disease (AD), Parkinson's disease, Amyotrophic Lateral Sclerosis (ALS) and other forms of dementia are a class of late

onset diseases that affect the central nervous system²⁶. These diseases are characterized by the presence of abnormal proteinaceous deposits in the brain. Morphology of these deposits are predominantly fibrillar with characteristic properties common for all neurodegenerative diseases. These deposits are formed by misfolded and aggregated proteins. Proteins aggregate due to mutation in the sequence, genetic alteration of the protein or external stresses and aging. Protein folding in diseased condition has been proposed to be a cause of neurodegeneration, rather than an effect. This is supported by pathological and genetic information about the disease, animal model studies and also biophysical characterization of aggregated proteins *in vitro*²⁷.

The most common form of neurodegenerative disease is ALS. It was first described by French neurobiologist Jean-Martin Charcot in 1869 and was known as Charcot's sclerosis until the 1930's when baseball player Lou Gehrig was affected by this debilitating condition²⁸. ALS is now familiarly known as Lou Gehrig's disease. The hallmarks of this fatal condition are progressive muscle weakness, spasticity, hyperreflexia, and atrophy due to degeneration of motor neurons in the brain and spinal cord²⁹. 5-10% of this disease is linked to genetic mutations meaning that they are familial, however almost 90% is sporadic in nature.

Mutations in the superoxide dismutase 1 (SOD1) protein contribute to about 20% of familial ALS cases. SOD1 is a 32 kDa homodimeric metallo-enzyme that is predominantly found in the cytosol³⁰. Each subunit in the homodimer, form eight stranded Greek key β -barrel containing active sites for binding copper and zinc³¹. The

main function of SOD1 is free radical scavenging in the body, thereby preventing oxidative and mitochondrial damage to cells.

Of the more than 153 SOD1 mutants that have been studied so far, most of them are missense mutations, with only a few C-terminal truncated forms³². As a part of the cells' protein degradation system, misfolded proteins are targeted through ubiquitylation. These misfolded proteins have toxic effects on the proteosomal pathway and autophagy leading to the presence of ubiquitylated aggregates in patient brains³³.

The relationship between presence of misfolded SOD1 in the brain and occurrence of ALS is unknown. However, it is established that aggregated SOD1 confers a toxic gain-of-function rather than a loss of normal function. This property is independent of the various SOD1 mutants that have been reported so far. The observations supporting this argument have been drawn from mice model studies³⁴. However, SOD1 null mice did not develop the disease at all. Transgenic mice studies have been carried out for more than 100 mutations and all SOD1 mutant aggregates have been ubiquitin positive with presence of intracellular aggregates of the protein in motor neurons and astrocytes. Effects of SOD1 mutations are disparate where some mutants (G37R) do not lose normal activity. Additionally, no correlation has been found between activity of SOD1 protein and the disease onset and progression. These lack of correlations make the scenario of understanding how SOD1 causes motor neuron death highly overwhelming.

Apart from SOD1, several other proteins have been identified within the pathophysiology of ALS. Approximately 90% of sporadic ALS cases have been identified to be a result of mutations in the TARDBP gene³⁵. This leads to aggregation

of TAR DNA binding protein (TDP-43) and is a major component of ubiquitinated inclusions in ALS patients. Aggregates of the wild type TDP-43 are also found in the motor neurons of the spinal cord, hippocampal neurons and glial cells of various motor neuron diseases. Hence, diseases linked to TDP-43 deposits have been termed as *TDP-43 proteinopathies*.

TDP-43 is a 43 kDa nuclear protein that binds to the RNA and DNA via its RNA recognition motifs (RRM). The entire protein has 2 RRMs, a nuclear localization sequence (NLS) and a glycine rich flexible C-terminal that mediates protein-protein interactions. The main function of TDP-43 is RNA splicing regulation and modulation of microRNA biogenesis³². It is also involved in cytoplasmic stress granule response and is a key component of dendritic RNA transport granules in neurons.

Currently around 38 mutations in the TARDBP gene have been reported in ALS cases and most mutations are on the region encoding the C-terminal end of the protein³⁶. In the diseased conditions, TDP-43 localizes to the cytoplasm from the nucleus accompanied by C-terminal degradation of the protein. Hyper-phosphorylated forms of the protein are a common hallmark of ALS, and the extent of disease progression is generally gauged by the spread of proteinaceous deposits from spinal and cortical neurons to other cortical regions³⁷. However, the relationship between TDP-43 proteinopathy and pathology is still unknown. A number of studies using animal models have confirmed that TDP-43 aggregation not only leads to toxic gain-of-function (like SOD1), but also causes loss of normal function of the protein. While both

loss and gain plays a role in disease progress, the link between the two is yet to be established.

1.5 The Model Proteins

Protein folding is a complex process and the complexity increases when they misfold and aggregate. The main aim of studying the molecular mechanism of protein aggregation is to understand the disease process associated with these proteins for possible therapeutics. Owing to the complexity of disease-related proteins and the vast pool of unanswered questions, dissection of the molecular mechanism behind aggregation is very difficult. Therefore it is worthwhile to study the mechanism of aggregation in proteins that are well characterized, in order to provide a broader picture of the aggregation process. The motivation of studying model proteins and extrapolating the information to disease causing complex proteins is due to the generic mechanism of aggregation that is observed among unrelated proteins. A large number of proteins such as carbonic anhydrase, myoglobin, β -lactoglobulin, lysozyme, bovine serum albumin (BSA), insulin etc. have been candidates for model protein. All these proteins aggregate under destabilizing conditions that can be controlled and the aggregation mechanism can be elucidated.

We chose to study human insulin, lysozyme, and BSA as our model proteins due to their diversity in molecular weight, disulfide bond status and secondary structure. Also, all three proteins have a very high propensity to aggregate and two of these (insulin and lysozyme) have been linked to some form of amyloid diseases.

1.5.1 Human Insulin

Insulin was discovered in 1920 by Dr. Frederick Banting, and is considered as one of the most important advances in the field. This protein played a pivotal role in advancing knowledge in peptide chemistry, structural biology and recombinant DNA technology. This 5.8 kDa hormone continues to be a subject of research owing to its use by insulin-dependent diabetes patients and its aggregation propensity under various conditions *in vivo* and *in vitro*. Still today, insulin serves as a model protein for studying protein aggregation due to its small size, presence of highly conserved disulfide bonds and well defined secondary structure (Figure 1.6).

Figure 1.6 Ribbon structure of human insulin hexamer. Figures generated using PyMOL 1.3 from PDB file 1AI0

Insulin is a 51 amino acid residue polypeptide with 2 chains A and B³⁸. It is synthesized in the body as proinsulin (longer chain) and is cleaved off to release the active insulin. This messenger hormone is secreted by the pancreas into the blood where it binds to its receptor and helps in monitoring blood glucose in the body.

Comparison of phylogenetic data and studies using various biophysical techniques show that the majority of amino acid residues are conserved in the insulin structure, across 100 vertebrates that were studied³⁹. The conserved areas are those that bind to the insulin receptor for proper functioning of the protein, suggesting that the structure and function have been conserved throughout evolution. The structural features that have been conserved are the chain length, position as well as number of cysteine residues and 16 of the 51 amino acid residues.

As with other eukaryotic organisms, insulin undergoes post-translational modifications that are required for its native fold as well as function. One of the important modifications is the formation of disulfide bonds that help maintain the stability and biological activity of insulin. This is ensured by the fact that insulin disulfides are thoroughly conserved between residues A6-A11, A7-B7 and A20-B19 as they hold the structure of the protein. The A chain forms a U-shape and sits on the B chain with hydrophobic interactions between 5 conserved leucine residues holding them together.

The importance of each of the conserved disulfide bonds has been elucidated using protein-engineering methods. Substituting the cysteine in the A6-A11 bond leads to decreased binding affinity of insulin. This has been attributed to the loss of hydrophobic interaction in the A chain that majorly rests on the disulfide bridge between A6-A11.

Substitution of cysteine in A7-B7 has led to diverse outcomes. Studies have shown that these insulin analogs show substantially low biological activity, binding affinity and 40% potency in mouse convulsion assay. Deletion or substitution of the C-terminal disulfide bridge A20-B19 has led to very poor expressions in yeast systems. Overall, the inter-chain disulfide residues help in maintaining the parallel orientation of chains A and B whereas the intra-chain disulfide bridge directs the folding of the protein by bringing the N and C terminal in close proximity that forms the insulin receptor binding site.

1.5.2 Lysozyme

Lysozyme was accidentally discovered by Dr. Alexander Fleming in 1921 during his search for medical antibiotics. It is a 14.3 kDa globular protein with 129 amino acid residues. The relatively small protein is abundantly found in tears and hen egg white. Lysozyme is a glycosidase enzyme that attacks the peptidoglycan cell wall of bacteria and hydrolyses them, thereby imparting anti-bacterial properties.

This protein consists of two subunits, one predominantly α -helical and the other mainly β -sheet structure (Figure 1.7). The three-dimensional structure of lysozyme is stabilized by four disulfide bridges between C6-C127, C30-C115, C64-C80, and C76-C94. The disulfide bridge between C6-C127 keeps the N and C-terminal of the protein in close contact thus maintaining the protein stability and fold. Hence C6-C127 is partially solvent-exposed⁴⁰. However, the other three disulfide bonds are buried within the

hydrophobic core of the protein. Lysozyme serves as a model protein, not only because it is well characterized but also because it has been linked with systemic amyloidosis.

Figure 1.7 Ribbon structure of hen egg-white lysozyme. Figures generated using PyMOL 1.3 from PDB file 1UCO.

1.5.3 Serum Albumin

Serum albumin is abundant in the circulatory system with a concentration of 35-50 mg/mL in the blood plasma. It is a 66 kDa protein that transports various endogenous and exogenous compounds such as fatty acids, metals, amino acids, drugs etc. through the blood stream⁴¹. Transportation of various ligands occur by the binding of water-insoluble molecules to the hydrophobic pocket of BSA. Apart from being a transport

protein, BSA also helps in maintaining blood pH and osmotic pressure. BSA is majorly α -helical with three homologous domains (I, II and III), 17 intra chain disulfide bridges and one free cysteine at position 35. Most of the disulfide bonds are buried within the core of this hydrophobic protein around neutral pH and the native fold is characterized by formation of nine disulfide loops. However, with increase in pH and temperature beyond physiological, some disulfide bridges become solvent accessible due to partial unfolding of the protein⁴⁰.

BSA maintains 76% homology with human serum albumin suggesting the importance of conserving the function of serum albumin protein. Serum albumin is considered a model protein due to its high molecular weight, high hydrophobicity and presence of large number of disulfide bonds that stabilize the structure of this globular protein.

Figure 1.8 Ribbon structure of bovine serum albumin. Figures generated using PyMOL 1.3 from PDB file 4F5S.

1.6 Aims of Research

Protein misfolding and aggregation have been linked to Alzheimer's disease, Parkinson's disease, Amyotrophic Lateral Sclerosis (ALS) and other forms of dementia that are neurodegenerative in nature. The molecular mechanism of these diseases needs to be investigated in order to continue the search for therapeutics. The main players identified in these diseases are SOD1, TDP-43, α -synuclein, and A β -42, among others. However, information on the structure and folding of these proteins are very limited. This makes understanding the aggregation process of these proteins and its relationship to diseases very complex.

Proteins aggregate due to various destabilizing effects, and one of the most crucial destabilizing factors is breakage of disulfide bonds. Disulfide bonds are covalent posttranslational modifications that predominantly have two functions. First, they influence the thermodynamics of protein folding by destabilizing and lowering the entropy of the unfolded protein and shifting the equilibrium to the native fold. Second, they maintain protein integrity, protect against proteolysis and increase shelf life of proteins. However, the reducing cytosolic environment or deviations in pH and temperature compromise the disulfide bonds that trigger protein aggregation.

The main aim of this research is to understand the effects of reducing environment on the aggregation kinetics of the model protein insulin (Chapter 3). Understanding the aggregation of insulin in the presence of a reducing agent would be helpful since the protein is known to aggregate under storage conditions via disulfide bond scrambling. Insulin is also a disease related protein since it aggregates at the site of repeated injections in diabetes patients leading to injection amyloidosis. Aggregated insulin becomes the nucleus for further protein aggregation. We studied the kinetics of nucleation dependent insulin aggregation using preformed insulin *seeds*.

Alternate forms of protein aggregation have been found within amyloid plaques *in vivo*, however their toxicity is still unknown. Diverse forms of protein aggregates have also been observed in *in vitro* studies using various proteins. The possibility of formation of alternate protein aggregates leads to the second aim of this research (Chapter 4). It is to characterize the morphology of different types of aggregates formed at diverse pH and temperature range, and correlate them to the cellular toxicity.

The aim of studying protein misfolding and aggregation is to understand the mechanism of protein aggregation and stabilize the native fold of the protein. In the last part of this research, we study the effect of various macromolecular crowding agents on model protein - lysozyme aggregation (Chapter 5). These crowding agents are high molecular weight polymers that are known to protect the protein structure against aggregation and thus increase stability. We study the effect of these macromolecules on lysozyme aggregation under mild reducing as well as thermal stress, in order to investigate the protective role of these crowding agents against disulfide bond scrambling and aggregation.

Insulin aggregation kinetics under various conditions was studied using UV/Visible Spectroscopy, Fluorescence Spectroscopy, Fourier-transform infrared spectroscopy (FTIR), Mass Spectrometry (MALDI-MS). Morphology of the aggregates and nucleation dependent insulin aggregation was investigated using Scanning Electron Microscopy and Atomic Force Microscopy. All techniques used to accomplish the aims of this research are discussed in Chapter 2

References

- (1) Nelson, D.L.; Cox, M.M.; Lehninger, A.L. *Lehninger Principles of Biochemistry*. Fifth Edition ed.; W.H. Freeman and Company: New York, 2008.
- (2) Rose, G. D.; Fleming, P. J.; Banavar, J. R.; Maritan, A. A Backbone-Based Theory of Protein Folding. *PNAS* **2006**, *103* (45), 16623-33;
- (3) Morange, M. The Protein Side of the Central Dogma: Permanence and Change. *Hist. Phil. Life Sci.* **2006**, *28* (4), 513-524.
- (4) Zwanzig, R.; Szabo, A.; Bagchi, B. Levinthal's paradox. *PNAS* **1992**, *89*, 20-22.
- (5) Dinner, A. R.; Šali, A.; Smith, L. J.; Dobson, C. M.; Karplus, M. Understanding Protein Folding via Free-Energy Surface from Theory and Experiment. *Trends Biochem. Sci.* **2000**, *25*(7), 331-339
- (6) Dobson, C. M. Protein Folding and Misfolding. *Nature* **2003**, *426*, 884-890.
- (7) Onuchic, J. N.; Wolynes, P. G. Theory of Protein Folding. *Curr. Opin. Struct. Biol.* **2004**, *14* (1), 70-5.
- (8) Chi, E.Y.; Krishnan, S.; Randolph, T.W.; Carpenter, J.F. Physical Stability of Proteins in Aqueous Solution: Mechanism and Driving Forces in Nonnative Protein Aggregation. *Pharm. Res.* **2003**, *20* (9), 1325-1336.
- (9) Jahn, T. R.; Radford, S. E. Folding Versus Aggregation: Polypeptide Conformations on Competing Pathways. *Arch. Biochem. Biophys.* **2008**, *469* (1), 100-17.
- (10) Onuchic, J. N.; Socci, N. D.; Luthey-Schulten, Z.; Wolynes, P. G. Protein folding funnels: The Nature of the Transition State Ensemble. *Fold. Des.* **1996**, *1* (6), 441-450.

- (11) Dobson, C.M.; Šali, A.; Karplus, M. Protein Folding: A Perspective from Theory and Experiment. *Angew Chemie Intl Ed* **1998**, *37*, 868-893.
- (12) Chiti, F.; Dobson, C. M. Protein Misfolding, Functional Amyloid, and Human Disease. *Annu. Rev. Biochem.*, **2006**; *75*, 333-366.
- (13) Uversky, V. N. Amyloidogenesis of Natively Unfolded Proteins. *Curr. Alz. Res.* **2008**, *5* (3), 260-287.
- (14) Nilsson, M. R. Techniques to Study Amyloid Fibril Formation In Vitro. *Methods* **2004**, *34* (1), 151-160.
- (15) Toyama, B. H.; Weismann, J. S. Amyloid Structure: Conformational Diversity and Consequences. *Ann. Rev. Biochem.* **2011**, *80*, 557-585
- (16) Bucciantini, M.; Giannoni, E.; Fabrizio, C.; 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.
- (17) Auer, S.; Meersman, F.; Dobson, C. M.; Vendruscolo, M. A Generic Mechanism of Emergence of Amyloid Protofilaments from Disordered Oligomeric Aggregates. *PloS Comput. Biol.* **2008**, *4* (11), e1000222.
- (18) de la Paz, M. L.; de. Mori, G. M. S.; Serrano, L.; Colombo, G. Sequence Dependence of Amyloid Fibril Formation: Insights from Molecular Dynamics Simulations. *J. Mol. Biol.* **2005**, *349* (3), 583-596.
- (19) Frieden, C. Protein aggregation processes: In search of the Mechanism. *Prot. Sci.* **2007**, *16* (11), 2334-2344.

- (20) Krebs, M. R.; Domike, K. R.; Donald, A. M. Protein Aggregation: More than Just Fibrils. *Biochem. Soc. Trans.* **2009**, *37* (Pt 4), 682-6.
- (21) Krebs, M. R. H.; Devlin, G. L.; Donald, A. M., Protein Particulates: Another Generic Form of Protein Aggregation? *Biophys. J.* **2007**, *92* (4), 1336-1342.
- (22) Bromley, E. H. C.; Krebs, M. R. H.; Donald, A. M., Aggregation Across the Length-Scales in Beta-Lactoglobulin. *Farad. Discuss.* **2005**, *128*, 13-27.
- (23) Domike, K. R.; Donald, A. M. Thermal Dependence of Thermally Induced Protein Spherulite Formation and Growth: Kinetics of β -lactoglobulin and Insulin. *Biomacromolecules* **2007**, *8*, 3930-3937.
- (24) Rogers, S. S.; Krebs, M. R.; Bromley, E. H.; van der Linden, E.; Donald, A. M. Optical Microscopy of Growing Insulin Amyloid Spherulites on Surfaces In Vitro. *Biophys. J.* **2006**, *90* (3), 1043-54.
- (25) Krebs, M. R.; Bromley, E. H.; Rogers, S. S.; Donald, A. M. The Mechanism of Amyloid Spherulite Formation by Bovine Insulin. *Biophys. J.* **2005**, *88* (3), 2013-21.
- (26) Exley, C.; House, E.; Collingwood, J. F.; Davidson, M. R.; Cannon, D.; Donald, A. M. Spherulites of Amyloid-Beta 42 In Vitro and in Alzheimer's Disease. *J. Alzheimers Dis.* **2010**, *20* (4), 1159-65.
- (27) Ross, C.A.; Poirier, M. A. What is the Role of Protein Aggregation in Neurodegeneration? *Nat. Rev.* **2005**, *6*, 891-898.
- (28) Caughey, B.; Lansbury, P. T. Protofibrils, Pores, Fibrils, and Neurodegeneration: Separating the Responsible Protein Aggregates from the Innocent bystanders. *Ann. Rev. Neurosci.* **2003**, *26*, 267-98.

- (29) Cleveland, D.W.; Rothstein, J. D. From Charcot to Lou Gehrig: Deciphering Selective Motor Neuron Death in ALS. *Nat. Rev. Neurosci.* **2001**, *2*, 806-819.
- (30) Bruijn, L.; Miller, T.M.; Cleveland, D.W. Unraveling the Mechanism Involved in motor Neuron Degeneration in ALS. *Ann. Rev. Neurosci.* **2004**, *27*, 723-749.
- (31) Banci, L.; Bertini, I.; Boca, M.; Girotto, S.; Martinelli, M.; Valentine, J. S.; Vieru, M. SOD1 and Amyotrophic Lateral Sclerosis: Mutations and Oligomerization. *PLoS One* **3** (2), e1677.
- (32) Tiwari, A.; Liba, A.; Sohn, S. H.; Seetharaman, S. V.; Bilsel, O.; Matthews, C. R.; Hart, P. J.; Valentine, J. S.; Hayward, L. J. Metal Deficiency Increases Aberrant Hydrophobicity of Mutant Superoxide Dismutases That Cause Amyotrophic Lateral Sclerosis. *J. Biol. Chem.* **2009**, *284* (40), 27746-27758.
- (33) Robberecht, W.; Phillips, T. The Changing Scene of Amyotrophic Lateral Sclerosis. *Nat. Rev. Neurosci.* **2013**, *4*, 248-264.
- (34) Cleveland, D.W.; Liu, J. Oxidation Versus Aggregation — How do SOD1 Mutants Cause ALS? *Nat. Med.* **2000**, *6*, 1320-1321.
- (35) Bruijn L. I.; Houseweart, M. K.; Kato, S.; Anderson, K. L.; Anderson, S. D.; Ohama, E.; Reaume A.G.; Scott, R.W.; Cleveland D.W. Aggregation and Motor Neuron Toxicity of an ALS-Linked SOD1 Mutant Independent from Wild-Type SOD1. *Science* **1998**, *281* (5384), 1851-1854.
- (36) Scotter, E. L.; Chen, H. J.; Shaw, C.E. TDP-43 Proteinopathy and ALS: Insights into Disease Mechanism and Therapeutic Targets. *Neurotherapeutics* **2015**, *12*, 352-363.

- (37) Pesiridis, G.S.; Lee, V.M.; Trojanowski, J.Q. Mutations in TDP-43 Link Glycine-Rich Domain Functions to Amyotrophic Lateral Sclerosis. *Hum. Mol. Genet.* **2009**, *18* (2), R156-R162.
- (38) Neumann, M.; Kwong, K.; Kremmer, E.; Flatley, A.; Xu, Y.; Forman, M. S.; Troost, D.; Kretzschmar, H. A.; Trojanowski, J. Q. Lee, V.M. Phosphorylation of S409/410 of TDP-43 is a Consistent Feature in all Sporadic and Familial Forms of TDP-43 Proteinopathies. *Acta Neuropathol.* **2009**, *117*, 137-149.
- (39) Hua, Q. X. Insulin: a Small Protein with a Long Journey. *Protein & Cell* **2010**, *1* (6), 537-551.
- (40) Mayer, J. P.; Zhang, F.; DiMarchi, R. D., Insulin Structure and Function. *Biopolymers* **2007**, *88* (5), 687-713.
- (41) Yang, M.; Dutta, C.; Tiwari, A., Disulfide-bond Scrambling Promotes Amorphous Aggregates in Lysozyme and Bovine Serum Albumin. *J. Phys. Chem.B* **2015**, *119* (10), 3969-81.
- (42) Huang, B. X.; Kim, H. Y.; Dass, C., Probing Three-Dimensional Structure of Bovine Serum Albumin by Chemical Cross-Linking and Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2004**, *15* (8), 1237-47.

Chapter 2. Methods

Biophysical and biochemical characterization of protein aggregation kinetics along with its morphology was carried out using various spectroscopic, spectrometric and microscopic techniques. Mass assisted laser desorption mass spectrometry (MALDI-MS) and gel electrophoresis was used to analyze various species of protein aggregates present. Protein aggregation was monitored by measuring turbidity of protein solution at 600 nm using UV-visible spectroscopy. Protein unfolding, conformational change and exposure of hydrophobic pocket was monitored using intrinsic and extrinsic fluorescence. Secondary structure of protein aggregates were probed using Fourier transform infrared spectroscopy (FTIR). Morphology of protein aggregates was determined using scanning electron microscopy (SEM) and atomic force microscopy (AFM). The basic working principle of the techniques is discussed below.

2.1 Spectroscopy

Spectroscopy entails the study of the interactions between electromagnetic waves and molecules. Electromagnetic waves are radiations with electrical and magnetic forces. Spectroscopic measurements involve the effects associated with the electrical effects of the electromagnetic spectrum. The path of the electromagnetic wave is traced as a sinusoidal curve (Figure 2.1) where λ is the wavelength and A is the maximum amplitude of the wave.

Figure 2.1 Wave-like propagation of light. Light propagates in a sinusoidal path where one cycle is the distance between two troughs and the depth of a trough denotes the amplitude.

The electromagnetic spectrum spans a vast energy range between 10^{-5} nm (cosmic rays) to 10^{14} nm (radio waves). Application of spectroscopy for studying proteins involves using radiation within the ultraviolet (200-380 nm), visible (380-780 nm) and infrared (780-3000 nm) regions¹.

2.1.1 Ultraviolet (UV)-Visible Spectroscopy

UV-Visible spectroscopy is a form of absorbance spectroscopy that constitutes electromagnetic radiations from 200-780 nm (UV and visible). This form of spectroscopy assumes light as a stream of particles (photons) instead of waves.

Molecules exist with discrete energy levels. Change in energy levels occurs in the

molecule by absorption of energy in the form of a photon. The energy (E) of a photon is defined as

$$E = h \frac{c}{\lambda} \quad (2.1)$$

where h is the Planck's constant, c is the speed of light and λ is the wavelength of energy used.

The main principle of absorption spectroscopy is explained by Beer-Lambert's Law. It states that the fraction of incident monochromatic light that is absorbed by a molecule is proportional to the number of molecules in its path (Figure 2.2)

Figure 2.2 Schematic representation of Beer-Lambert's Law. Incident light (I_0) is absorbed by a solution and transmitted light (I) leaves the solution. Concentration (C) of the solution is calculated using the path length l and extinction coefficient of the solute in the solvent.

When light passes through a solution, the amount of light absorbed or transmitted is an exponential function of the concentration of solute in the solution². Amount of light absorbed also depends on the length of the path of radiation through the solution. Mathematically, this can be explained as

$$\log \frac{I_o}{I} = \epsilon cl \quad (2.1)$$

Where I_o is the intensity of the incident light, I is the intensity of light transmitted through the solution, c is the concentration of solute in moles/L, l is the path length of the sample in cm and ϵ is the molar absorptivity or molar extinction coefficient in $M^{-1}cm^{-1}$. Molar extinction coefficient signifies how strongly a molecule absorbs light within a particular solvent at a specific wavelength. The logarithm of the ratio of the intensity of incident light to that of transmitted light is the absorbance (A) of the molecule. Hence, equation 2.1 can be rewritten at

$$A = \epsilon cl \quad (2.2)$$

Molar absorption coefficient plays a very important role in analyzing protein using UV-Vis spectroscopy. Proteins absorb energy between 275 and 290 nm (Near-UV range) because of the absorbance of aromatic residues Tryptophan (Trp), Tyrosine (Tyr), and Phenylalanine (Phe). Cysteine absorbs to a minor extent as well. Absorbance of aromatic residues depends on the environment and it is red-shifted when the environment changes from polar to non-polar. In the near-UV range, molar absorbance of Phe is significantly smaller than Trp and Tyr. Calculating the molar extinction coefficient from the amount of Trp, Tyr and Cys residues at a particular wavelength provide the basis of calculating protein concentration within aqueous media³.

Absorbance of a molecule linearly depends on the concentration of the solute, making this technique suitable for quantitative measurements. The strength of absorbance not only depends on the chemical structure of the solute, but also on the environment

around the solute. Hence, this technique can be effectively used to measure protein concentrations and quantify protein aggregation. Protein aggregation can be quantified by measuring the absorbance of aggregates within the visible range at 350 or 600 nm (depending on the hydrodynamic radius of the aggregates). This is based on the increase in turbidity and absorbance of protein solution with increase in aggregated species.

2.1.2 Fluorescence Spectroscopy

Emission of light from substances occurs at their electronically excited states. This phenomenon is known as luminescence. Luminescence has two forms – fluorescence and phosphorescence, depending on the nature of excited state of the substance. In the case of fluorescence, the excited electron reaches a higher energy orbital with another electron of opposite spin in the ground state. This leads to return of the excited electron to ground state with emission of photon at a rate of approximately 10^8 s^{-1} . Fluorescence is indicated in a Jablonski diagram (Figure 2.3) as a straight line going down along the energy axis between electronic states. Fluorescence is a relatively slow process hence most likely observed between the first excited state and the ground state of a molecule. Energy of photons emitted during fluorescence is equal to the energy difference between corresponding Eigen states, however the energy of the emitted photon is less than that of the excited photons due to energy lost in internal conversion. This difference of energy between absorption and emission is known as Stokes shift.

The electronic states are depicted by S_0 , S_1 , S_2 etc where the molecules can exist in a number of vibrational energy levels depicted by V_0 , V_1 , V_2 etc. After absorbing light

in the form of a photon, molecules are excited to a higher energy level (S_1 , S_2 or higher). The molecule relaxes instantly (in most cases) and reaches a lower vibrational energy level by a process called internal conversion, which precedes emission. Hence, emission occurs from a thermally equilibrated excited state⁴.

Figure 2.3 Simplified Jablonski diagram showing the possible transitions during intrinsic fluorescence. Solid purple line shows absorption, followed by internal conversion, fluorescence (solid green line) and phosphorescence (solid red line). Adapted from Principles of Fluorescence Spectroscopy. Third Edition. Lakowicz, J.R.

Molecules that exhibit fluorescence are known as fluorophores. Intrinsic fluorophores are those that exhibit fluorescence naturally. Intrinsic fluorescence is dominant in the aromatic residues of amino acids: mainly indole group of tryptophan, followed by tyrosine and phenylalanine to a lesser extent (Figure 2.4). This is because tryptophan

has the highest absorption coefficient ($5.6 \times 10^3 \text{ M}^{-1}\text{cm}^{-1}$) followed by tyrosine ($1.4 \times 10^3 \text{ M}^{-1}\text{cm}^{-1}$) and phenylalanine ($0.2 \times 10^3 \text{ M}^{-1}\text{cm}^{-1}$). At 280 nm, only tryptophan and tyrosine absorb and hence the observed fluorescence is contributed by both amino acids. Unlike tyrosine, fluorescence of indole group of tryptophan residue may be blue-shifted (shorter wavelength) when the residue is buried within the native protein, whereas red-shift is observed in an unfolded protein. This property is used to measure the kinetics of protein unfolding and aggregation. In a completely folded native protein, intrinsic fluorescence shows an increasing trend due to buried tryptophan residues. In an unfolded protein, tryptophan residues are exposed to the aqueous environment leading to fluorescence quenching.

Figure 2.4 Structures of amino acid residues that exhibit intrinsic fluorescence. Figures generated using Chem3D Pro 12.0

Extrinsic fluorophores are mostly covalently attached to proteins or ligands for fluorescence energy transfer and anisotropic measurements. However, a class of extrinsic fluorophores are those that interact with proteins and its degradation products via non-covalent interactions, mainly electrostatic and hydrophobic. These

commercially available fluorophores (dyes) are highly sensitive to the environment. These dyes have been extensively used in protein characterization to monitor protein folding and unfolding, to detect molten globule intermediates, to determine mechanism of aggregation and to assess surface hydrophobicity. Dyes used for protein characterization are tabulated in Table 2.1 and their structures are shown in Figure 2.5. Proteins are characterized using dyes in conjunction with steady state fluorescence spectroscopy. Light is absorbed by the dye that leads to excitation and transfer of electrons from the ground state (S_0) to higher singlet excited state. Excitation is followed by loss of energy by several processes, leading to Stokes shift. These processes are vibrational relaxation, internal conversion, solvent relaxation, and intermolecular charge transfer (ICT) or twisted intermolecular charge transfer (TICT). These electronic transitions are explained using a schematic in Figure 2.6. After the electron reaches the lowest vibrational energy level of S_1 or $S_{(T)ICT}$ in case of charge transfer process, the molecule relaxes to ground state by fluorescence emission or non-radiative decay occurs.

Table 2.1. Extrinsic fluorescent dyes used for protein characterization and aggregation.

Adapted from Hawe, A. et.al. (2007) *Pharm.Res.* 25(7): 1487-1499.

Dye	Application	Stock Solution	Extinction Coefficient (M-1cm-1)	Excitation (nm)	Emission (nm)
ANS	Surface hydrophobicity, protein folding/unfolding kinetics	Ethanol, Aqueous	5000 at 350 nm in water	350-380	475-505
Bis- ANS	Surface hydrophobicity, protein folding/unfolding kinetics, conformational tightening	Ethanol, Aqueous	16790 at 350 nm in water	385-400	480-515
ThT	Aggregate morphology	Ethanol	26620 at 416 nm	450	480-490

Figure 2.5 Structure of extrinsic fluorescent dye ANS, Bis-ANS and ThT, used for protein characterization aggregation studies. Structures made using Chem 3D Pro 12.0.

These electronic transitions are accompanied by change in dipole moment of the dye. If the solvent also possesses a dipole moment, the solvent molecules reorient around the dye molecules to energetically favorable positions. This causes lowering of energy level S_1 and energy level S_0 rises, leading to increase in Stokes shift. This is known as solvent relaxation and depends on solvent polarity. Fluorescence properties of extrinsic dyes are also dependent on ICT and/or TICT. In both situations, an electron is transferred from a donor (amino group) to an acceptor (aromatic group) within the excited dye molecule. Both ICT and TICT induce charge separation causing an increase in dipole moment of the excited state $S_{(T)ICT}$. T(ICT) is favored in polar solvents and relaxation occurs by solvent relaxation or non-radiative decay. This explains the weak fluorescence of dyes in polar solvents. Another factor that alters dye fluorescence is

hydrogen bonding. Solvents that are capable of hydrogen bonding shift the emission spectra of a dye significantly.

Figure 2.6 Simplified Jablonski diagram showing the energy levels of the ground state and excited state, as well as electronic transitions in case of fluorescent dyes. Radiationless vibrational relaxation and internal conversion process (dashed arrows), solvent relaxation and TICT (dotted arrows) and fluorescence (solid arrows) are the possible outcomes when electrons absorb energy and reach excited state S_1 , S_2 from ground state S_0 . Adapted from Hawe, A. et al. (2007) *Pharm.Res.* 25(7): 1487-1499

In the 1950s chemists found that introducing various substituted aniline groups in naphthalene derivatives, resulted in molecules like 1-anilinonaphthalene-8-sulfonate (ANS) and its dimeric analog 4,4'-bis-1-anilinonaphthalene-8-sulfonate (Bis-ANS). ANS and Bis-ANS fluorescence is sensitive to solvent polarity, viscosity and temperature. Blue shift and increase in quantum yield occurs with decreasing dielectric

constant of solvent (decrease in polarity). In the excited state, both dyes undergo TICT with charge transfer from amino-aryl group to sulfonated naphthalene group. Fluorescence properties of ANS and Bis-ANS depend on their interaction with protein molecules. Both dyes bind to proteins via hydrophobic and electrostatic interactions. Ion pairing between negatively charged sulfonate groups with positively charged amino acids (lysine, histidine, and arginine) is dominant in ANS-protein interaction. These ion pairs are stabilized by van der Waals interactions. In case of Bis-ANS, hydrophobic interaction is more dominant. While ANS is used to determine the exposure of hydrophobic residues in a protein molecule during the aggregation process, Bis-ANS fluorescence also depends on conformational tightening of the protein⁵.

ANS and Bis-ANS are mainly used to monitor the mechanism of protein unfolding during aggregation and the conformation of the resultant aggregates respectively. Understanding the aggregation mechanism is accompanied by the question of aggregate morphology. Thioflavin T (ThT) is a classic dye used to understand whether protein aggregates are fibrillar structures or disordered amorphous forms. The dye binds to both aggregate forms with significant differences in the fluorescence intensity. When ThT binds to fibrils, fluorescence intensity is almost 10 times when compared to amorphous aggregate binding. Quantum yield of ThT is affected by solvent viscosity and rigidity of the microenvironment. The dye is almost non-fluorescent in solvents of low viscosity. This has been explained by the formation of a non-fluorescent TICT state accompanied by a change in the angle between benzothiazole and the benzene ring from 37° to 90°⁵.

ThT was originally used as a histochemical stain for amyloid-like deposits in tissues. Later the application was extended for quantification of amyloid fibrils *in vitro*. Amyloid fibrils are highly ordered protein aggregates that are rich in β -sheet structures. In the presence of these filamentous structures, ThT exhibits absorption peak at 450 nm and emits at 480 nm. The fluorescence is due to the formation of micelles binding to grooves of the fibrils. These grooves have a cavity size of less than 8 Å. Hence, amorphous aggregates with a similar cavity size also lead to fluorescence due to dye binding⁶. In bound form, ThT remains in flat conformation leading to increase in quantum yield due to inhibition of $S_{(T)ICT}$ state.

2.1.3 Fourier-Transform Infrared Spectroscopy

Infrared spectroscopy is a classical method for secondary structure determination of proteins due to its sensitivity to the chemical composition and architecture of molecules. Absorption of energy in the infrared region excites the vibrational transitions in a molecule. The vibrational frequency and probability of absorption depend on the polarity and strength of vibrating bonds, which is, related to the inter- and intra-molecular effects. The position of absorption band in infrared spectroscopy is determined by vibration masses, type of bonds involved, and position of electron accepting and electron donating groups. The strength of absorption increases with increase in polarity of the vibrating bonds.

In 1950, Elliot *et al.*⁷ demonstrated that there is an empirical relationship between the vibrational frequencies of amide I and amide II absorptions of a protein and its secondary structure^{8, 9}. Studies with proteins show that predominantly α -helical

proteins exhibit amide I and amide II spectral ranges of 1652 to 1657 cm^{-1} and 1545 to 1551 cm^{-1} respectively, in aqueous medium. Proteins with β -sheet structures absorb between 1628-1635 cm^{-1} and 1521-1525 cm^{-1} . Secondary structures such as turns, loops and random structures also absorb at specific wavenumbers. A list of wavenumbers and assignments to secondary structures in water is shown in Table 2.2.

Miyazawa and co-workers later described the amide modes in the 1960s. It was elucidated that absorption in the amide I range was 80% due to C=O stretching in the amide group. The other 20% could be attributed to C-N stretching vibration. The amide II absorption arises from N-H bending and C-H stretching vibrations. The frequency of the amide I and II absorptions also depends on the strength of hydrogen bonds involved in amide C=O and N-H groups. Each amide group is associated with distinct secondary structures in a protein, due to its association with characteristic hydrogen bonding pattern between C=O and N-H groups. This forms the basis of distinct absorption bands for each secondary structure of proteins and is the main principle behind using infrared spectroscopy for secondary structure determination^{10, 11}.

Table 2.2. Deconvoluted amide I band frequencies and assignments to secondary structure for protein in water. Adapted from Kong, J. *et.al. Acta. Biochim.et. Biophys.Sin.* (2007), 39(8): 549-559

Mean frequencies in H ₂ O	Secondary Structure Assignment
1624±1.0	β-sheet
1627±2.0	β-sheet
1633±2.0	β-sheet
1638±2.0	β-sheet
1642±1.0	β-sheet
1648±2.0	Random
1656±2.0	α-helix
1663±3.0	3 ₁₀ Helix
1667±1.0	β-turn
1675±1.0	β-turn
1680±2.0	β-turn
1685±2.0	β-turn
1691±2.0	β-sheet
1696±2.0	β-sheet

Figure 2.7 Scheme of a Fourier transform infrared spectrometer. Adapted from Barth, A. *Biochim.et.Biophys.Acta* (2007), 1767: 1073-1101

Modern infrared spectrometers are Fourier-transform infrared spectrometers (FTIR). The detector signal of the spectrometer is related by a Fourier transformation to the measured spectrum. The scheme of an FTIR spectrometer is shown in Figure 2.7. Quantification of protein secondary structure using FTIR is based on the assumption that any protein can be considered as a linear sum of individual secondary structure elements. The percentage of these elements is related to the spectral intensity. The simplicity of using FTIR to determine secondary structure of a protein is made complex by side-chain absorption. The contribution of side-chain in a typical globular protein is

roughly between 10-30% of the overall absorption. This is superimposed on the peptide absorption and might lead to changes in the amide I spectral band. Amino acid residues like lysine, arginine, histidine, asparagine, and glutamine absorb intensely in the amide I region. Hence it is important to quantify the amino acids in side-chains in order to refine the secondary structure determination of proteins⁸.

2.2 Microscopy

Proteins aggregate due to aberrant misfolding. Although the underlying mechanism of protein aggregation is similar, they form diverse aggregate forms. The pathophysiology of protein aggregation diseases is also dependent on the morphology of these aggregates that may influence their toxicity. The structure and morphology of these aggregates have been revealed by microscopy techniques such as scanning electron microscopy (SEM), transmission electron microscopy (TEM), and atomic force microscopy (AFM).

2.2.1 Scanning Electron Microscopy

Scanning Electron Microscopy (SEM) bombards a beam of high-energy electrons on a sample surface. The electron beam interacts with the solid sample and emit electrons from the sample. The type of electron emission is dependent on the interaction type and the depth of penetration of the incident electrons, as shown in Figure 2.8. As shown in the figure, emission of secondary and backscattered electrons is captured by the SEM detector. Secondary electrons are produced when an incident electron excites an electron in the sample and loses most of its energy in the process. The excited electron

moves towards the surface of the sample undergoing elastic and inelastic collisions until it reaches the surface, where it can escape if it still has sufficient energy. Production of secondary electrons is very topography related and is used to understand sample morphology. Backscattered electrons consist of high-energy electrons originating in the electron beam that are reflected or back-scattered out of the specimen interaction volume. The production of backscattered electrons varies directly with the specimen's atomic number. These differing production rates cause higher atomic number elements to appear brighter than lower atomic number elements, revealing the composition of a specimen¹².

Figure 2.8 Interactions of incident electron beam with solid sample. Incident electron beam leads to emission of electrons from the sample depending on the interaction. Adapted from Scanning Electron Microscopy and X-ray Microanalysis. Third Edition Goldstein, J., Newbury, D.E., Joy, D.C., Lyman, C.E., Echlin, P., Lifshin, E., Sawyer, L., Michael, J.R.

The working of the microscope involve four basic steps (i) Stream of electrons is formed by the electron gun in vacuum, (ii) Electrons are accelerated towards the samples (with a positive electrical potential), (iii) Electrons are confined and focused using metal apertures and electromagnetic lenses into a thin monochromatic beam, and (iv) The sample is irradiated by the beam and the interactions are detected and transformed into an image. The converging of the electron beam after passing through various electromagnetic lenses is shown in Figure 2.9.

While imaging biomolecules such a proteins using SEM, it is important that the sample does not degrade due to the heating effects associated with the high-energy electron bombardment. This necessitates the use of a cold field emission (FE) electron gun. The FE electron gun (cathode) is usually a wire of single-crystal tungsten fashioned into a sharp point and spot welded to a tungsten hairpin. The significance of the small tip radius, about 100 nm or less, is that an electric field can be concentrated to an extreme level. If the tip is held at negative 3-5 kV relative to the anode, the applied electric field at the tip is so strong that the potential barrier for electrons becomes narrow in width. This narrow barrier allows electrons to “tunnel” directly through the barrier and leave the cathode without requiring any thermal energy to lift them over the work function barrier.

When nonconductive specimens such as biomolecules are directly illuminated with an electron beam, the electrons collect locally (specimen charge-up), thus preventing normal emission of secondary electrons. This charge-up causes some unusual phenomena such as abnormal contrast and image deformation and shift. Charging of

specimen is avoided by coating the sample with a conductive material such as platinum. The average coating thickness varies from 1 to 20 nm, depending on the stability of the sample.

Both SEM and TEM are electron microscopy techniques in which electrons are bombarded on the sample to obtain morphological information. While TEM provides information in the nanometer range (100-200 nm), its use can be limited to more stable fibrils that do not disintegrate due to the high voltage used (100-200 kV). On the other hand, FESEM can be used to image relatively unstable biomolecules since it uses lower acceleration voltage (1-10 kV).

Figure 2.9 Schematic of the working principle of scanning electron microscope. Adapted from Scanning Electron Microscopy and X-ray Microanalysis. Third Edition Goldstein, J., Newbury, D.E., Joy, D.C., Lyman, C.E., Echlin, P., Lifshin, E., Sawyer, L., Michael, J.R.

2.2.2 Atomic Force Microscopy

Atomic Force Microscopy (AFM) was invented by Binnig *et.al.* in 1986 and has been used since then to study molecular interactions, mechanical forces and surface topography¹¹. It uses a force-sensing cantilever and has a very sharp tip that scans over the sample surface. As the distance between tip and surface decreases, the close-range attractive force between the surface and the tip causes the cantilever to deflect towards the sample. However, as the distance decreases significantly and the cantilever tip comes in contact with the sample surface, repulsive forces become dominant, causing the cantilever to deflect away from the sample. A typical schematic of AFM is shown in Figure 2.10.

The tip of the cantilever scans through the sample surface and the tip-surface interaction is monitored by the deflection of the laser beam on the cantilever. This deflection is measured by a position sensitive quadrant photo diode and information is transferred to the computer. The tip rasters along x and y direction and the force between the surface and tip along z direction is measured as a function of x and y¹³.

Figure 2.10 Schematic drawing of AFM apparatus. The set up consists of a sharp tip integrated into the apex of a micro-fabricated cantilever spring, a solid substrate surface (piezo scanner) to support the sample, an optical beam apparatus to detect the cantilever's deflection, and an electronic system to record and analyze data. Adapted from Barattin R. *et al.* (2008) *Chem. Comm.* 1513-1532

AFM uses two modes to investigate the sample surface – contact mode and tapping mode. In contact mode, the tip remains in contact with the sample and is very sensitive even to the slightest of changes in the sample topography. Tapping mode is a dynamic mode where the tip oscillates close to its resonance frequency and scans the sample surface. The contact between the tip and surface during tapping results in the reduction of the cantilever vibration amplitude. In both contact and tapping mode, the deflections caused by the tip-sample interaction are measured by laser deflection¹³.

New generation AFM allows sample imaging under ambient conditions as well as within aqueous solutions at physiological conditions. This has greatly improved its usage for imaging biomolecules. Either contact or tapping mode can be used for biomolecular imaging with pros and cons for both. Contact mode provides higher resolution images, although the sample needs to be immobilized on a surface in order to avoid deformation or damage. Immobilization can be a tedious process thereby limiting the use of contact mode. Tapping mode reduces sample deformation and damage, thus providing better accuracy^{14, 15}.

2.3 Cell Based Assays

Cell based assays are mainly used to screen drug compounds by quantifying their effect on cell viability and proliferation. These assays are widely used to monitor organelle functioning and to measure receptor binding and signal transduction. In order to understand the mechanism of protein aggregation and the toxicity of aggregates with diverse morphology, it is important to study the effect of these aggregates on cells.

A variety of cell based assays are available that measure an aspect of metabolism or enzymatic activity of viable cells. Dead cells are unable to carry out cellular function, and hence can be differentiated from live cells. These assays require incubating viable cells with a reagent that converts the substrate to a colored fluorescent product. The product is detected using a plate reader. The absorbance signal of the product at 490 nm is roughly proportional to the amount of viable cells present¹⁴.

Tetrazolium salt based reduction assays are most commonly used. Until recently, the most commonly used compound was (3-(4,5-dimethylthiazol-2-yl)-2,5-

diphenyltetrazolium bromide) (MTT). It is a positively charged compound that penetrates into viable cells. MTT was later developed into (3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium (MTS), which is a negatively charged compound that does not readily penetrate into cells. The mechanism of identifying viable cells using MTS reagent is shown in Figure 2.11. It involves an intermediate electron acceptor that can transfer electrons from the plasma membrane or cytoplasm, in order to facilitate the reduction of tetrazolium salt and formation of colored formazan product via mitochondrial succinate dehydrogenase enzyme^{16, 17}.

MTS reagent has an advantage over MTT because it is directly soluble in cell culture medium. This eliminates a liquid handling step of adding a second reagent to solubilize the formazan precipitates. MTS reagent uses an intermediate electron acceptor such as phenazine methyl sulfate (PMS) or phenazine ethyl sulfate (PES) that can penetrate viable cells. These get reduced in the cytoplasm or cell surface and exit the cells where they are converted into soluble formazan products.

Figure 2.11 Reaction mechanism of MTS Assay. Intermediate electron acceptor phenazine ethyl sulfate (PES) transfers electron from NADH in the cytoplasm to reduce MTS in the culture medium into an aqueous soluble formazan.

2.4 Polyacrylamide Gel Electrophoresis

Polyacrylamide gel electrophoresis (PAGE) is a versatile technique to separate macromolecules like proteins based on their molecular weight, conformation and net charge. The electrophoretic mobility of a protein is directly proportional to voltage and charge on the molecule and inversely proportional to the molecular weight, hydrodynamic radius, shape and size. A typical gel-electrophoresis set up is shown in

figure 2.12. Polyacrylamide gel is formed by the polymerization of acrylamide crosslinked by N,N'-methylene-bis-acrylamide (BIS). The polymerization reaction is initiated by ammonium persulfate (APS) and catalyst -N,N,N',N'-tetramethylene diamine (TEMED). The concentration of acrylamide and BIS can be varied to obtain a median pore size between 0.5 and 3 nm. Additionally, anionic detergent sodium dodecyl sulfate (SDS) is also added. It denatures and binds to the protein and provides a uniform net negative charge to the protein. This allows separating proteins based only on their molecular weight and is known as SDS-PAGE. This type of electrophoresis also involves denaturing the protein by heating it in a buffer containing thiol-modifying agent 2-mercaptoethanol (β -ME). β -ME reduces the disulfide bonds in the protein to free sulfhydryl groups and SDS disrupts all the inter- and intra-molecular interactions.

Figure 2.12 A typical polyacrylamide gel electrophoresis apparatus set up. Protein samples are loaded in the wells and they migrate from cathode (-) to anode (+) on application of electric field¹. From: Lehninger Principles of Biochemistry 6E, by David L. Nelson and Michael M. Cox, Copyright 2013, by W.H. Freeman and Company. Used by permission of the publisher.

A modification of SDS PAGE is a non-reducing SDS PAGE where already reduced disulfide bonds are prevented from re-forming by blocking the free sulfhydryls by iodoacetamide. After blocking, samples are mixed with buffer containing SDS sample buffer (without β -mercaptoethanol). This is important the protein already has reduced disulfide bonds and maintaining them is key for the investigation.

In an alternate form of PAGE, SDS is not added and the protein is resolved on the gel based on its molecular weight as well hydrodynamic radius. This is known as Native PAGE since the protein samples are prepared in a non-reducing, non-denaturing buffer that helps in maintaining the secondary structure and shape of the protein.

Typical gels used for protein separation are discontinuous gels. These are constructed with two different acrylamide gels on top of one another. The upper gel (stacking gel) is a loose gel with 4-5% acrylamide content. The lower gel is resolving gel with a higher acrylamide content (8%-20%). Stacking and resolving gel are maintained at different pH and this discontinuity leads to a current discontinuity as voltage is applied. The combination of stacking and resolving gels increase the resolution of protein separation by concentrating the sample to an ultrathin zone at the boundary of stacking and resolving gel. The protein is mixed with running buffer that contains glycerol. High density of glycerol ensures that the samples remain within the “well” in the gel. The top of the apparatus forms the cathode whereas bottom forms the anode. When voltage is applied, proteins migrate downward to the positively charged electrode due to the net negative charge on the protein (in case of SDS PAGE). In case of native PAGE, (protein’s net charge depends on content of acidic and basic amino acid) most proteins

that have isoelectric point (pI) in the 3-8 range migrate towards the positive electrode. If the protein of interest has a pI greater than 8, then native PAGE can be effectively performed by reversing the electrode polarity. The stacking gel has low ionic strength that increases the electrical resistance allowing the protein to migrate faster in the running gel^{18, 19}.

In all gels, a tracking dye called bromophenol blue (BPB) is used to determine the electrophoresis time where BPB serves as a visual indicator. BPB is a small molecule that migrates faster than the proteins and can be used effectively without losing proteins in the process.

2.5 Mass Assisted Laser Desorption/Ionization (Time-of-Flight) Mass Spectrometry

The main principle of mass spectrometry (MS) is to produce, separate and detect gas-phase ions. Sample to be analyzed by this method needs to ionize to produce gas-phase ions. Traditionally, samples were ionized using electron impact or chemical ionization methods. However, these methods were not suitable for handling biomolecules since they fragment and degrade the molecules significantly during the ionization process. This called for development of 'soft' ionization techniques such as Mass Assisted Laser Desorption/Ionization (Time-of-Flight) Mass Spectrometry (MALDI-MS) for analysis of biomolecules²⁰ as shown in figure 2.13

Figure 2.13 Typical set up of MALDI-TOF Mass Spectrometer. Analyte (biomolecules) is mixed with matrix and the co-crystals are allowed to form on the target and loaded into a vacuum chamber. Analyte molecules are desorbed by high energy radiation and accelerated through electric field causing the analyte ions to fly to the detector where the mass to charge ratio is obtained as the mass spectrum output. Adapted from Kim,S. *et al.* (2003) *Nat. Rev. Genetics* 4, 1001-1008.

MALDI involves the use of a small crystalline organic compound (matrix) in which the sample of interest is embedded. The mix of sample and matrix is deposited on a conductive sample support (target). The matrix and sample co-crystals are irradiated with a nanosecond UV laser beam with a wavelength of 226 or 337 nm. This contributes

to an energy range between 1×10^7 and 5×10^7 W/cm². This energy decomposes the co-crystals and creates a particle cloud from which ions are extracted by an electric field. The ions in gas-phase are accelerated through the electric field followed by a drift in a field-free path and finally reaches the detector. This desorption process is mainly understood as the conversion of laser energy to vibrational collision energy of the crystal molecules, leading to disintegration of the crystal molecules. A variety of mass analyzers are available to study biomolecules such as proteins. These range from hybrid-quadrupole time-of-flight (Q-Q-ToF) and tandem time-of-flight (ToF-ToF). MALDI provides information in the form of mass to charge ratios (m/z) that is calculated by measuring the typical time-of-flight (TOF) through a tube of specified length in vacuum²⁰⁻²². TOF is longer for larger molecules provided their initial energies are identical. MALDI predominantly generates single charged, non-fragmented ions, hence making data processing less complicated.

References

- (1) Thermo Spectronic (2001) *Basic UV-Vis Theory, Concepts and Applications*.
Consulta: 16 de mayo de, 8-9
- (2) Jaffe, H.H.; Miller, A.L. The Fates of Electronic Excitation Energy. *J. Chem. Educ.* **1966**, *43* (9), 469-473.
- (3) Schmid, F.X.; Biological Macromolecules: UV-visible Spectrophotometry. *Encyclopedia of Life Sciences*, 2001.
- (4) Lakowicz, J., *Principles of Fluorescence Spectroscopy*. Third Edition ed.; Springer: 2010.
- (5) Hawe, A.; Sutter, M.; Jiskoot, W. Extrinsic Fluorescent Dyes as Tools for Protein Characterization. *Pharm. Res.* **2008**, *25* (7), 1487-1499.
- (6) Yang, M.; Dutta, C.; Tiwari, A. Disulfide-Bond Scrambling Promotes Amorphous Aggregates in Lysozyme and Bovine Serum Albumin. *J. Phys. Chem. B* **2015**, *119* (10), 3969-81.
- (7) Elliott, E.; Ambrose, E. J.; Robinson, C. Chain Configurations in Nated and Denatured Insulin: Evidence from Infra-Red Spectra. *Nat. Rev.* **1950**, *166* (194).
- (8) Jackson, M.; Mantsch, H. H. The Use and Misuse of FTIR Spectroscopy in the Determination of Protein Structure. *Crit. Rev. Biochem. Mol. Biol.* **1995**, *30* (2), 95-120.
- (9) Surewicz, W.K.; Mantsch, H. H.; Chapman, D. *Biochemistry* **1993**, *32* (2), 389-394.
- (10) Kong, J.; Yu, S. Fourier Transform Infrared Spectroscopic Analysis of Protein Secondary Structures *Acta Biochim. Biophys. Sin.* **2007**, *39* (8), 549-559.

- (11) Barth, A. Infrared Spectroscopy of Proteins. *Biochim. Biophys. Acta* **2007**, 1767 (9), 1073-1101.
- (12) Goldstein, J.; Newbury, D.; Joy, D. C.; Charles E. Lyman, Patrick Echlin, Eric Lifshin, Linda Sawyer, J.R. Michael, *Scanning Electron Microscopy and X-Ray Microanalysis*. Third Edition ed.; Kluwer Academic/Plenum Publishers: 2009.
- (13) Binnig, G.; Quate, C. F.; Gerber, Ch. Atomic Force Microscope. *Phys. Rev. Lett.* **1986**, 56 (9), 930-934.
- (14) Giessibl, F. J. Advances in Atomic Force Microscopy. *Rev. Mod. Phys.* **2003**, 75, 949-983.
- (15) Barattin, R.; Voyer, N. Chemical Modifications of AFM Tips for the Study of Molecular Recognition Events. *Chem. Comm.* **2008**, 1513-1532.
- (16) Riss, T. L.; Niles, A. L.; Benink, H. E.; Worzella, T. J.; Minor, L.; Storts, D.; Reid, Y. Cell Viability Assays. In *Assay Guidance Manual*, Coussens, N. P., Ed. Eli Lilly & Company and the National Center for Advancing Translational Sciences: 2015.
- (17) Catalano, J. G.; Menking, D.; Hubbard, K.; Gnatt, A. L. Standardization of the 3-(4,5-dimethylthiazole-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulphophenyl)-2H-tetrazolium, inner salt (MTS) Assay for the SK-N-SH, KYSE-30, MCF-7, and HeLa Cell Lines; *Chemical Biological Center, U.S. Army Research*. 2006.
- (18) Nelson, D.L.; Cox, M. *Lehninger Principles of Biochemistry*. Fifth Edition ed.; W.H. Freeman and Company: New York, 2008.
- (19) Chrambach, A.; Rodbard, D. Polyacrylamide Gel Electrophoresis. *Science* **1971**, 172 (3982), 440-451

- (20) Jurinke, C.; Oeth, P.; van den Boom, D. MALDI-TOF Mass Spectrometry: A Versatile Tool for High-Performance DNA Analysis. *Mol. Biotechnol.* **2004**, *26*, 147-163.
- (21) Kim, S.; Ruparel, H. D.; Gilliam, T. C.; Ju, J. Digital Genotyping using Molecular Affinity and Mass Spectrometry. *Nat. Rev. Genet.* **2003**, *4*, 1001-1008.
- (22) Domon, B.; Aebersold, R. Mass Spectrometry and Protein Analysis. *Science* **2006**, *312*, 212-217.

Chapter 3. Preformed Seeds Modulate Native Insulin Aggregation Kinetics¹

Colina Dutta¹, Mu Yang¹, Fei Long², Reza Shahbazian-Yassar^{2,3} and Ashutosh Tiwari^{1*}

¹Department of Chemistry, Michigan Technological University, Houghton, MI 49931

²Department of Mechanical Engineering, Michigan Technological University, Houghton, MI 49931

³Department of Mechanical and Industrial Engineering, University of Illinois at Chicago, Chicago, IL 60607

***Correspondence:** Ashutosh Tiwari, Department of Chemistry, Michigan Technological University, Houghton, MI, USA,

Tel:(906) 487-1840; Fax: (906) 487-2061; Email: tiwari@mtu.edu

¹The material contained in this chapter was previously published in The Journal of Physical Chemistry B. Reprinted (adapted) with permission from “Dutta, C.; Yang, Mu; Long, F.; Yassar, R.S.; Tiwari, A. Preformed Seeds Modulate Native Insulin Aggregation Kinetics (2015) *J. Phys. Chem B.* **119**, 15089-99. Copyright (2015) American Chemical Society.

Abstract

Insulin aggregates under storage conditions via disulfide interchange reaction. It is also known to form aggregates at the site of repeated injections in diabetes patients, leading to injection amyloidosis. This has fueled research in pharmaceutical and biotechnology industry as well as in academia to understand factors that modulate insulin stability and aggregation. The main aim of this study is to understand the factors that modulate aggregation propensity of insulin under storage conditions. We explored the aggregation kinetics of insulin at pH 7.2 and 37 °C in the presence of disulfide-reducing agent dithiothreitol (DTT), using spectroscopy (UV-visible, fluorescence, and Fourier transform infrared spectroscopy) and microscopy (scanning electron microscopy, atomic force microscopy) techniques. We prepared insulin ‘seeds’ by incubating disulfide-reduced insulin at pH 7.2 and 37 °C for varying lengths of time (10 min to 12 h). These seeds were added to the native protein and nucleation-dependent aggregation kinetics was measured. Aggregation kinetics was fastest in the presence of 10 min seeds suggesting they were *nascent*. Interestingly, *intermediate* seeds (30 min to 4 h incubation) resulted in formation of transient fibrils in 4 h that converted to amorphous aggregates upon longer incubation of 24 h. Overall, the results show that insulin under disulfide reducing conditions at pH and temperature close to physiological favors amorphous aggregate formation and seed ‘maturity’ plays an important role in nucleation dependent aggregation kinetics.

Introduction

More than 29 million people have diabetes in USA alone and this number is on the rise with more than one million new cases of diabetes diagnosed every year¹. Nearly 14% of these patients take insulin, a globular hormone that has been used as treatment of insulin-dependent diabetes for more than seven decades¹. Studies in the last two decades show that insulin plays an important role in cognition and memory. Insulin is also known to regulate metabolism of beta-amyloid peptide and phosphorylation of tau-protein that is robustly linked to Alzheimer's disease². Even though, insulin has been used for diabetes for several decades, its high cost due to poor stability and ease of forming aggregates when stored is a major challenge for the pharmaceutical industry. Insulin undergoes self-association by disulfide interchange reaction and aggregates under storage conditions^{3,4}. Previous works show that insulin tends to precipitate when reacted with a disulfide interchange enzyme in presence of reducing agent. This aggregation was attributed to scrambling of disulfide bonds^{5,6}. In addition, insulin can also aggregate at the site of injection in the subcutaneous tissue which has a pH of ~ 7.2 ^{7,8}. Even though, several insulin aggregation studies have been carried out at neutral pH either at high temperature or in the presence of chaotropic agents (e.g. urea or guanidine hydrochloride)^{9,10}, very few studies have explored insulin aggregation under conditions close to physiological¹¹⁻¹³. Therefore, it is important to understand the factors that govern insulin stability and aggregation at neutral pH at which insulin is generally stored. Understanding the mechanism of insulin denaturation and aggregation during storage conditions will help in formulating effective solutions to increase the shelf life of the protein and prevent injection amyloidosis.

Insulin consists of two polypeptide chains: a 21-residue A-chain and a 30-residue B-chain. It has 3 disulfide bridges that stabilize the protein structure; two inter-chain disulfide bonds (A7-B7, A20-B19) and one intra-A-chain (A6-A11) disulfide bond^{14,15}. At acidic conditions, insulin is predominantly a monomer and the hexamer prevails at neutral pH¹⁶ (Figure 3.1). Insulin fibrils have been reported in case of human insulinoma as well as in pancreatic tumors¹⁷. Classically, the aggregation propensity of this protein has been studied at acidic pH and elevated temperatures¹⁸⁻²¹. Under these conditions, it has been shown to form fibrils that exhibit characteristic amyloid properties that are assumed to start from partially unfolded monomeric intermediates^{22, 23}. Changes in pH, ionic strength, anions, and stabilizers affect the lag time and fibril growth formation in insulin suggesting the importance of hydrophobic and electrostatic interactions during initial aggregation process¹³. Studies involving solvation effects on insulin fibrillation exhibited aggregation in the presence of ethanol and trifluoroethanol²². All these studies point towards a general hypothesis that any external perturbation contributes towards insulin aggregation^{18, 20, 23-27}.

Figure 3.1 Ribbon structure of insulin hexamer at neutral pH. Firebrick red sticks are used to denote the disulfide bonds (-S-S-) between A7-B7, A20-B19 and A6-A11 positions in each monomer. The backbone is shown in gray color. The structures were generated using PyMOL 1.3 and PDB file 1AI0²⁹.

Liquid insulin formulations prepared for diabetics, come with an expiration time span of 18-24 months. This long-term storage of insulin can lead to instability of the protein resulting in formation of unwanted products and possible complications such as injection amyloidosis when administered to patients¹⁴. One of the major causes of loss of native protein structure and aggregation is loss of critical S-S bonds or formation of

scrambled disulfide bonds that destabilize the protein structure^{26, 27}. Previous studies on insulin with reduced and alkylated disulfide bonds showed increased aggregation rate when compared to insulin with intact disulfide bonds^{27, 28}. This underscores the importance of disulfide bonds in the overall stability and activity of the protein. Therefore, it is important to study how reducing environment near physiological pH can affect protein stability and hence modulate its aggregation over a period of time. In this study we show reduction of disulfide bonds of insulin leads to rapid aggregation within 10 min. Interestingly, we did not observe any fibrillar aggregates under these conditions. Instead, we observed amorphous insulin aggregates that stayed amorphous even after 6 months of incubation time. AFM images of native insulin in the presence of 30 min to 1 h insulin seeds show fibrillar aggregates after 4 h of incubation that converts to amorphous aggregates upon incubation for 24 h. Native insulin incubated with 10 min seeds form amorphous aggregates only. This study provides a mechanistic insight into insulin instability and how seeds from different stages of protein aggregation can affect native protein's aggregation kinetics and the nature of aggregates. This may help us better understand nucleation dependent aggregation process at the molecular level for improperly stored insulin and injection amyloidosis in diabetes patients.

Materials and Methods

Unless otherwise indicated, all materials were used as supplied by the manufacturer without any further purification. Human recombinant insulin, dithiothreitol (DTT), thioflavin T (ThT), 8-anilino-1-naphthalenesulfonic acid (ANS) dye, 4,4'-Dianilino-1,1'-binaphthyl-5,5'-disulfonic acid (bis-ANS) dye were purchased from Sigma.

Preparation of insulin samples - Insulin was dissolved in 0.025 M HCl at a concentration of 10 mg/mL. The protein was allowed to equilibrate overnight at 4 °C with mild agitation. To the insulin solution 0.025 M NaOH was added to neutralize the HCl. Then, insulin aliquots were made to the final concentration of 1 mg/mL (unless otherwise indicated) in 20 mM sodium phosphate buffer at pH 7.2 having 150 mM NaCl. All samples were prepared either in presence or absence of 10 mM dithiothreitol (DTT). The samples were prepared on ice followed by incubation at 37 °C for the indicated time periods (see figures for details). Blanks (contained everything except protein) were similarly prepared and incubated as samples and were used for background subtraction.

Seeding activity of disulfide-reduced insulin aggregates – Insulin seeds were prepared by incubating 1 mg/mL (172 µM) insulin in presence of 10 mM DTT in phosphate buffer, for 10 min, 30 min to 4 h (30 min intervals), and 4 h to 12 h (1 h intervals) at 37°C. Post incubation, seeds were washed with distilled water using Millipore Amicon® Ultra centrifugal filters (3000 Da cutoff) and were added to freshly prepared native insulin sample (all disulfide bonds intact) at a concentration of 5%, 15% and 45% v/v. Negative controls were prepared using identical conditions as samples but in

the absence of seeds. All samples were prepared on ice and then incubated at 37 °C for the indicated time. In addition to incubating insulin at 1 mg/mL (172 μ M), samples were also diluted to a final protein concentration of 90 μ M and 20 μ M in phosphate buffer to study the effect of dilution on aggregation kinetics. All measurements were done in triplicates at 25 °C.

Matrix Assisted Laser Desorption/Ionization Time of Flight Mass Spectrometry (MALDI-TOF MS) – Microflex LRF MALDI-TOF mass spectrometer (Bruker) equipped with microScout Ion Source was used to characterize disulfide states in the insulin samples. Insulin was incubated with 10 mM DTT for 0, 15 min, and 4 h at 37 °C. Fresh native insulin was used as control. Post incubation, samples (and control) were incubated with freshly prepared 100 mM iodoacetamide (IAA) for 2 h at room temperature to block the free sulfhydryl groups. Low molecular weight contaminants were removed using C₁₈ ZipTip[®]. Samples were prepared on the target using the direct droplet method. In this method, a saturated matrix solution of α -Cyano-4-Hydroxycinnamic acid was mixed with solvent (0.1% trifluoroacetic acid (TFA) with 33% acetonitrile (ACN)) having sample in 1:1 ratio v/v (matrix:solvent). From this mixture 1 μ L of sample was aliquoted and dried at room temperature on the target before analyzing.

UV-visible absorbance spectroscopy- Absorbance measurements were performed using a Perkin Elmer Lambda 35 UV-visible Spectrophotometer. After completion of incubation, absorbance of samples was measured at 600 nm in order to assess turbidity

of insulin samples due to aggregation. All measurements were done at room temperature and in triplicates.

Intrinsic and extrinsic fluorescence- Fluorescence measurements were carried out using a Horiba Jobin Yvon spectrofluorometer (Fluoromax-4). Excitation and emission slit widths were set at 2 nm, each. For intrinsic fluorescence, emission spectra were acquired from 285 to 450 nm with excitation at 280 nm. For extrinsic fluorescence, emission spectra were acquired in presence of 10 μ M ThT, 10 μ M ANS, or 1 μ M bis-ANS. The concentrations of dye stocks were determined using their extinction coefficients: ANS $\epsilon_{350\text{ nm}} = 5000\text{ M}^{-1}\text{ cm}^{-1}$, bis-ANS $\epsilon_{385\text{ nm}} = 16790\text{ M}^{-1}\text{ cm}^{-1}$, and ThT $\epsilon_{416\text{ nm}} = 26,620\text{ M}^{-1}\text{ cm}^{-1}$. Protein samples were equilibrated with respective fluorophores on ice, in dark for 30 min before acquiring the spectra. For ThT fluorescence, emission spectra were collected from 460 to 650 nm with excitation at 450 nm. Fluorescence spectra from 400 to 700 nm were collected with excitation at 380 nm for ANS and at 360 nm for bis-ANS. All measurements were done in triplicates.

Field Emission Scanning Electron Microscopy (FESEM) – Electron microscope images were collected using Hitachi S-4700 field-emission scanning electron microscope (FESEM). Samples incubated for indicated times were diluted with double-distilled water and aliquoted in Millipore Amicon[®] Ultra centrifugal filters (3000 Da cut off). The diluted samples were centrifuged at 7000 x g at 4 °C (three repeats after dilution with water) to remove low molecular weight impurities (salt, buffer etc). The washed samples were aliquoted on SEM stubs and air-dried. The samples were

platinum coated using a sputter coater and imaged using an acceleration voltage of 10 kV.

Atomic Force Microscopy (AFM) – Aggregates of insulin were analyzed by aliquoting 7 μ L of sample on a freshly cleaved mica surface at 25 °C. AFM imaging was performed on Dimension ICON AFM system (Bruker, CA, USA) using Peakforce Tapping mode. Scan-Asyst-Air cantilevers (Bruker, CA, USA) with spring constant 0.4 N/m and resonant frequency 70 kHz were selected as recommended by the manufacturer. AFM image analysis was carried out using NanoScope Analysis software from Bruker, CA, USA.

Nano-mechanical measurements were performed using Dimension ICON AFM (Bruker, USA) under ambient conditions using Peakforce Tapping mode. RTESPA cantilevers (Bruker, USA) were selected with calibrated spring constant and tip radius of 55.6 N/m and 12 nm, respectively. For nano-mechanical measurements, 1 mg/mL insulin was incubated with 45% v/v seeds (10 min, 30 min, 1 h, 4 h) for 4 hours. Samples were prepared on freshly cleaved mica similar to AFM imaging. Derjaguin-Mueller-Toporov (DMT) modulus analysis was carried out using NanoScope software from Bruker, CA, USA. The Young's modulus value for each sample was calculated by averaging the moduli of each aggregate in a sample using the Nanoscope software.

Fourier Transform-Infrared Spectroscopy (FT-IR) – Secondary structure determination of insulin seeds were performed using Perkin Elmer Frontier FT-IR spectrometer. Seeds were prepared by incubating 1 mg/mL insulin with 10 mM DTT at 37 °C for 10 min, 30 min, 1 h and 4 h followed by washing as discussed earlier. A

drop of sample was placed on the crystal and allowed to dry at room temperature. Data was acquired at a resolution of 4 cm^{-1} and 128 scans were averaged per sample. Data was analyzed between 1800 and 1600 cm^{-1} after subtracting the background. Raw spectra were deconvoluted using Gaussian model to obtain the component peaks and the contribution from each peaks was calculated by integrating area under the curves using Origin 9.1.

Results

In this work, we monitored aggregation of insulin (Figure 3.1) in presence and absence of 10 mM DTT at pH 7.2 and at $37\text{ }^{\circ}\text{C}$. We also measured the aggregation kinetics of native insulin in presence of “*seeds*” prepared by incubating disulfide-reduced insulin (insulin incubated with 10 mM DTT) for different time periods (10 min to 12 h). The kinetics, morphology, and secondary structure of the aggregates were studied using a variety of biophysical and analytical methods. Susceptibility of insulin to disulfide bond reduction was investigated using MALDI-TOF (Figure 3.2). Samples incubated for 15 min and 4 h with 10 mM DTT (Figure 3.2C, 3.2D) had the highest proportion of Insulin B-chain (3.5 kDa) along with dimers of full-length protein as well as B-chain (11.6 and 7.1 kDa respectively). Aggregated forms of B-chain dimer along with A-chain were also observed at a molecular mass of 9.6 kDa . Trace amounts of A-chain (2.5 kDa) and insulin monomer were also visible. Fresh insulin in the absence or presence of reducing agent predominantly showed the existence of intact insulin (Figure 3.2A, 3.2B). However, in the presence of 10 mM DTT, the inter-chain disulfide

bonds in insulin were cleaved rapidly showing increase in B chain content (Figure 3.2C, 3.2D).

The extent of protein aggregation was monitored by measuring the turbidity (optical density (O.D.)) at 600 nm, using a UV-visible spectrophotometer (Figure 3.3). The absorbance of samples containing 10 mM DTT increased with increasing incubation time but was mostly unchanged for the control groups (proteins not treated with reducing agent). Significant increase in O.D. was observed within the first 1 h of incubation.

Figure 3.2 Mass spectrometry (MALDI-TOF) analysis showing effect of 10 mM DTT on disulfide status of insulin. Insulin was incubated with 10 mM DTT for 0 min, 15 min, and 4 h at 37 °C. Fresh insulin was used as control. Post incubation, samples (and control) were alkylated with 100 mM IAA for 2 h before running on MALDI-TOF. Fresh insulin in the presence and absence of reducing agents were similarly alkylated using IAA and were run as controls.

Figure 3.3 Protein aggregation monitored by UV absorbance. Insulin was incubated at 37° C for the indicated time periods and UV absorbance of the samples were measured at 600 nm. Error bars indicate $\pm$ SD.

Insulin misfolding, conformational change, hydrophobic exposure, and aggregation were monitored by intrinsic and extrinsic fluorescence. Protein misfolding and conformational change for disulfide-reduced insulin was investigated by measuring the change in intrinsic fluorescence (Figure 3.4). We used extrinsic fluorophores such as 8-anilino-1-naphthalenesulfonic acid (ANS), to measure misfolding and extent of hydrophobic exposure; and 4,4'-Dianilino-1,1'-binaphthyl-5,5'- disulfonic acid (bis-ANS) as a conformational tightening agent and measure of hydrophobic exposure

(Figure 3.5). Thioflavin T (ThT) was used to monitor the nature of aggregates (Figures 3.8). The morphology (Figure 3.9 to 3.11, 3.21. 3.22) and secondary structure (Figure 3.28) of the aggregates were further characterized by microscopy techniques (SEM, AFM) and FT-IR respectively.

Insulin samples incubated with 10 mM DTT show a rapid decline in fluorescence as a function of time, whereas no measurable change in fluorescence intensity was observed for the protein in the control groups (0 M DTT). The steepest drop in fluorescence intensity for DTT treated samples is observed in the first 20 min (inset in Figure 3.4). Fluorescence intensity continues to decrease rapidly up to 48 h and then slows down and the intensity stabilizes. Misfolding and hydrophobic exposure of insulin was monitored using ANS and bis-ANS (Figure 3.5). Both ANS and bis-ANS show a rapid increase in fluorescence intensity in the first 20 min and then was steady beyond 1 h of incubation time. Control samples (samples with no DTT) did not show any change in fluorescence upon addition of the dyes.

Figure 3.4 Insulin unfolding monitored by intrinsic fluorescence. Fluorescence spectra for 172 μ M insulin incubated with/without 10 mM DTT were collected from 285 to 450 nm with excitation at 280 nm. Peak emission wavelength of 306 nm was selected. Peak fluorescence intensities are shown using open symbols (without DTT) and closed symbols (with 10 mM DTT). Inset shows a plot for the first 1 h of incubation. Error bars indicate $\pm$ SD.

Figure 3.5 Protein misfolding and hydrophobic exposure monitored by ANS and bis-ANS fluorescence. Insulin samples at 172 μM concentration were incubated with 10 μM ANS or 1 μM bis-ANS for 15 min before acquiring spectra. Emission spectra were collected from 400 to 700 nm with excitation at 380 nm for (A) ANS and 360 nm for (B) bis-ANS. Average peak emission wavelengths of 467 nm and 482 nm were selected for ANS and bis-ANS, respectively. Peak fluorescence intensities are shown using open symbols (without DTT) and closed symbols (with 10 mM DTT). Inset shows a plot for the first 1 h of incubation. Error bars indicate $\pm$ SD.

Figure 3.6 ANS fluorescence peak intensity and wavelength shift as a function of time. Insulin samples at 172 μM protein concentration were incubated with 10 μM ANS, and emission spectra were acquired as detailed in figure 4. Peak fluorescence intensities are shown in black and wavelength for peak maxima are in red. Error bars indicate $\pm$ SD.

Figure 3.7 ANS fluorescence emission spectra, peak intensity wavelength, and wavelength shifts as a function of time. Insulin samples at 172 μM protein concentration were incubated with 10 mM DTT and then were diluted with phosphate buffer (20 mM, pH 7.2) to a final concentration of 20 μM and 90 μM before analysis. Protein samples were incubated with 10 μM ANS and emission spectra was acquired for (A) 20 μM and (C) 90 μM insulin as detailed in figure 4. Wavelength shifts were plotted for (B) 20 μM and (D) 90 μM protein. Fluorescence intensities are shown in black and wavelength for peak maxima are in red. For emission spectra, peak fluorescence intensities are shown using open symbols (without DTT) and closed symbols (with 10 mM DTT). Inset shows a plot for the first 1 h of incubation. Error bars indicate $\pm$ SD.

The nature and morphology of the aggregates were further characterized by ThT fluorescence, SEM and AFM imaging. ThT is known to bind to amyloid fibrils as well

as large aggregates and fluoresces^{30, 31}. Fluorescence intensity for DTT treated insulin increased rapidly in the first 30 min of incubation, followed by a gradual increase upon longer incubation (Figure 3.8). Whereas, insulin incubated similarly but in absence of DTT (control groups) did not show any change in fluorescence intensity over the length of time. SEM images of insulin samples treated with DTT and incubated at 37 °C for different time periods (15 min, 2 h, 4 h, 168 h, or 6 months) were taken (Figures 3.9, 3.10, and 3.11). Freshly incubated insulin samples with and without DTT were imaged as controls (Figure 3.10, panels A and B). All the aggregates were amorphous in nature with an average subunit diameter of 200 ± 100 nm (Figure 3.9)³². The morphology of aggregates remained unaltered even after a prolonged incubation of 6 months (Figure 3.11).

Figure 3.8 Protein aggregation monitored by ThT fluorescence. Emission spectra for 57 μ M insulin in the presence of 10 μ M ThT were acquired from 460 to 650 nm with excitation at 450 nm. Average emission peak wavelength of 486 nm was selected. Peak fluorescence intensities are shown using open symbols (without DTT) and closed symbols (with 10 mM DTT). Inset shows a plot for the first 1 h of incubation. Error bars indicate $\pm$ SD.

Figure 3.9 Scanning electron microscope images of insulin aggregates. Aggregates of insulin were imaged using SEM. Insulin samples were incubated at 37 °C with 10 mM DTT for (A) 15 min, (B) 2 h, (C) 4 h, and (D) 168 h, respectively. Scale bar for all panels is 500 nm.

Figure 3.10 Scanning electron microscope images of insulin aggregates incubated for different times. Insulin aggregates were made by incubating 172 μ M insulin in the absence (A) and presence (B-F) of 10 mM DTT at 37 $^{\circ}$ C, at pH 7.2. Insulin images are (A) 0 h (no DTT) and samples incubated with 10 mM DTT for (B) 0 h (C) 15 min (D) 2 h, (E) 4 h and (F) 168 h, respectively. Scale bars are 50 μ m, 10 μ m, 5 μ m and 500 nm from left to right.

Figure 3.11 Insulin samples incubated for 6 months show amorphous aggregates. Insulin was incubated at 37°C in presence of 10 mM DTT for 6 months before imaging the aggregates with SEM at indicated magnifications. Scale bars are shown in each panel.

The effect of electrostatic interaction on insulin aggregation was studied in the absence as well as presence (150 or 300 mM) of NaCl. In addition, insulin concentration was also varied (20, 90, or 172 μM) (Figures 3.12 to 3.17). Increasing the salt concentration up to 300 mM did not alter the overall aggregation kinetics (Figures 3.12 to 3.17). Furthermore, change in insulin concentration (20, 90, or 172 μM) did not affect the overall trend for the fluorescence.

Figure 3.12 Effect of electrostatic interactions on insulin misfolding monitored by intrinsic fluorescence over time. Intrinsic fluorescence spectra were collected from 285 to 450 nm with excitation at 280 nm. Peak emission wavelength of 306 nm was selected. Insulin samples at 172 μ M protein concentration were incubated in the absence or presence of NaCl. Peak fluorescence intensities are shown using open symbols (without DTT) and closed symbols (with 10 mM DTT). Inset shows a plot for the first 1 h of incubation. Error bars indicate $\pm$ SD.

Figure 3.13 Effect of electrostatic interactions on insulin misfolding monitored by intrinsic fluorescence over time. Intrinsic fluorescence spectra were collected from 285 to 450 nm with excitation at 280 nm. Peak emission wavelength of 306 nm was selected. Insulin samples at 172 μ M protein concentration were incubated in the absence or presence of NaCl and spectra were acquired at a final concentration of (A) 90 μ M and (B) 20 μ M. Peak fluorescence intensities are shown using open symbols (without DTT) and closed symbols (with 10 mM DTT). Inset shows a plot for the first 1 h of incubation. Error bars indicate $\pm$ SD.

Figure 3.14 Effect of electrostatic interactions on hydrophobic exposure monitored by Bis-ANS fluorescence over time. Fluorescence spectra were collected from 400 to 700 nm with excitation at 360 nm. Peak emission wavelength of 482 nm was selected. Insulin samples at 172 μ M protein concentration were incubated in the absence or presence of NaCl. Peak fluorescence intensities are shown using open symbols (without DTT) and closed symbols (with 10 mM DTT). Inset shows a plot for the first 1 h of incubation. Error bars indicate $\pm$ SD.

Figure 3.15 Effect of electrostatic interactions on hydrophobic exposure monitored by Bis-ANS fluorescence over time. Fluorescence spectra were collected from 400 to 700 nm with excitation at 360 nm. Peak emission wavelength of 482 nm was selected. Insulin samples at 172 μ M protein concentration were incubated in the absence or presence of NaCl and spectra were acquired at a final concentration of (A) 90 μ M and (B) 20 μ M. Peak fluorescence intensities are shown using open symbols (without DTT) and closed symbols (with 10 mM DTT). Inset shows a plot for the first 1 h of incubation. Error bars indicate $\pm$ SD.

Figure 3.16 Effect of electrostatic interactions on insulin aggregation monitored by ThT fluorescence over time. Fluorescence spectra were collected from 460 to 650 nm with excitation at 450 nm. Peak emission wavelength of 486 nm was selected. Insulin samples at 172 μ M protein concentration were incubated in the absence or presence of NaCl. Peak fluorescence intensities are shown using open symbols (without DTT) and closed symbols (with 10 mM DTT). Inset shows a plot for the first 1 h of incubation. Error bars indicate $\pm$ SD.

Figure 3.17 Effect of electrostatic interactions on insulin aggregation monitored by ThT fluorescence over time. Fluorescence spectra were collected from 460 to 650 nm with excitation at 450 nm. Peak emission wavelength of 486 nm was selected. Insulin samples at 172 μ M protein concentration were incubated in the absence or presence of NaCl and spectra were acquired at a final concentration of (A) 90 μ M and (B) 20 μ M. Peak fluorescence intensities are shown using open symbols (without DTT) and closed symbols (with 10 mM DTT). Inset shows a plot for the first 1 h of incubation. Error bars indicate $\pm$ SD.

Seeding experiments were carried out by incubating native insulin with *seeds* for varying lengths of time at 37 °C. *Seeds* were prepared by incubating disulfide-reduced insulin at pH 7.2 at 37 °C for 10 min, 30 min to 4 h (30 min intervals) and 5 h to 12 h (1 h intervals). Fresh insulin was incubated with 5%, 15% and 45% v/v of various *seeds* prepared. The kinetics and morphology of insulin aggregates in presence of seeds were analyzed by ThT fluorescence and AFM (Figures 3.18 to 3.27, Table 3.1). The aggregation kinetics of native insulin in the presence of *seeds* were calculated by fitting the ThT fluorescence data at 45% v/v seed concentration (Figure 3.23 to 3.25) till 4 h and the results are shown in Table 3.1. Aggregation of native insulin was fastest in the presence of 10 min seeds (Figure 3.23, Table 3.1) followed by seeds incubated for 30 min, 1, 1.5 and 4 h (Figure 3.24, Table 3.1). Seeds prepared by incubating insulin for longer time (5 to 12 h) showed a much slower aggregation kinetics (Figures 3.25, Table 3.1). Control samples (fresh insulin incubated without seeds) did not show any change over the experiment period of 7 days. Since, 10 min seeds showed fast and extreme aggregation efficiency, effect of seed concentrations at 5%, 15% and 45% v/v by ThT fluorescence on kinetics of native insulin aggregation was measured (Figure 3.27). Seed concentrations alter the aggregation kinetics with 5% seeds showing a gradual but continuous increase in fluorescence, 15% seeds showing a rapid initial increase in fluorescence followed by gradual increase in fluorescence, and 45% seeds showing a very rapid increase in fluorescence intensity that peaks by 4 h of incubation time and then stabilizes. Insulin aggregates were prepared in the presence of seeds (10 min, 30 min, 1 h and 4h) at a concentration of 45% v/v and were visualized using AFM with peak force tapping mode in air (Figures 3.21, 3.22). The elastic strength of the

aggregates along the radial direction (Young's modulus) was measured using peak force QNM mode of AFM. This intrinsic stiffness of each aggregate in a sample was calculated using Nanoscope software and the average number was used to represent the Young's moduli of aggregates generated using different timed *seeds*. The elastic strength of aggregates formed by incubating native insulin with different timed *seeds* (10 min, 30 min, 1 h, and 4 h) were comparable and ranged between 4 ± 0.6 and 6 ± 0.2 GPa. Aggregate population was maximum for protein incubated with 10 min seeds and it decreased with seed incubation time (Figure 3.21).

Secondary structure of the protein seeds (10 min, 30 min, 1 h and 4 h) were analyzed using FT-IR (Figure 3.28) in the amide I band region ($1700\text{-}1600\text{ cm}^{-1}$). The raw spectra were deconvoluted using Gaussian model to obtain the component peaks. Native insulin showed a peak at 1656 cm^{-1} whereas seeds exhibited peaks at 1656 cm^{-1} and 1625 cm^{-1} upon deconvolution. Area under the peaks was integrated to obtain the relative contribution of component peaks at 1656 cm^{-1} and 1625 cm^{-1} , respectively.

Figure 3.18 Effect of seeding (by different time seeds) on native protein aggregation monitored by ThT fluorescence. Seeding concentration used for all experiments was 5% seed v/v. Seeds were prepared by incubating 172 μ M insulin for 10, 30 min, 1 or 4 h, in presence of 10 mM DTT. The spectra were acquired at (A) 172 μ M protein concentration and also after dilution of protein to (B) 90 μ M and (C) 20 μ M concentrations. ThT fluorescence experiments were performed as detailed in Figure 3.8. Error bars indicate $\pm$ SD.

Figure 3.19 Effect of seeding (by different time seeds) on native protein aggregation monitored by ThT fluorescence. Seeding concentration used for all experiments was 15% seed v/v. Seeds were prepared by incubating 172 μ M insulin for 10, 30 min, 1 or 4 h, in presence of 10 mM DTT. The spectra were acquired at (A) 172 μ M protein concentration and also after dilution of protein to (B) 90 μ M and (C) 20 μ M concentrations. ThT fluorescence experiments were performed as detailed in Figure 3.8. Error bars indicate $\pm$ SD.

Figure 3.20 Effect of seeding (by different time seeds) on native protein aggregation monitored by ThT fluorescence. Seeding concentration used for all experiments was 45% seed v/v. Seeds were prepared by incubating 172 μ M insulin for 10, 30 min, 1 or 4 h, in presence of 10 mM DTT. The spectra were acquired at (A) 172 μ M protein concentration and also after dilution of protein to (B) 90 μ M and (C) 20 μ M concentrations. ThT fluorescence experiments were performed as detailed in Figure 3.8. Error bars indicate $\pm$ SD.

Figure 3.21 AFM Images of insulin aggregates after incubating with different timed seeds for 4 h. Insulin samples at 172 μM protein concentration were incubated for 4 h in presence of seeds prepared by incubating for (A) 10 min (*nascent* seeds) (B) 30 min (*intermediate* seeds) (C) 1 h (*intermediate* seeds) and (D) 4 h (*intermediate* seeds). For panel A, B, C, and D the scale bar is 500 nm and the zoomed-in image in panel C is at 100 nm. Arrows indicate proto-fibrillar structures.

Figure 3.22 AFM images of aggregated insulin after incubating with seeds for 24 h. Insulin samples at 172 μ M protein concentration were incubated for 24 h in presence of seeds prepared by incubating disulfide-reduced insulin at pH 7.2 and 37 $^{\circ}$ C for (A) 10 min (B) 30 min (C) 1 h and (D) 4 h. Scale bar is 500 nm.

Figure 3.23 Linear regression analysis of ThT fluorescence data. ThT fluorescence plots of native insulin aggregation in the presence of nascent *seeds* were fitted up to 4 h and kinetics of aggregation was calculated using Origin 9.1

Figure 3.24 Linear regression analysis of ThT fluorescence data. ThT fluorescence plots of native insulin aggregation in the presence of intermediate *seeds* were fitted up to 4 h and kinetics of aggregation was calculated using Origin 9.1

Figure 3.25 Linear regression analysis of ThT fluorescence data. ThT fluorescence plots of native insulin aggregation in the presence of mature *seeds* were fitted up to 4 h and kinetics of aggregation was calculated using Origin 9.1

Figure 3.26 Effect of seed incubation time on native insulin aggregation monitored by ThT fluorescence. Seeds were prepared by incubating 172 μ M insulin for 10 min, 30 min to 4 h (30 min intervals) and 4 h to 12 h (1 h intervals) in presence of 10 mM DTT. The spectra were acquired as detailed in Figure 3.8. Error bars indicate $\pm$ SD.

Table 3.1 Aggregation kinetics in the presence of seeds incubated from 10 min to 12h.

Seed Incubation Time (h)	Slope after Linear Fit ($\times 10^3$)
0.2	235.92 $\pm$ 4.86
0.5	178.18 $\pm$ 21.10
1	214.29 $\pm$ 29.52
1.5	147.62 $\pm$ 5.41
2	84.68 $\pm$ 11.65
2.5	78.19 $\pm$ 8.20
3	131.11 $\pm$ 39.12
3.5	101.87 $\pm$ 15.82
4	161.77 $\pm$ 41.66
5	114.32 $\pm$ 14.71
6	89.17 $\pm$ 22.77
7	124.85 $\pm$ 75.78
8	35.82 $\pm$ 6.06
9	33.11 $\pm$ 6.04
10	146.00 $\pm$ 42.22
11	130.25 $\pm$ 32.00
12	98.37 $\pm$ 25.46

Figure 3.27 Effect of seeding concentration of *nascent* seeds on aggregation kinetics. *Nascent* Seeds prepared by incubating insulin with 10 mM DTT for 10 min were used at a concentration of 5%, 15% and 45% v/v for seeding experiments. Protein aggregation was monitored by ThT fluorescence and was performed as detailed in Figure 3.8.

Figure 3.28 FT-IR Spectra for the amide I region of insulin aggregates. Aggregates were prepared by incubating 1 mg/mL insulin with 10 mM DTT for 10 min, 30 min, 1 h and 4 h along with spectra of native protein. (A) Raw FT-IR spectra showing that amide bands are observed near 1656 and 1625 wavenumbers. (B) Raw FT-IR spectra (solid lines) and their component peaks (dotted lines) for the seeds.

Discussion

Insulin can easily aggregate due to destabilizing influences and the nature and morphology of these aggregates are diverse^{11, 18, 19, 33-35}. Under extremes of conditions such as pH, temperature, or denaturants, insulin forms beta-sheet rich fibrils^{11, 36}. Several other aggregation studies at physiological pH and temperature in combination with external perturbations show that proteins can misfold and form a range of aggregate forms^{19, 22, 37}. Both amyloid formation and amorphous aggregates have been reported in the disease process and have been shown to be dependent on external

factors. In addition, importance of conserved disulfide bond providing stability to proteins has been explored in several studies³⁸⁻⁴⁰. It is known that insulin aggregates at storage conditions by a beta-elimination reaction (scrambled disulfide bonds) and this may lead to aggregation of insulin at the injection site of diabetes patients leading to injection amyloidosis. Therefore, to better understand the kinetics and mechanism of insulin aggregation with compromised disulfide bonds; we studied the effect of different preformed *seeds* on the aggregation kinetics of native insulin. In this work, we used DTT as disulfide reducing agent to accelerate the disulfide scrambling process.

We measured the extent of disulfide reduction using MALDI-TOF mass spectrometry (Figure 3.2). Fresh insulin sample spectra showed that it was predominantly monomer (Figure 3.2A, 3.2B). Insulin incubated with 10 mM DTT for 15 min and 4 h show a significant increase in B-chain monomers (3.5 kDa) with appearance of a new peak at 7.1 kDa that indicates B-chain dimers. Relative abundance peak for insulin monomer (5.8 kDa) decreased in presence of DTT when incubation time was increased from 15 min to 4 h. (Figure 3.2C, 3.2D). Appearance of a new peak was observed at 9.6 kDa for insulin incubated with DTT at 15 min and 4 h samples suggesting presence of aggregated forms of B-chain dimer with A-chain or presence of tetramer of A-chain. The data suggests cleavage of inter-chain disulfide bonds in insulin at positions A7-B7 and A20-B19 leading to generation of insulin A and B chains. An increase in molecular weight of ~20-30 Da for the samples could be due to the residual buffer ions present as contaminants in the sample.

We monitored aggregation kinetics and secondary structure of insulin at pH 7.2, 37 °C using various spectroscopic techniques (Figures 3.3 to 3.8) and morphology of aggregates were characterized by SEM and AFM (Figures 3.9 to 3.11 and 3.21, 3.22).

We investigated effect of 10 mM DTT on insulin instability and aggregation. As turbidity is a direct measure of the amount of aggregation, the optical density of the aggregated samples were measured at 600 nm. This method is very useful to detect aggregates with a hydrodynamic radius greater than the wavelength of light used^{24, 41}. The aggregation propensity was highest in the first 20 min, after which the process slowed down considerably. Proteins that were incubated for shorter time periods showed a very sharp increase in absorbance within the first 10 min, implying rapid initial aggregation (inset in Figure 3.3). This suggests increased instability and unfolding of insulin in presences of DTT that leads to misfolding and aggregation of the protein within 10 min of incubation. Controls similarly incubated in the absence of DTT did not exhibit any turbidity or increase in absorbance.

Insulin does not have any tryptophan residues therefore intrinsic fluorescence of insulin was measured by monitoring the exposure of tyrosine (Tyr) residues at 306 nm^{42,43}. Decrease in fluorescence intensity at 306 nm (Figure 3.4) suggests exposure of Tyr residues to polar (aqueous) environment leading to fluorescence quenching as a consequence of protein unfolding⁴⁴. Fluorescence intensity decreases rapidly in first 20 min followed by a gradual decrease up to 48 h, which is in line with the UV absorbance data. This suggests that insulin in presence of disulfide reducing agent unfolds rapidly exposing the hydrophobic core (aromatic residues) followed by

aggregation. This can be a direct result of catalytic amount of reducing agent promoting ‘*disulfide scrambling*’ and drastically altering the protein stability^{21, 25, 26}. The effect of 10 mM DTT on protein misfolding and hydrophobic exposure was measured using fluorescent dye ANS and its dimeric analog bis-ANS. ANS is highly fluorescent and blue shifts (peak maximum shifts to 470 nm) upon binding to the hydrophobic core of the protein⁴⁵. The dimeric analog of ANS, bis-ANS shares similar properties to ANS but has a higher binding affinity for molten globule like structures of proteins (Figure 3.5). Both dyes exhibit similar trends with rapid increase in fluorescence intensity in the first h followed by high but steady fluorescence emission at longer incubation times. In presence of disulfide-reduced insulin at 172 μ M concentration, ANS shows rapid increase in fluorescence in the first few min of incubation that peaks and then steadies in 1 h (Figure 3.5A). The peak wavelength for ANS stays steady at $\sim$ 470 nm for the length of incubation (Figure 3.6). The disulfide reduced insulin at 90 and 20 μ M also showed emission spectra similar to that seen for 172 μ M protein (Figure 3.7A, 3.7C). Interestingly, at these lower protein concentrations the peak wavelength for ANS that is initially around 500 nm shifts to $\sim$ 470 nm in about 20 min and stays steady at $\sim$ 470 nm even upon longer incubation (Figure 3.7B, 3.7D). This suggests that disulfide reduced insulin unfolds and the hydrophobic amino acid residues that are initially solvent exposed collapse and aggregate burying the hydrophobic residues (Figure 3.7). At protein concentration of 172 μ M the unfolded protein burden is very high and it rapidly undergoes hydrophobic collapse aggregating with peak wavelength observed at $\sim$ 470 nm (Figure 3.6). Controls (proteins with intact disulfide bonds) did not show change in fluorescence as a function of time. This is in line with other protein

aggregation studies where exposure of hydrophobic amino acids is followed by collapse and burial of the hydrophobic core^{35, 46-48}.

ThT was used to investigate the nature of insulin aggregates. ThT showed very rapid increase in fluorescence in the first 1 h of incubation with disulfide reduced insulin followed by gradual increase in fluorescence for the rest of the length of incubation (Figure 3.8). This observed trend of fluorescence increase for ThT is similar to earlier protein aggregation studies suggesting amyloid formation^{27, 49-51}. The interaction between ThT molecules and amyloid β -sheets results in high fluorescent emission at 480 nm⁴⁵. However, some other studies have shown that ThT dye can also interact with large amorphous aggregates by stacking in between the aggregates, leading to increased fluorescence^{52, 53}. Therefore, ThT fluorescence alone cannot accurately provide structural information about protein aggregates and needs further independent verification. We used SEM to image the aggregates and visualize the morphology of these aggregates. Electron microscope images show that the aggregates are amorphous in nature (Figure 3.9 to 3.11) that stay amorphous even upon longer incubation (Figure 3.11). Size of aggregates ranged between 100-200 nm, depending on the length of incubation (Figures 3.9 to 3.11)³⁰. Comparing the aggregates formed during the first 15 min (Figure 3.9A) with longer incubation time, the density of aggregates was seen to increase dramatically with time (Figure 3.10) but the nature of aggregates remained amorphous. As shown, large insulin aggregates were made of smaller units, suggesting a mechanism similar to '*chain polymerization*' (Figure 3.10). This implies that unfolded insulin monomer comes together to form multimers that in turn join together to form macromolecular aggregates. SEM image of fresh insulin with DTT (control

sample, Figure 3.10) shows very few aggregates that rapidly increases in size and number of aggregates in the first 15 min, after which the change is not as dramatic. This is in line with the observed data for unfolding, misfolding, and aggregation kinetics of insulin (Figures 3.3 to 3.8). Studies by several groups have suggested that amorphous aggregates are precursors of amyloid fibrils^{31, 37, 54, 55}. In order to test if disulfide-reduced insulin at physiological pH that is initially observed as amorphous aggregates turn into fibril; we imaged similarly incubated insulin (pH 7.2, 37 °C) samples after 6 months (Figure 3.11). We only observed amorphous aggregates even after such a long incubation time under the experimental conditions (Figure 3.11). These results are in line with other reports showing formation of amorphous aggregates even upon longer incubation^{31, 56, 57}.

Salt has been shown to accelerate protein aggregation^{58, 59}. Effect of salt concentration was illustrated in the absence and presence of NaCl (150 or 300 mM) (Figures 3.12 to 3.117). Aggregation kinetics showed trends that were similar for proteins with/without NaCl and the results were comparable to our earlier results (Figure 3.3 to 3.8), at all protein concentrations (Figures 3.12 to 3.17). This suggests that insulin aggregation under reducing conditions is independent of ionic strength at pH 7.2. These results are in line with another study where chaotropic role of NaCl was absent in the process of insulin aggregation, both at acidic and at neutral pH¹¹.

Insulin has been shown to aggregate at repeated injection sites of diabetic patients leading to a critical medical condition known as injection amyloidosis⁸. Many studies have shown that aggregation of proteins is propagated by a nucleation-dependent

polymerization method where a “nucleus” is formed by assembly of monomers. Although, this process is thermodynamically and kinetically unfavorable, it decreases the energy barrier for aggregation, reducing the lag time and shifting the equilibrium towards unfolded state of the protein^{24, 60}. To understand how nucleation can impact aggregation of native insulin we prepared “*seeds*” by incubating insulin in the presence of 10 mM DTT for 10 min, 30 min to 4 h (30 min interval), and 5 h to 12 h (1 h interval). Effect of seeding on native insulin aggregation was monitored by ThT fluorescence (Figure 3.26). The ThT fluorescence data for different timed *seeds* up to 4 h was fitted and analyzed by linear regression (Figure 3.23 to 3.25). The slope of the curves shows the kinetics of aggregation (Table 3.1). Aggregation kinetics of native insulin in the presence of 10-min *seeds* was maximum with a slope of 235.92 ± 4.86 (Table 3.1); followed by *seeds* made by incubating insulin for 30 min and 4 h. Aggregation of native insulin in the presence of *seeds* prepared by incubating protein for 5 h or longer (up to 12 h) had the slowest aggregation kinetics (Figure 3.25). Therefore, these seeds were grouped into *nascent* (10 min), *intermediate* (30 min to 4 h) and *mature* (5 h to 12 h) *seeds* based on the difference in their aggregation kinetics (Table 3.1, Figure 3.23 to 3.25).

The rate of aggregation increases continuously in the presence of *nascent* seeds at a concentration of 5% and 15% v/v (Figure 3.18, 3.19). However at a concentration of 45% (Figure 3.20), the aggregation rate is independent of “stage” of seeds (*nascent*, *intermediate*, or *mature* seeds) used to initiate aggregation. This maybe a result of saturation of the number of nucleation sites available for aggregation. This establishes the critical concentration of seeds necessary for proteins to aggregate *in vitro* and that

still shows a difference in aggregation kinetics. In summary, *nascent* seeds had the shortest lag time and fastest aggregation kinetics. This suggests that *nascent* seeds of insulin have high number of soluble misfolded proteins with exposed hydrophobic surface providing very high number of nucleation sites for native insulin to bind and aggregate compared to seeds prepared by longer incubation. The kinetics of aggregation was also observed to linearly depend on the seeding concentration in the presence of *nascent* seeds (Figure 3.27). This suggests that at higher seed concentration (e.g. 15% seed v/v) the number of available nucleation sites for seeds are more compared to seed concentration of 5% v/v and hence higher aggregation rate observed at higher concentration (Figure 3.27). This is in line with another study that suggests partially-folded protein intermediates promote aggregation and not the mature aggregates³⁵.

Morphology of the aggregates formed as a result of seeding were characterized by atomic force microscopy (Figure 3.21, 3.22). The aggregates formed by native insulin incubated with *nascent* or *intermediate* seeds increased in size from 10 nm to 400 nm. Stiffness of aggregates as measured by Young's modulus ranged between 4 ± 0.6 to 6 ± 0.2 GPa for the aggregates formed by *nascent* and *intermediate* seeds, which is comparable to values reported in literature⁶¹⁻⁶⁴. These values suggest that the aggregates are relatively rigid and the high tensile strength of these aggregates maybe a result of strong hydrogen bonding as well as hydrophobic interactions contributing to the mechanical stiffness of the structure^{64, 65}. Interestingly, native insulin when incubated with *intermediate* seeds formed fibrils (Figure 3.21 B, C) that converted to amorphous aggregates upon longer incubation (24 h or more) (Figure 3.22). The

secondary structure of *nascent* and *intermediate* seeds along with native insulin (control) was investigated using FT-IR^{66, 67} to understand the morphological diversity of the seeds (Figure 3.28). A peak at 1656 cm⁻¹ was observed for native insulin that is characteristic of α -helical proteins⁶⁸⁻⁷¹. All the *seeds* show a peak at 1625 cm⁻¹ that is characteristic of β -sheet structure. The raw spectra were deconvoluted and fitted using Gaussian function to obtain the major component peaks. Area under the deconvoluted peaks were integrated to calculate the percent contribution from α -helical or β -sheet structures, for each seeds. The α -helical content decreased with increase in incubation time. *Nascent seeds* (10 min) had nearly 63% of α -helical and 37% of β -sheet structures (Figure 3.28). This ratio changed to ~50% α -helical and ~50% β -sheet for *intermediate seeds* (4 h). This suggests that the *nascent* seeds have significant native structure (α -helix) that can contribute to native insulin aggregation in a nucleation dependent manner. In *nascent seeds* the α -helical structure acts as the site of nucleation for native insulin and the β -sheet component propagates the aggregation process in a template-dependent manner. *Intermediate seeds* (30 min to 4 h) could also promote aggregation of native insulin due to the presence of significant α -helical structure (50% to 60%) present in them. AFM images of *intermediate* seeds incubated with native protein show formation of protofibrillar to fibrillar structures. This observation is complemented by FTIR data wherein increase in β -sheet structure is observed for both 1 and 4 h seeds. Kinetic data shows that 1 h seed is most efficient among *intermediate* seeds, in aggregating native insulin, which suggests that protofibrils are more efficient than fibrils in aggregating native insulin (Table 3.1, Figure 3.28). Conversion of fibrils to amorphous aggregates within 24 h of incubation under our experimental conditions

shows the transient nature of these aggregates (Figure 3.29). This is in line with amorphous aggregates observed for disulfide-reduced insulin incubated at 37 °C for six months in this study. Overall, this study and an earlier study shows that proteins under conditions that are reducing and close to physiological favor formation of amorphous aggregates⁵⁷.

Figure 3.29 Plausible aggregation mechanism model based on insulin aggregation in presence of different timed seeds. *Nascent* seeds promote rapid amorphous aggregate formation; *intermediate* seeds lead to formation of fibrillar structure in 4 h that converts to amorphous upon longer incubation of 24 h. Aggregation in the presence of *mature* seeds is very slow when compared to *nascent* and *intermediate* seed kinetics.

Conclusion

In summary, this study measures the aggregation kinetics and characterizes the nature of aggregates of insulin formed under disulfide reducing conditions at pH 7.2 and 37 °C. The results show that insulin aggregates rapidly under these conditions and forms amorphous aggregates that stay amorphous even upon incubation for six months. *Seeds* affect native insulin aggregation kinetics; with *nascent* seeds showing the fastest aggregation kinetics. Interestingly, incubation of native insulin with *intermediate* seeds for shorter period (4 h) results in fibril formation that converts to amorphous upon longer incubation (24 h). These results suggest that insulin favors amorphous aggregate formation under disulfide reducing conditions at physiological pH and temperature.

References

- (1) Centers for Disease Control and Prevention. *National Diabetes Statistics Report: Estimates of Diabetes and Its Burden in the United States*, **2014**. Atlanta, GA: U.S. Department of Health and Human Services.
- (2) Gasparini, L.; Netzer, W. J.; Greengard, P.; Xu, H. X. Does Insulin Dysfunction Play a Role in Alzheimer's Disease? *Trends Pharmacol. Sci.* **2002**, *23*, 288-293.
- (3) Brems, D. N.; Brown, P. L.; Bryant, C.; Chance, R. E.; Green, L. K.; Long, H. B.; Miller, A. A.; Millican, R.; Shields, J. E.; Frank, B. H. Improved Insulin Stability Through Amino-Acid Substitution. *Protein Engg.* **1992**, *5*, 519-525.
- (4) Wang, W. Protein Aggregation and its Inhibition in Biopharmaceutics. *Int. J. Pharm.* **2005**, *289*, 1-30.
- (5) Givol, D.; Lorenzo F.D.; Goldberger, R.F.; Anfinsen, C.B. Disulfide Interchange and the Three-Dimensional Structure of Proteins. *PNAS* **1965**, *53*, 676-684.
- (6) Morris, H.R.; Pucci, P.; Panico, M.; Marino, G. Protein Folding/Refolding Analysis by Mass Spectrometry. Scrambling of Disulphide Bridges in Insulin. *Biochem. J.* **1990**, *268*, 803-806.
- (7) Gerweck, L.E.; Seetharaman, K. Cellular pH Gradient in Tumor Versus Normal Tissue: Potential Exploitation for the Treatment of Cancer. *Cancer Res.* **1996**, *56*, 1194-1198.

- (8) Dische, F.E.; Wernstedt, C.; Westermark, G.T.; Westermark, P.; Pepys, M.B.; Rennie, J.A.; Gilbey, S.G.; Watkins, P.J. Insulin as an Amyloid-Fibril Protein at Sites of Repeated Insulin Injections in a Diabetic Patient. *Diabetologia* **1988**, *31*, 158-161.
- (9) Ahmad, A.; Millett, I. S.; Doniach, S.; Uversky, V. N.; Fink, A. L. Stimulation of Insulin Fibrillation by Urea-Induced Intermediates. *J. Bio. Chem.* **2004**, *279*, 14999-15013.
- (9) Jiang, C. T.; Chang, J. Y. Unfolding and Breakdown of Insulin in the Presence of Endogenous Thiols. *FEBS Letters* **2005**, *579*, 3927-3931.
- (10) Arora, A.; Ha, C.; Park, C. B. Insulin Amyloid Fibrillation at Above 100 Degrees C: New Insights into Protein Folding Under Extreme Temperatures. *Prot. Sci.* **2004**, *13*, 2429-2436.
- (11) Ahmad, A.; Millett, I. S.; Doniach, S.; Uversky, V. N.; Fink, A. L. Partially Folded Intermediates in Insulin Fibrillation. *Biochemistry* **2003**, *42*, 11404-11416.
- (12) Sluzky, V.; Tamada, J. A.; Klibanov, A. M.; Langer, R. Kinetics of Insulin Aggregation in Aqueous-Solutions upon Agitation in the Presence of Hydrophobic Surface. *PNAS* **1991**, *88*, 9377-9381.
- (13) Brange, J.; Andersen, S.; Laursen, E.D.; Meyn. G.; Rasmussen, E. Towards Understanding Insulin Fibrillation. *J. Pharm. Sci.* **1997**, *97*, 517-525
- (14) Hua, Q. X. Insulin: a Small Protein with a Long Journey. *Protein & Cell* **2010**, *1*, 537-551.
- (15) Nielsen, L.; Khurana, R.; Coats, A.; Frokjaer, S.; Brange, J.; Vyas, S.; Uversky, V. N.; Fink, A. L. Effect of Environmental Factors on the Kinetics of Insulin Fibril

Formation: Elucidation of the Molecular Mechanism. *Biochemistry* **2001**, *40*, 6036-6046.

(16) Westermark, P.; Andersson, A.; Westermark, G. T. Islet Amyloid Polypeptide, Islet Amyloid, and Diabetes Mellitus. *Physiol. Rev.* **2011**, *91*, 795-826.

(17) Bouchard, M.; Zurdo, J.; Nettleton, E. J.; Dobson, C. M.; Robinson, C. V. Formation of Insulin Amyloid Fibrils Followed by FTIR Simultaneously with CD and Electron Microscopy. *Prot. Sci.* **2000**, *9*, 1960-1967.

(18) Dzwolak, W.; Ravindra, R.; Lendermann, J.; Winter, R. Aggregation of Bovine Insulin Probed by DSC/PPC Calorimetry and FTIR Spectroscopy. *Biochemistry* **2003**, *42*, 11347-11355.

(19) Hong, D. P.; Ahmad, A.; Fink, A. L. Fibrillation of Human Insulin A and B Chains. *Biochemistry* **2006**, *45*, 9342-9353.

(20) Hua, Q. X.; Weiss, M. A. Mechanism of Insulin Fibrillation - the Structure of Insulin under Amyloidogenic Conditions Resembles a Protein-Folding Intermediate. *J. Bio. Chem.* **2004**, *279*, 21449-21460.

(21) Grudzielanek, S.; Jansen, R.; Winter, R. Solvational Tuning of the Unfolding, Aggregation and Amyloidogenesis of Insulin. *J. Mol. Bio.* **2005**, *351*, 879-894.

(22) Dzwolak, W.; Smirnovas, V.; Jansen, R. Winter, R. Insulin Forms Amyloid in a Strain-Dependent Manner: an FT-IR Spectroscopic Study. *Prot. Sci.* **2004**, *13*, 1927-1952

- (23) Jarett, J.T.; Lansbury, P.T. Jr. Seeding “One-Dimensional Crystallization” of Amyloid: A Pathogenic Mechanism in Alzheimer’s Disease and Scrapie? *Cell* **1993**, *73*, 1055-1058
- (24) Toichi, K.; Yamanaka, K.; Furukawa, Y. Disulfide Scrambling Describes the Oligomer Formation of Superoxide Dismutase (SOD1) Proteins in the Familial Form of Amyotrophic Lateral Sclerosis. *J. Bio. Chem.* **2012**, *288*, 4970-4980.
- (25) Trivedi, M.V.; Laurence, J.S.; Siahaan, T.J. The Role of Thiols and Disulfides in Protein Chemical and Physical Stability. *Curr. Protein Pept. Sci.* **2009**, *10*, 614-625.
- (26) Li, Y.; Gong, H.; Sun, Y.; Yan, J.; Cheng, B. A.; Zhang, X.; Huang, J.; Yu, M.; Guo, Y.; Zheng, L.; Huang, K. Dissecting the Role of Disulfide Bonds on the Amyloid Formation of Insulin. *Biochem. Biophys. Res. Commun.* **2012**, *423*, 373-378.
- (27) Huang, K.; Maiti, N. C.; Phillips, N. B.; Carey, P. R.; Weiss, M. A. Structure-Specific Effects of Protein Topology on Cross-Beta Assembly: Studies of Insulin Fibrillation. *Biochemistry* **2006**, *45*, 10278-10293.
- (28) Chang, X.; Jorgensen. A.M.; Bardrum, P.; Led, J.J. Solution Structures of the R6 Human Insulin Hexamer. *Biochemistry* **1997**, *36*, 9409-9422.
- (29) Krebs, M. R. H.; Devlin, G. L.; Donald, A. M. Protein Particulates: Another Generic Form of Protein Aggregation? *Biophys. J.* **2007**, *92*, 1336-1342.
- (30) Stefani, M.; Rigacci, S. Protein Folding and Aggregation into Amyloid: The Interference by Natural Phenolic Compounds. *Int. J. Mol. Sci.* **2013**, *14*(6), 12411-12457

- (31) Westermark, P. Aspects on Human Amyloid Forms and their Fibril Polypeptides. *FEBS J.* **2005**, *272*, 5942-5949.
- (32) Khurana, R.; Coleman, C.; Ionescu-Zanetti, C.; Carter, S.A.; Krishna, V.; Grover, R.K.; Roy, R.; Singh, S. Mechanism of Thioflavin T Binding to Amyloid Fibrils. *J. Struc. Bio.* **2005**, *151*, 229-238.
- (33) Ahmad, A.; Uversky, V. N.; Hong, D.; Fink, A. L. Early Events in the Fibrillation of Monomeric Insulin. *J. Bio. Chem.* **2005**, *280*, 42669-42675.
- (34) Nielsen, L.; Frokjaer, S.; Brange, J.; Uversky, V. N.; Fink, A. L. Probing the Mechanism of Insulin Fibril Formation with Insulin Mutants. *Biochemistry* **2001**, *40*, 8397-8409.
- (35) Groenning, M.; Frokjaer, S.; Vestergaard, B. Formation Mechanism of Insulin Fibrils and Structural Aspects of the Insulin Fibrillation Process. *Curr. Protein Pept. Sci.* **2009**, *10*, 509-528.
- (36) Fink, A. L. Protein Aggregation: Folding Aggregates, Inclusion Bodies and Amyloid. *Fold Des.* **1998**, *3*, R9-R23.
- (37) Yoshimura, Y.; Lin, Y. X.; Yagi, H.; Lee, Y. H.; Kitayama, H.; Sakurai, K.; So, M.; Ogi, H.; Naiki, H.; Goto, Y. Distinguishing Crystal-Like Amyloid Fibrils and Glass-Like Amorphous Aggregates from their Kinetics of Formation. *PNAS* **2012**, *109*, 14446-14451.

- (38) Tiwari, A.; Hayward, L. J. Familial Amyotrophic Lateral Sclerosis Mutants of Copper/Zinc Superoxide Dismutase are Susceptible to Disulfide Reduction. *J. Bio. Chem.* **2003**, *278*, 5984-5992.
- (39) Lee, J. Y.; Hirose, M. Partially Folded State of the Disulfide-reduced Form of Human Serum Albumin as an Intermediate for Reversible Denaturation. *J. Bio. Chem* **1992**, *267*, 14753-14758.
- (40) Creighton, T. E. Disulphide Bonds and Proteins Stability. *BioEssays* **1988**, *8*, 57-63.
- (41) Jarett, J.T.; Lansbury, P.T. Jr. Amyloid Fibril Formation Requires a Chemically Discriminating Nucleating Event: Studies of an Amyloidogenic Sequence from the Bacterial Protein OsmB. *Biochemistry* **1992**, *31*(49), 12345-52
- (42) Vincentelli, R.; Canaan, S.; Campanacci, V.; Valencia, C.; Maurin, D.; Frassinetti, F.; Scappucini-Calvo, L.; Bourne, Y.; Cambillau, C.; Bignon, C. High-Throughput Automated Refolding Screening of Inclusion Bodies. *Prot. Sci.* **2004**, *13*, 2782-2792.
- (43) Cromwell, M. E. M.; Hilario, E.; Jacobson, F. Protein Aggregation and Bioprocessing. Paper Presented at the Proceedings of the 2005 AAPS Biotec Open Forum on Aggregation of Protein Therapeutics. San Francisco, CA.
- (44) Vivian, J. T.; Callis, P. R. Mechanisms of Tryptophan Fluorescence Shifts in Proteins. *Biophys. J.* **2001**, *80*, 2093-2109.
- (45) Hawe, A.; Sutter, M.; Jiskoot, W. Extrinsic Fluorescent Dyes as Tools for Protein Characterization. *Pharm. Res.* **2008**, *25*, 1487-1499.

- (46) 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. Bio.* **2005**, *350*, 379-392.
- (47) Dobson, C. M. Protein Folding and Misfolding. *Nature* **2003**, *426*, 884-890.
- (48) Ali, V.; Prakash, K.; Kulkarni, S.; Ahmad, A.; Madhusudan, K. P.; Bhakuni, V. 8-Anilino-1-naphthalene Sulfonic Acid (ANS) Induces Folding of Acid Unfolded Cytochrome C to Molten Globule State as a Result of Electrostatic Interactions. *Biochemistry* **1999**, *38*, 13635-13642.
- (49) Makin, O.M.; Serpell, L.C. Structure of Amyloid Fibrils. *FEBS J.* **2005**, *272*, 5950-5961.
- (50) Krebs, M.R.H.; Bromley, E.H.; Donald, A.M. The Binding of Thioflavin-T to Amyloid Fibrils: Localisation and Implications. *J. Struc. Bio.* **2005**, *149*, 30-37.
- (51) Groenning, M.; Olsen, L.; Van de Weert, M.; Flink, J. M.; Frokjaer, S; Jørgensen, F. S. Study on the Binding of Thioflavin T to Beta-Sheet-Rich and Non-Beta-Sheet Cavities. *J. Struc. Bio.* **2007**, *158*, 358-369.
- (52) Ellis, R. J.; Pinheiro, T. J. Medicine: Danger -- Misfolding proteins. *Nature* **2002**, *416*, 483-484.
- (53) Chiti, F.; Dobson, C. M. Protein Misfolding, Functional Amyloid, and Human Disease. *Annu. Rev. Biochem.* **2006**, *75*, 333-366.
- (54) Kopito, R. R. Aggresomes, Inclusion Bodies and Protein Aggregation. *Trends Cell Bio.* **2000**, *10*, 524-530.

- (55) Stefani, M. Biochemical and Biophysical Features of both Oligomer/Fibril and Cell Membrane in Amyloid Cytotoxicity. *FEBS J.* **2010**, 277, 4602-4613.
- (56) Yang, M.; Dutta, C.; Tiwari, A. Disulfide Bond Scrambling Promotes Amorphous Aggregates in Lysozyme and Bovine Serum Albumin. *J. Phys. Chem. B* **2015**, 119(10), 3969-81
- (57) Muzaffar, M.; Ahmad, A. The Mechanism of Enhanced Insulin Amyloid Fibril Formation by NaCl Is Better Explained by a Conformational Change Model. *PLoS ONE* **2011**, 6(11), e27906.
- (58) Chi, E.Y.; Krishnan, S.; Kendrick, B.S.; Chang, B.S.; Carpenter, J.F.; Randolph, T.W. Roles of Conformational Stability and Colloidal Stability in the Aggregation of Recombinant Human Granulocyte Colony-Stimulating Factor. *Prot. Sci.* **2003**, 12, 903-913.
- (59) Harper, J.D.; Lansbury, P.T. Jr. 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
- (60) Adamcik, J.; Lara, C.; Usov, I.; Jeong, J.S.; Ruggeri, F.S.; Dietler, G.; Lashuel, H.A.; Hamley, I.W.; Mezzenga, R. Measurement of Intrinsic Properties of Amyloid Fibrils by Peak Force QNM Method. *Nanoscale* **2012**, 4, 4426-4429
- (61) Adamcik, J.; Berquand, A.; Mezzenga, R. Single-step direct measurement of amyloid fibrils stiffness by peak force quantitative nanomechanical atomic force microscopy. *Appl. Phys. Lett.* **2011**, 98, 193701

- (62) Adamecik, J.; Mezzenga, R. Study of amyloid fibrils via atomic force microscopy. *Curr. Opin. Colloid Interface Sci.* **2012**, 17(6), 369-376
- (63) Smith, J.F.; Knowles, T.P.; Dobson, C.M.; MacPhee, C.E.; Welland, M.E. Characterization of nanoscale properties of individual amyloid fibrils. *PNAS* **2006**, 103(43), 15806-15811
- (64) Knowles, T.P.; Fitzpatrick A.W.; Meehan, S.; Mott, H.R.; Vendruscolo, M.; Dobson, C.M., Welland, M.E. Role of intermolecular forces in defining material properties of protein nanofibrils. *Science* **2007**, 318, 1900-1903
- (65) Srivastava, A.; Arya, P.; Goel, S.; Kundu, B.; Mishra, P.; Fnu, A. Gelsolin Amyloidogenesis Is Effectively Modulated by Curcumin and Emetine Conjugated PLGA Nanoparticles. *PloS ONE* **2015**, 10(5) e0127011
- (66) Srivastava, A.; Arya, P.; Vasaikar, S.V.; Mukherjee, G.; Mishra, P.; Kundu, B. Selective Interception of Gelsolin Amyloidogenic Stretch Results in Conformationally Distinct Aggregates with Reduced Toxicity. *ACS Chem.Neurosci.* **2014**, 5, 982-992
- (67) Shivu, B.; Seshadri, S.; Li, J.; Oberg, K.A.; Uversky, V.N.; Fink, A.L. Distinct β -sheet Structure in Protein Aggregates Determined by ATR-FTIR Spectroscopy. *Biochemistry* **2013**, 52, 5176-5183
- (68) 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

- (69) Kong, J.; Yu, S. Fourier Transform Infrared Spectroscopic Analysis of Protein Secondary Structures. *Acta Biochim. Biophys.* **2007**, 39(8), 549-559
- (70) Amenabar, I; Poly, S.; Nuansing, W.; Hubrich, E.H.; Govyadinov, A.A.; Huth, F.; Krutokhvostov, R.; Zhang, L.; Knez, M.; Heberle, J.; Bittner, A.M.; Hillenbrand, R. Structural Analysis of Mapping of Individual Protein Complexes by Infrared Nanospectroscopy. *Nat. Comm.* **2013**, 4(2890)

Chapter 4. Disulfide Bond Scrambling at Acidic pH Promotes Insulin Aggregate Toxicity²

Colina Dutta, and Ashutosh Tiwari*

Department of Chemistry, Michigan Technological University, Houghton, MI 49931

***Correspondence:** Ashutosh Tiwari, Department of Chemistry, Michigan Technological University, Houghton, MI, USA,

Tel:(906) 487-1840; Fax: (906) 487-2061; Email: tiwari@mtu.edu

² The material contained in this chapter will be published in RSC Publications.

Abstract

Observation of protein aggregates in end stage tissues is hallmark of several neurodegenerative diseases such as Amyotrophic Lateral Sclerosis, Alzheimer's and Parkinson's disease. However, the relationship between morphology of aggregated proteins and their associated toxicity is not clear. To better understand the relationship between distinct aggregated forms of proteins and their toxicity, we used insulin as a model protein. Morphologically distinct insulin aggregates were prepared by incubating the protein at various pH (2.0, 5.3, 7.0 and 8.8) and temperature (37 °C or 65 °C) in presence or absence of disulfide reducing agent. The morphology and unique physic-chemical properties of the aggregates were characterized using spectroscopy and electron microscopy techniques. Insulin aggregates within 4 h in the presence of reducing agent at all pH studied. At acidic pH (pH 2.0 or pH 5.3) it forms *rod-like* aggregates at 37 °C and *fibrils* at 65 °C in absence of reducing agent. In the presence of reducing agent, insulin forms amorphous aggregates with distinct morphology at pH 2.0 and 7.0. At pH 5.3 and 8.8, *filamentous* structures of varying lengths were observed. The cytotoxicity of these distinct aggregates were measured using human neuroblastoma (SH-SY5Y) cells. Overall, the results show that aggregates formed under disulfide reducing conditions at acidic pH were more toxic.

Introduction

Proteins are functional within a narrow range of temperature and pH, deviation from which leads to physical and/or chemical instability causing them to aggregate. Neurodegenerative diseases such as Amyotrophic Lateral Sclerosis, Parkinson's and Alzheimer's disease are a result of protein misfolding and aggregation¹⁻⁴. However, the general mechanism by which these aggregates cause toxicity is not well understood. This is further complicated by the fact that there is no clear relationship between the morphology and toxicity of these aggregates. This diversity has been observed between different proteins as well as within mutants of the same protein^{5, 6}. Majority of these disease causing proteins form β -sheet rich structures *in vivo* that are collectively known as amyloid fibrils, however structurally distinct fibrils have been observed in A β ₁₋₄₂ and prion protein aggregates^{7, 8}. These fibrils were also shown to possess diverse interaction patterns. Heterogeneous aggregate forms have also been observed in tissue-derived aggregates of lysozyme, apolipoprotein and tau protein with disparate cytotoxicity⁷. Amyloid deposits have been occasionally found to *contain alternate aggregate forms* within amyloid plaques. These include Creutzfeld-Jacob disease, Down's syndrome and rat model of Alzheimer's disease⁹⁻¹³. *In vitro* studies have also reported the formation of *alternate aggregate forms* in pathogenic Ig light chain, human serum albumin, lysozyme, β -lactoglobulin, insulin, and in solution of 40-residue A β -peptide¹⁴⁻²¹. Studies with A β protein have shown that the longer peptide (1-42) is more toxic to neuronal cells than A β ₁₋₄₀ and form amylospheroid like structures²². Mature aggregate forms of A β ₁₋₄₂ exhibited high toxicity even at nano-molar concentrations when compared to immature oligomers²³. Other disease related proteins such as

superoxide dismutase 1 (SOD1) also exhibit toxicity due to lack of copper and/or zinc in both wild type and mutant protein forms²⁴. Although, a host of information is available on aggregate morphology and cytotoxicity from both *in vitro* and *in vivo* studies, yet the relationship between structure and toxicity of these aggregates is still a matter of debate. These observations suggest that the morphology of protein aggregates is highly dependent on aggregation conditions such as temperature, pH and other stresses in the cellular environment.

To better understand role of misfolded proteins in the disease process, we need to characterize the structural changes in proteins due to change in pH, temperature, and external stresses and correlate aggregate morphology to its toxicity. The goal is to generate distinct aggregate forms from a single protein by manipulating the proteins local environment and test its toxicity. For this we chose a commonly studied model protein insulin as it can generate distinct aggregate forms. Insulin is an aggregation prone protein that forms fibrils at low pH and forms spherulites at elevated temperatures (65 – 75 °C)^{18, 25-29}, and amorphous aggregates near neutral pH under disulfide reducing conditions⁴⁴.

Insulin is a 51-residue protein consisting of two polypeptide chains A and B with 21 and 30 residues respectively. There are two inter-chain disulfide bonds (A7-B7, A20-B19) and one intra-chain disulfide bond (A6-A11) that substantially constrain the structure and native fold of insulin^{30, 31}. The intra-chain disulfide bonds are buried within the hydrophobic core of the protein. Insulin aggregation has been studied extensively only at acidic pH due to ease of fibril formation, whereas fewer studies

have been done at physiological pH where amorphous aggregates and particulates are prevalent.³²⁻³⁷. Insulin stability is dictated by its conserved residues (Val^{A3}, Ile^{A13}, Val^{B12}, Val^{B18}, Phe^{B24}, and Tyr^{B26}) that are solvent exposed. At acidic pH, the net charge on these amino acid residues become positive owing to their isoelectric point ranging between 5 and 6, leading to destabilization of the native fold and subsequent aggregation³⁸. It is known that disulfide bridges are retained in the insulin fibrillation process and also the individual chains are capable of fibrillation that are structurally different from the intact insulin fibril^{39, 40}. Partial reduction of disulfide bonds in the presence of tris(2-carboxyethyl)phosphine (TCEP) at low pH lead to the formation of flexible filaments of bovine insulin⁴⁰. Toxicity of filaments formed from reduced bovine insulin at pH 1.6 is considerably lower than fibrils formed from insulin with intact disulfide bridges⁴⁰. Insulin oligomers and fibrils at acidic pH have been found to be toxic to PC-12 and SH-SY5Y cell lines^{41, 42}. However, a systematic study on insulin aggregates formed at other pHs and temperature toxicity information of aggregates formed at different pH and temperature under disulfide reducing conditions is lacking. Hence it is important to understand the factors that trigger the formation of these alternate forms of aggregates, their morphology and the relationship to cytotoxicity. Having a better understanding of different aggregated forms of insulin and its relationship to cellular toxicity can help shed light on other protein aggregation diseases owing to the generic mechanism of protein aggregation³.

In this study, we prepared aggregates at acidic, neutral, and basic pH and also at the isoelectric point (pH 5.3)⁴³ of insulin. Aggregation of insulin at pH above and below its isoelectric point is important for understanding the effect of net charge on the protein

and its contribution towards protein unfolding and subsequent aggregation. The aggregates were prepared by incubating insulin at physiological (37 °C) and elevated (65 °C) temperature to investigate the temperature dependence of aggregate morphology and aggregation kinetics. Disulfide bonds are central to protein stability and reduction of disulfide bonds can lead to aggregation in protein including insulin^{44,45}. Therefore, to understand the effect of disulfide bond reduction on insulin aggregation and toxicity, we used 10 mM tris(2-carboxyethyl)phosphine (TCEP) – a non-thiol based reducing agent. Comparing and contrasting the distinct aggregate forms of insulin and their toxicity at different pHs, temperature, and oxidizing/reducing conditions will lead to a better understanding of the relationship between different aggregate forms and their toxicity.

In this study, we demonstrate that insulin forms fibrils at acidic pH (2.0 and 5.3) but forms largely amorphous structures at basic pH (8.8), in the absence of TCEP. Insulin incubated without TCEP at pH 7 and 37 °C did not show any aggregation. In the presence of the reducing agent short, branched and long filaments were formed at pH 5.3 and 8.8, respectively. Amorphous aggregate forms dominate at acidic and neutral pH. Comparison of the toxicity levels of the various aggregates show that insulin aggregates formed in the presence of reducing agent are more toxic to SH-SY5Y cells compared to non-reduced protein aggregates. We also observed that loose amorphous and globular aggregates formed at pH 2.0 are most toxic followed by the filamentous (short, branched and long) structures formed at the isoelectric point as well as basic pH, respectively. Interestingly, incubation temperature plays an important role in

accelerating the aggregation process, rather than altering the overall morphology and cytotoxicity of the aggregates.

Materials and Methods

Unless otherwise indicated, all materials were used as supplied by the manufacturer without any further purification. Human recombinant insulin, tris(2-carboxyethyl)phosphine (TCEP), thioflavin T (ThT), 8-anilino-1-naphthalenesulfonic acid (ANS) dye, 4,4'-Dianilino-1,1'-binaphthyl-5,5'-disulfonic acid (Bis-ANS) dye were purchased from Sigma.

Preparation of insulin samples - Insulin stock solutions were prepared as described earlier⁴⁴. All insulin samples were prepared to a final concentration of 1 mg/mL in 20 mM of respective buffers at pH 2.0 (Glycine-HCl), pH 5.3 and 8.8 (Tris-HCl), pH 7.0 (HEPES), either in presence or absence of 10 mM tris(2 carboxyethyl)phosphine (TCEP). The samples were prepared on ice followed by incubation both at 37 °C and 65 °C for the indicated time periods (see figures for details). Blanks (contained everything except insulin) were similarly prepared and incubated as samples and were used for background subtraction.

Intrinsic and extrinsic fluorescence- Fluorescence measurements were carried out using a Horiba Jobin Yvon spectrofluorometer (Fluoromax-4). Excitation and emission slit widths were set at 2 nm, each. For intrinsic fluorescence, emission spectra were acquired from 285 to 450 nm with excitation at 280 nm. For extrinsic fluorescence, emission spectra were acquired in presence of 10 μ M ThT, 5 μ M ANS, or 1 μ M Bis-ANS. The concentrations of dye stocks were determined using their extinction

coefficients: ANS $\epsilon_{350 \text{ nm}} = 5000 \text{ M}^{-1} \text{ cm}^{-1}$, bis-ANS $\epsilon_{385 \text{ nm}} = 16790 \text{ M}^{-1} \text{ cm}^{-1}$, and ThT $\epsilon_{416 \text{ nm}} = 26,620 \text{ M}^{-1} \text{ cm}^{-1}$. Protein samples were diluted to a final concentration of 90 μM using 20 mM of respective buffers for intrinsic, ANS and Bis-ANS fluorescence measurements. A final sample concentration of 30 and 15 μM were used for ThT fluorescence measurements. After timed incubation, diluted protein samples were equilibrated with respective fluorophores on ice, in dark for 30 min before acquiring the spectra. For ThT fluorescence, emission spectra were collected from 460 to 650 nm with excitation at 450 nm. Fluorescence spectra from 400 to 700 nm were collected with excitation at 380 nm for ANS and at 360 nm for bis-ANS. All measurements were done in triplicates.

Fitting of ThT Fluorescence Data – ThT fluorescence data in the absence of reducing agent at pH 2.0 and 5.3 at 65 °C was fitted using a Gompertz function given by $y(t) = A \cdot e^{-e^{-\frac{(t-t_i)}{b}}}$ where t_i is the lag time and b is the apparent aggregation rate ⁶³. This function is used to fit sigmoidal growth/aggregation curves with relatively long lag times.

Field Emission Scanning Electron Microscopy (FESEM) – Electron microscope images were collected using Hitachi S-4700 field-emission scanning electron microscope (FESEM). Samples incubated for indicated times were diluted 4-fold with MQ water and aliquoted in Millipore Amicon[®] Ultra centrifugal filters (3 kDa cut off). The diluted samples were centrifuged at 7000 x g at 4 °C (three repeats after dilution with water) to remove low molecular weight impurities (salt, buffer etc). The washed

samples were aliquoted on SEM stubs and air-dried. The samples were platinum coated using a sputter coater and imaged using an acceleration voltage of 5 kV.

Cell Culture – Human neuroblastoma (SH-SY5Y) cells were maintained in DME/F12 medium containing 10% fetal bovine serum, 100 U/mL penicillin and 100 µg/mL streptomycin in 5% CO₂ at 37 °C.

Toxicity Assay – Cell viability was determined using CellTiter 96® AQueous One Solution Cell Proliferation Assay (MTS). It is based on the formation of formazan from tetrazolium salt by mitochondrial dehydrogenase. This reaction changes the color of viable cells to dark purple-blue and its absorbance is measured spectrophotometrically at 490 nm. The percentage of viable cells was determined using the absorbance of control cells as 100%.

Insulin samples were washed with double-distilled water and diluted to a final concentration of 10 µM using cell-growth medium. Cells were plated in 96-well plates at a density of 15000 cells/well and grown overnight. The cells were then incubated with 100 µL of medium with and without insulin aggregates and incubated for 48 hours in 5% CO₂ at 37 °C. Blanks containing insulin samples and medium were also incubated similarly, and used for background subtraction. After incubation, 20 µL of MTS reagent was added to each well and incubated for 4 hours at 37 °C, 5% CO₂ incubator. The absorbance of the cells was read using a micro-plate reader at 490 nm. All measurements were done in triplicates.

Figure 4.1 Ribbon structure of insulin (A) monomer at pH 2.0, (B) dimer at pH 5.3, and >8, and (C) hexamer at pH 7.0 showing the disulfide bonds (S–S, sand-yellow). The backbone is shown in gray color. The structures were generated using PyMOL 1.3 and PDB files 1GUJ for monomer and 1AI0 for dimer and hexamer.

Figure 4.2 Protein misfolding and hydrophobic exposure monitored by ANS fluorescence. Insulin samples incubated at a concentration of 172 μM were diluted to a final concentration of 90 μM and incubated with 5 μM ANS for 15 minutes before acquiring spectra. Emission spectra were collected from 400-650 nm with excitation at 380 nm for (A) pH 2.0 (B) pH 5.3 (C) pH 7.0 and (D) pH 8.8. Average peak emission wavelength of 467 nm was selected. Peak fluorescence intensities are shown using open symbols (without TCEP) and closed symbols (10 mM TCEP). Data is representative of triplicates with error bars showing $\pm$ SD.

Figure 4.3 Protein misfolding and conformational tightening monitored by Bis-ANS fluorescence. Insulin samples incubated at a concentration of 172 μ M were diluted to a final concentration of 90 μ M and incubated with 1 μ M Bis-ANS for 15 minutes before acquiring spectra. Emission spectra were collected from 400-650 nm with excitation at 360 nm for (A) pH 2.0 (B) pH 5.3 (C) pH 7.0 and (D) pH 8.8. Average peak emission wavelength of 482 nm was selected. Peak fluorescence intensities are shown using open symbols (without TCEP) and closed symbols (10 mM TCEP). Data is representative of triplicates with error bars showing $\pm$ SD.

Figure 4.4 Protein aggregation monitored by ThT fluorescence. Insulin samples incubated at a concentration of 172 μM were diluted to a final concentration of 90 μM and incubated with 10 μM ThT for 30 minutes before acquiring spectra. Emission spectra were collected from 460–650 nm with excitation at 450 nm for (A) pH 2.0 (B) pH 5.3 (C) pH 7.0 and (D) pH 8.8. Average peak emission wavelength of 486 nm was selected. Peak fluorescence intensities are shown using open symbols (without TCEP) and closed symbols (10 mM TCEP). Data is representative of triplicates with error bars showing $\pm$ SD.

Figure 4.5 Protein misfolding and hydrophobic exposure monitored by ANS fluorescence. Insulin samples incubated at a concentration of 172 μ M were diluted to a final concentration of 90 μ M and incubated with 5 μ M ANS for 15 minutes before acquiring spectra. Emission spectra were collected from 400-650 nm with excitation at 380 nm for (A) pH 2.0 (B) pH 5.3 (C) pH 7.0 and (D) pH 8.8. Average peak emission wavelength of 467 nm was selected. Peak fluorescence intensities are shown using open symbols (without TCEP) and closed symbols (1 mM TCEP).

Figure 4.6 Protein aggregation monitored by ThT fluorescence. Insulin samples incubated at a concentration of 172 μ M were diluted to a final concentration of 90 μ M and incubated with 10 μ M ThT for 30 minutes before acquiring spectra. Emission spectra were collected from 460–650 nm with excitation at 450 nm for (A) pH 2.0 (B) pH 5.3 (C) pH 7.0 and (D) pH 8.8. Average peak emission wavelength of 486 nm was selected. Peak fluorescence intensities are shown using open symbols (without TCEP) and closed symbols (1 mM TCEP).

Figure 4.7 Rate of protein aggregation calculated using ThT fluorescence data. Insulin samples incubated at a concentration of 172 μM were diluted to a final concentration of 45 μM and incubated with ThT as detailed in Figure 4. Apparent rate of aggregation and lag time for (A) pH 2.0 and (B) pH 5.3 in the absence of TCEP at 65 °C were calculated after fitting the data to Gompertz growth function. The data points are experimental values fitted using the function. Data is representative of triplicates with error bars showing $\pm$ SD.

Figure 4.8 FE-SEM Images of insulin aggregates without TCEP. One hundred seventy two μM insulin was incubated at 37 °C (A-D) and 65 °C (E-H) for 4 h at pH 2.0 (A, E), pH 5.3 (B, F), pH 7.0 (C, G) and pH 8.8 (D, H). Scale bar is 1 μm for all images.

Figure 4.9 FE-SEM Images of insulin aggregates with 10 mM TCEP. One hundred seventy two μM insulin was incubated at 37 °C (A-D) and 65 °C (E-H) for 4 h at pH 2.0 (A, E), pH 5.3 (B, F), pH 7.0 (C, G) and pH 8.8 (D, H). Scale bar is 1 μm for all images

Figure 4.10 FE-SEM Images of insulin aggregates without TCEP. One hundred seventy two μM insulin was incubated at 37 °C (A-D) and 65 °C (E-H) for 72 h at pH 2.0 (A, E), pH 5.3 (B, F), pH 7.0 (C, G) and pH 8.8 (D, H). Scale bar is 1 μm for all images.

Figure 4.11 FE-SEM Images of insulin aggregates with 10 mM TCEP. One hundred seventy two μM insulin was incubated at 37 °C (A-D) and 65 °C (E-H) for 72 h at pH 2.0 (A, E), pH 5.3 (B, F), pH 7.0 (C, G) and pH 8.8 (D, H). Scale bar is 1 μm for all images.

Figure 4.12 FE-SEM Images of insulin aggregates without TCEP. One hundred seventy two μM insulin was incubated at 37 °C (A-D) and 65 °C (E-H) for 7 days at pH 2.0 (A, E), pH 5.3 (B, F), pH 7.0 (C, G) and pH 8.8 (D, H). Scale bar is 1 μm for all images.

Figure 4.13 FE-SEM Images of insulin aggregates with 10 mM TCEP. One hundred seventy two μM insulin was incubated at 37 °C (A-D) and 65 °C (E-H) for 7 days at pH 2.0 (A, E), pH 5.3 (B, F), pH 7.0 (C, G) and pH 8.8 (D, H). Scale bar is 1 μm for all images.

Figure 8

Figure 4.14 ThT emission spectra and SEM Images of insulin aggregates incubated for 4 weeks. One hundred seventy two μM insulin was incubated at pH 2.0 (A), pH 5.3 (B), pH 7.0 (C) and pH 8.8 (D) in the presence and absence of 10 mM TCEP at 37 and 65 $^{\circ}\text{C}$. Samples were diluted to a final concentration of 15 μM and incubated with 10 μM ThT for 30 min on ice. Emission spectra was acquired from 460–650 nm after excitation at 450 nm. Insets show SEM images after 4 weeks of incubation for each pH in the absence (a, b; e, f; i, j; m, n) and presence (c, d; g, h; k, l; o, p) of 10 mM TCEP at 37 $^{\circ}\text{C}$ (a, c; e, g; i, k; m, o) and 65 $^{\circ}\text{C}$ (b, d; f, h; j, l; n, p). Scale bar is 5 μm for all images. Data is representative of average of triplicates.

Figure 4.15 Cytotoxicity effects of insulin aggregates by MTS assay. Aggregates were prepared by incubating 172 μ M protein for 7 days, at all the indicated conditions and were diluted with cell growth medium. SH-SY5Y cells were plated in 96 well plates at a density of 15000 cells per well and grown overnight. Cells were then incubated in the absence (control) and presence of 10 μ M insulin aggregates. Data is representative of triplicates with error bars showing $\pm$ SD.

Results & Discussion

Insulin aggregation has been extensively studied and serves as a model protein for non-neuropathic systemic amyloidosis and other amyloid diseases⁴⁷. Insulin is known to aggregate rapidly and form fibrils at acidic pH as well as in the presence of chaotropic agents such as urea, guanidine hydrochloride etc.^{33, 35, 36, 39}. Spherulite and filament formation of insulin has also been observed in *in vitro* studies at high temperatures and acidic pH²⁵⁻²⁷. Few studies have also shown formation of amorphous aggregates of insulin at physiological conditions⁴⁴. This suggests that insulin follows alternate aggregation pathways that are triggered by external conditions such as temperature, pH and oxidizing/reducing stresses, among others. The potential correlation between diverse aggregate forms to disease processes requires information about toxicity as well. Alternate aggregate forms of proteins such as spherulites and gels have been observed within amyloid plaques in certain diseases, although toxicity information of these aggregates are unknown¹². Insulin fibrils and pre-fibrillar oligomers at acidic pH are known to be cytotoxic but toxicity details on *alternate aggregate forms* are unavailable^{41, 42}. Therefore we wanted to study the mechanism of aggregation of various forms of insulin aggregates, measure its associated kinetics, and understand the relation between aggregate structure and its cytotoxicity.

We monitored insulin aggregation kinetics at pH 2.0, 5.3, 7.0 and 8.8 in the absence and presence of reducing agent (10 mM TCEP) at 37 and 65 °C using various spectroscopic techniques (Figures 4.2-4.7, 4.14). Morphology of various aggregate forms was characterized using SEM (Figures 4.8-4.13). Insulin forms morphologically diverse aggregates with distinct spectral properties in the absence and presence of

reducing agent. Cytotoxicity of the diverse aggregate forms was elucidated using human neuroblastoma SH-SY5Y cells (Figure 4.15).

Insulin structure changes with change in pH. Zn-insulin shifts from monomer to its dimeric form at acidic pH with the equilibrium shifting towards the dimer with increase in pH. The protein is predominantly hexamer at neutral/physiological pH (Figure 4.1)^{34, 48, 49}. Beyond pH 7.4 insulin dissociates to form dimers due to loss of zinc binding that is facilitated directly by histidine at position B10⁵⁰. These associations are predominantly dependent on the net charge on the protein and hence are dictated by the solution pH.

Insulin aggregation and conformational change due to hydrophobic exposure was monitored using extrinsic fluorescent dye ANS (Figure 4.2). This dye is highly sensitive to its local environment and causes increase in fluorescence intensity and a corresponding blue shift when it binds to hydrophobic pockets of proteins. Quenching of fluorescence intensity and red shift is observed upon exposure to polar environment. Insulin incubated in the absence of TCEP at 37 °C did not show increase in fluorescence for the entire incubation period. However, with increase in incubation temperature to 65 °C, fluorescence intensity does not increase up to 72 h and 7 days respectively for pH 2.0 and 5.3 (Figures 4.2A, 4.2B), beyond which fluorescence intensity increased very rapidly. Insulin exists as a monomer at acidic pH and it self-associates to form dimer near its isoelectric point. The monomer has two hydrophobic surfaces on each side that are capable of ANS binding and increase in fluorescence intensity. This is in line with the observation that fluorescence intensity rapidly increases after a lag time

of 72 h at pH 2.0, 65 °C. Dimerization of insulin with increase in pH, leads to burial of hydrophobic residues to some extent that is explained by the increase in lag time of aggregation from 72 h to 7 days in case of pH 5.3. Marginal increase in fluorescence is observed at neutral pH, however at basic pH it remained unaffected with increase in incubation temperature from 37 °C to 65 °C (Figures 4.2C, 4.2D). This suggests the absence of unfolding and exposure of hydrophobic residues at this pH due to presence of predominantly dimeric form of the protein at pH 8.8. Insulin aggregation at 37 °C was independent of the pH change, under our experimental conditions. This can be explained in light of previous studies that elevated temperature and/or other stresses are important to overcome the activation barrier in order to shift the equilibrium towards the unfolded state of the protein⁵¹. Earlier studies have shown that insulin aggregates at 37 °C within two weeks of incubation with continuous agitation⁴¹.

In the presence of reducing agent, ANS fluorescence of insulin (Figures 4.2A, 4.2B) showed a rapid increase for the first 4 h for all pH studied. Beyond 4 h fluorescence intensity decreased and steadied for both incubation temperatures, in case of pH 2.0 and 5.3. At pH 7.0 and 8.8, decrease in fluorescence intensity beyond 4 h was continuous till the entire incubation time in case of 65 °C (Figures 4.2C, 4.2D). The amount of reducing agent used is enough to scramble the disulfide bonds of insulin. This causes unfolding and exposure of hydrophobic residues within 4 hours followed by hydrophobic collapse and burial that is marked by decrease in fluorescence intensity. Similar trends were observed in an earlier study with insulin at physiological pH and temperature⁴⁴.

We used Bis-ANS, a dimeric analog of ANS to study the conformational flexibility of the insulin aggregates (Figure 4.3). Bis-ANS is known to fluoresce on binding to flexible molten-globule structures rather than highly rigid and hydrophobic structures. This property was used to assess the conformational tightening of insulin aggregates. Insulin in absence of reducing agent at 37 °C did not show significant change in fluorescence intensity over the 14 days incubation time. At 65 °C temperature, insulin in the absence of reducing agent showed fluorescence intensity increased rapidly within 4 h followed by a gradual decrease in intensity, at pH 5.3, 7.0 and 8.8 (Figures 4.3B, 4.3C, 4.3D). However, at pH 2.0 (Figure 4.3A), intensity dropped beyond 4 hours followed by gradual increase after 48 h of incubation time. This suggests that at pH 2.0 and 5.3 newly formed aggregates (initial 4 h) are flexible that bind to Bis-ANS however with time the aggregates reorient themselves to form tighter conformations that is signified by the drop in fluorescence intensity past 4 h of incubation. For TCEP-treated insulin at both temperatures (37 and 65 °C) and all pH (2.0, 5.3, 7.0 and 8.8) bis-ANS fluorescence increased rapidly for 4 h followed by a gradual decrease signifying the formation of flexible aggregates that reorient to eventually form rigid structures leading to decrease in fluorescence.

Morphology of insulin aggregates were monitored using ThT fluorescence (Figure 4.4, 4.7) and further confirmed by scanning electron microscopy (Figures 4.8-4.13). ThT is classically known to bind to amyloid fibrils with a characteristic fluorescence. Insulin aggregates in oxidizing environment at 37 °C (all pH) did not show any fluorescence for the entire incubation time. This supports the ANS and bis-ANS fluorescence data confirming the non-existence of hydrophobic pockets or large aggregate forms at

physiological temperature incubation conditions. Electron micrographs confirm the fluorescence data finding. At acidic (2.0 and 5.3) and basic pH (8.8), insulin formed *rod-like* and amorphous aggregates respectively, without the characteristic fluorescence. This suggests the possibility that at physiological temperature, the cavity of aggregates formed is $< 8 \text{ \AA}$ where ThT molecule could bind without inducing fluorescence⁵². No aggregation was observed at neutral pH at 37 °C that is in line with an earlier study⁴⁴.

Insulin incubated at elevated temperature (65 °C) in the absence of reducing agent, show drastic increase in ThT fluorescence after 4h, 72 h and 7 days of incubation in case of pH 2.0 and 5.3 (Figure 4.4). The behavior is similar to that of ANS fluorescence (Figure 4.2) confirming that the aggregates formed at acidic pH is either fibrils or large structures with significant hydrophobic exposure. SEM images of the aggregates confirmed fibril formation for both pH conditions (Figures 4.8, 4.10, 4.12). Fluorescence intensity at pH 7.2 and 8.8 were almost half compared to acidic pH and the aggregates observed by SEM were amorphous. Earlier studies have shown that ThT can bind to large amorphous structures and the fluorescence is almost one-fifth to that of fibrils, in this case⁵². Increased fluorescence intensity of amorphous aggregates of insulin at pH 8.8, as compared to the study by Yang *et al.* (2015) suggests at the aggregate cavity size is $> 8 \text{ \AA}$ leading to increase in ThT fluorescence⁴⁵.

TCEP-treated insulin at both incubation temperatures follows similar trends. At pH 2.0, 5.3 and 7.0, fluorescence intensity gradually increases for the 14 days incubation period, although at pH 8.8, ThT fluorescence decreases after 7 days of incubation. This

suggests that reduced insulin follows similar aggregation pattern that is more or less independent of the pH. The aggregate formation is slow with typically low cps counts that might be accounted by the small size of cavity of the ThT molecule to bind. Electron micrographs show amorphous aggregate formation in the presence of TCEP at pH 2.0 and 7.0 (both temperatures) whereas filamentous structures were observed at the isoelectric point and basic pH (Figures 4.9, 4.11, 4.13). An earlier study with TCEP induced bovine-insulin aggregation showed formation of filamentous aggregates at pH 1.6 and high temperature, which is contrary to the observation under our experimental conditions ⁴⁰. However, the authors found intact fibrils in the absence of TCEP that support our findings.

We also wanted to understand whether decreasing the concentration of reducing agent decreased the kinetics of aggregation or altered the mechanism as well. We characterized insulin aggregates formed in the presence of 1 mM TCEP using ANS and ThT fluorescence (Figures 4.5, 4.6). We observed overall decrease in fluorescence intensity although the fluorescence trend remained unaltered, thereby confirming slower aggregation kinetics without altering the aggregation mechanism.

The aggregation kinetics at pH 2.0 and 5.3 in the absence of TCEP at 65° C, followed a characteristic sigmoidal trend with a lag phase, growth phase and saturation phase of aggregate formation (Figure 4.7). Hence, ThT fluorescence data was fitted to a Gompertz growth function. This function provides the best fit to sigmoidal curves with long lag times and assumes first order reaction kinetics ⁵³⁻⁵⁶. The fitted function was used to calculate the lag time and apparent rate of aggregation in the fibrillation process

for both pH studied. At pH 2.0, insulin aggregated with a lag time of 137.7 h and apparent aggregation rate was 0.244 s^{-1} . In case of pH 5.3, lag time and aggregate rate were 327.5 h and 0.052 s^{-1} respectively. This suggests that lag time of aggregation almost doubles with increase in pH with significantly slower rate (roughly one-fourth). The driving force of insulin aggregation is the acidic environment that leads to insulin monomer formation followed by rapid aggregation.

Insulin formed diverse aggregate forms in oxidizing (0 mM TCEP) and reducing (10 mM TCEP) environment when incubated for 4 h (Figures 4.8, 4.9). 72 h (Figures 4.10, 4.11) and 7 days (Figures 4.12, 4.13). After 4 h of incubation, all the aggregate forms were amorphous. Morphology of the aggregates changed after 72 h incubation and they were similar to those observed after 7 days. In the absence of TCEP at physiological temperature, insulin formed *rod-like* structures at acidic pH (Figure 4.12 A, B) and amorphous aggregates at pH 8.8 (Figure 4.12 C). At neutral pH (Figure 10 D), no aggregation is observed. At elevated temperature, fibrillation is observed at pH 2.0 and 5.3 (Figure 4.12 E, F), however fibril morphology is different at both pH. At neutral pH (Figure 4.10 G), aggregates formed are amorphous in nature whereas *sponge-like* amorphous aggregates are observed at basic pH (Figure 4.12 H). In the presence of 10 mM TCEP, amorphous aggregates were observed at pH 2.0, 37 °C and at neutral pH for both incubation temperatures (Figure 4.13 A, C, G). However at 65 °C the aggregates were amorphous at pH 2.0 (Figure 4.13 E). At pH 5.3 and 8.8 insulin formed short, branched and long filaments, respectively at both physiological and elevated incubation temperatures (Figure 4.13 B, F, D, H). The stark difference in 4 h aggregates versus the later time points is also conspicuous from the ANS and Bis-ANS

fluorescence data. Increase in fluorescence (ANS, Bis-ANS) for the first 4 h of incubation suggests that these amorphous aggregates are transient and flexible that is formed by exposing the hydrophobic pockets. Gradual decrease of ANS, Bis-ANS fluorescence (Figures 4.2, 4.3) beyond 4 h, signifies reorientation and formation of stable, less flexible structures by hydrophobic collapse and burial. These observations also support earlier studies that amorphous aggregates convert into fibrils at low pH, high temperature and oxidizing environment.

Insulin is known to form fibrils at low pH and high temperature. Previous studies have shown that long-term incubation of insulin converts amorphous aggregates in fibrils. In order to compare the effect of incubation temperature towards fibril formation, insulin was incubated for 4 weeks at 37 and 65 °C in the presence and absence of 10 mM TCEP. The aggregates were characterized using ThT fluorescence and electron microscopy (Figure 4.14). Under reducing conditions at pH 2.0, insulin converted to filaments at 37 °C, and aggregate morphology did not change for 65 °C incubation temperature (Figures 4.14A – c, d). ThT fluorescence intensity for aggregates formed at both temperatures is comparable (Figure 4.14A). In the absence of TCEP, the rod-like filamentous structures converted to fibrils at 37 °C with a significantly high ThT fluorescence. Fibrils formed at 65 °C in 7-days converted into *tightly wound* fibrils after 4 weeks of incubation with ThT fluorescence intensity about ¼ to the fibrils at formed at physiological incubation temperature (Figures 4.14A – a, b). This can be explained in light of the structural differences between the fibrils formed at both temperatures. Fibrils formed after 4-weeks at 37 °C are *loosely wound* hence allowing space for ThT molecules to bind tightly and fluoresce. At 65 °C, the fibrils are tight and

twisted with minimal cavity for binding of ThT leading to low fluorescence and quantum yield. Kaye and coworkers observed thioflavin S negative results in human Alzheimer's disease brain although it was fibril-positive using antibodies, confirming that thioflavin class of dye requires a specific binding cavity into the fibril in order to fluoresce⁵⁷. Our finding contradicts the evidence in literature that insulin forms fibrils at elevated temperature and low pH since we observed *tight* fibrils at 37 °C after 4 weeks of incubation. This suggests that high temperature helps in accelerating the fibril formation and also affects the fibril packing and corresponding ThT fluorescence signal.

In case of pH 5.3, 7.0 and 8.8, aggregate forms did not change significantly and are comparable to those observed after 7 days of incubation (Figures 4.12, 4.13). ThT fluorescence intensity increased significantly in case of samples incubated with and without 10 mM TCEP for at pH 5.3 (65 °C) (Figure 4.13 B) whereas fluorescence was maximum in case of reduced insulin for both pH 7.0 and 8.8 at 65 °C and 37 °C respectively (Figures 4.13 C, D). This suggests that ThT fluorescence is independent of aggregate morphology and largely depends on the size and nature of grooves available in the aggregate for ThT binding. Short filamentous and branched structures do not bind to ThT molecule leading to lower fluorescence intensity as observed in case of insulin aggregates without TCEP at pH 5.3. Similar trend is observed in case of pH 8.8. In the absence of TCEP the aggregates are filamentous however, the larger amorphous aggregates contribute to maximum fluorescence intensity.

Insulin forms diverse aggregate forms when incubated at various pH and temperatures under reducing and non-reducing conditions. It is important to correlate aggregate morphology with its cytotoxicity in order to understand the stresses (pH, temperature, reducing agent) that contribute towards toxic aggregate formation. Insulin aggregates were formed by incubating the protein for 7 days at pH 2.0, 5.3, 7.0 and 8.8 at 37 and 65 °C in the absence and presence of 10 mM TCEP. SH-SY5Y cells were incubated with insulin aggregates for 48 h and the toxicity of the aggregates were assessed using MTS assay. Aggregates formed due to scrambling of disulfide bonds at acidic pH, in the presence of reducing agent were significantly more cytotoxic compared to the aggregates with intact disulfide bonds. Reduction of disulfide bonds and possible formation of non-native disulfide bonds is the driving force towards aggregation in the presence of reducing agent. Scrambled disulfide bonds leads to instability of the protein, which could be a critical factor apart from the pH. Toxic behavior was majorly independent of incubation temperature. Maximum toxicity was observed for aggregates formed at pH 2 (loose amorphous aggregates) followed by pH 5.3 (short filamentous aggregates). This suggests that larger amorphous aggregates can be more toxic to cells since they can sequester the proteins that are critical for maintaining homeostasis in cells. Similar trend was observed in case of inclusion bodies in case of Huntingtin protein aggregation⁵⁸. Longer fibrils/filaments and smaller amorphous aggregates (pH 8.8) are less cytotoxic when compared to shorter filaments since the latter can enter cells owing to its small size and disrupt cellular functions. Insulin aggregates formed at physiological and basic pH exhibited maximum cell viability compared to those at acidic pH. This suggests the insulin aggregates at acidic pH, disrupt cellular function

and confer toxicity in a multitude of ways, however reduced disulfide bonds add to the toxic nature of the aggregates (Figure 4.16).

Figure 4.16 Plausible model of pH-dependent insulin aggregation and toxicity. Insulin aggregates at pH 2.0 and 5.3 are relatively more toxic to SH-SY5Y cells, than those at pH 7.0 and 8.8. Scrambling of disulfide bonds lead to increase in toxicity at acidic pH. Cytotoxicity at neutral/basic pH is comparable for both oxidized and reduced forms of the protein. Surface charge is shown on protein at respective pH using Adaptive Poisson-Boltzmann Solver (ABPS) Software ⁵⁹.

Conclusion

In summary, this study monitors the kinetics and mechanism of insulin aggregation under various external stresses in the form of pH, temperature and reducing environment. Insulin aggregates in the presence of reducing agent within 4 h of incubation at both physiological (37 °C) and elevated (65 °C) temperatures, at all pH studied. In oxidizing environment, insulin aggregation kinetics and morphology is both pH and temperature dependent. Morphology of insulin aggregates were diverse and distinct ranging from amorphous, fibrils to *filamentous* structures. Aggregates formed in the presence of reducing agent at acidic environment were more toxic to human neuroblastoma (SH-SY5Y) cells compared to those formed in the absence of reducing agent, suggesting that pH and disulfide bond status play an important role in toxicity of the aggregated forms.

References

- (1) Ross, C.A.; M.A. Poirier. What is the Role of Protein Aggregation in Neurodegeneration? *Nat. Rev. Mol. Cell Biol.*, **2005**, 6(11), 891-898.
- (2) Ross, C.A.; M.A. Poirier. Protein Aggregation and Neurodegenerative Disease. *Nat Med.*, **2004**, 10 Suppl, S10-7.
- (3) Bucciantini, M.G.; Giannoni, E.; Chiti, Fabrizio.; 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. *Nat. Rev. Drug Discov.*, **2002**, 16, 507-511.
- (4) Tuomas, P. J.; Knowles, M.V.; Dobson, C.M. The Amyloid State and its Association with Protein Misfolding Diseases. *Nat. Rev. Mol. Cell Biol.*, **2014**, 15, 384–396.
- (5) Wang, Q.; Johnson, J.L.; Agar, N.Y.R.; Agar, J.N. Protein Aggregation and Protein Instability Govern Familial Amyotrophic Lateral Sclerosis Patient Survival. *PloS Bio*, **2008**, 6(7), e170.
- (6) Van Raaij, M.E.; Segers-Nolten, I.M.; Subramaniam, V. Quantitative Morphological Analysis Reveals Ultrastructural Diversity of Amyloid Fibrils from a-Synuclein Mutants. *Biophys. J: Biophys. Lett.*, **2006**, 91(11), L96-L98.
- (7) Meinhardt, J.; Sachse, C.; Hortschansky,P.; Grigorieff, N.; Fändrich, M. A β (1-40) Fibril Polymorphism Implies Diverse Interaction Patterns in Amyloid Fibrils. *J. Mol. Biol.*, **2009**, 386(3), 869-877.
- (8) Chiti, F.; Dobson, C.M. Protein Misfolding, Functional Amyloid, and Human Disease. *Ann. Rev.Biochem.* **2006**, 33-366.

- (9) Jin, L.W.; Claborn, K.A.; Kurimoto, M.; Geday, M.A.; Maezawa, I.; Sohraby, F.; Estrada, M.; Kaminsky, E.; Kahr, B. Imaging Linear Birefringence and Dichroism in Cerebral Amyloid Pathologies. *Proc. Natl. Acad. Sci. U S A*, **2003**, 100(26), 15294-8.
- (10) Manuelidis, L.; Fritch, W.; Xi, Y.-G. Evolution of a Strain of CJD That Induces BSE-Like Plaques. *Science*, **1997**, 277, 94-98.
- (11) Sipe, J.D., Amyloidosis. *Ann. Rev. Biochem.*, **1992**, 61, 947-975.
- (12) Snow, A.D.; Sekiguchi, R.; Nochlin, D.; Fraser, P.; Kimata, K.; Mizutani, A.; Arai, M.; Schreier, W.A.; Morgan, D. G. An Important Role of Heparan Sulfate Proteoglycan (Perlecan) in a Model System for the Deposition and Persistence of Fibrillar $\alpha\beta$ -Amyloid in Rat Brain. *Neuron*, **1994**, 12(1), 219-234.
- (13) Exley, C.; House, E.; Collingwood, J.F.; Davidson, M.R.; Cannon, D.; Donald, A.M. Spherulites of Amyloid-Beta42 in Vitro and in Alzheimer's Disease. *J Alzheimers Dis*, **2010**, 20(4), 1159-65.
- (14) Bromley, E.H.C.; Krebs, M.R.H.; Donald, A.M. Aggregation Ccross the Length-Scales in Beta-Lactoglobulin. *Farad. Discuss.*, **2005**, 128, 13-27.
- (15) Raffen, R.; Dieckman, L.J.; Szpunar, M.; Wunschl, C.; Pokkuluri, P.R.; Dave, P.; Wilkins Stevens, P.; Cai, X.; Schiffer, M.; Stevens, F.J. Physicochemical Consequences of Amino Acid Variations that Contribute to Fibril Formation by Immunoglobulin Light Chains. *Prot. Sci.*, **1999**, 8(3), 509-517.
- (16) Westlind-Danielsson, A.; Arnerup, G. Spontaneous In Vitro Formation of Supramolecular Beta-Amyloid Structures, "Beta Amy Balls", by Beta-Amyloid 1-40 Peptide. *Biochem.*, **2001**, 40(49), 14736-14743.

- (17) Sagis, L.; Veerman, C.; Ganzevles, R.; Ramaekers, M.; Bolder, S.; Van der Linden, E. Mesoscopic Structure and Viscoelastic Properties of Beta-Lactoglobulin Gels at Low pH and Low Ionic Strength. *Food Hydrocoll.*, **2002**, 16(3), 207-213.
- (18) Domike, K.R.; Donald, A.M. Thermal Dependence of Thermally Induced Protein Spherulite Formation and Growth: Kinetics of β -lactoglobulin and Insulin. *Biomacromol.*, **2007**, 8, 3930-3937.
- (19) Juarez, J.; Taboada, P.; Mosquera, V. Existence of Different Structural Intermediates on the Fibrillation Pathway of Human Serum Albumin. *Biophys. J.*, **2009**, 96(6), 2353-2370.
- (20) Bolder, S.G.; Hendrickx, H.; Sagis, L.M.C.; Van der Linden, E. Fibril Assemblies in Aqueous Whey Protein Mixtures. *J. Agric. Food Chem.*, **2006**, 54(12), 4229-4234.
- (21) Krebs, M.R.; Domike, K.R.; Donald, A.M. Protein Aggregation: More Than Just Fibrils. *Biochem. Soc. Trans.*, **2009**, 37(Pt 4), 682-6.
- (22) Hoshi, M.; Sato, M.; Matsumoto, S.; Noguchi, A.; Yasutake, K.; Yoshida, N.; Sato, K. Spherical Aggregates of β -Amyloid (Amylospheroid) Show High Neurotoxicity and Activate Tau Protein Kinase I/Glycogen Synthase Kinase-3 β . *PNAS*, **2003**, 100(11), 6370-6375.
- (23) Seilheimer, B.; Bohrmann, B.; Bondolfi, L.; Müller, F.; Stüber, D.; Döbeli, H. The Toxicity of the Alzheimer's β -Amyloid Peptide Correlates with a Distinct Fiber Morphology. *J. Struct. Bio.*, 1997, 119(1), 59-71.
- (24) Julien, J.P. Amyotrophic Lateral Sclerosis: Unfolding the Toxicity of the Misfolded. *Cell*, **2001**, 104(4), 581-591.

- (25) Krebs, M.R.H.; Bromley, E.H.C.; Rogers, S.S.; Donald, A.M. The Mechanism of Amyloid Spherulite Formation by Bovine Insulin. *Biophys. J.*, **2005**, 88(3), 2013-21.
- (26) Krebs, M.R.H.; MacPhee, C.E.; Miller, A.F.; Dunlop, I.E.; Dobson, C.M.; Donald, A.M. The Formation of Spherulites by Amyloid Fibrils of Bovine Insulin. *PNAS*, **2004**, 101(40), 14420-4.
- (27) Rogers, S.S.; Krebs, M.R.H.; Bromley, E.H.C.; Van der Linden, E.; Donald, A.M. Optical Microscopy of Growing Insulin Amyloid Spherulites on Surfaces In Vitro. *Biophys J.*, **2006**, 90(3), 1043-54.
- (28) Smith, M.I.; Foderà, V.; Sharp, J.S.; Roberts, C.J.; Donald, A.M. Factors Affecting the Formation of Insulin Amyloid Spherulites. *Colloids Surf B Biointerfaces*, **2012**, 89, 216-22.
- (29) Yagi, N.; Ohta, N.; Iida, T.; Inoue, K. A Microbeam X-ray Diffraction Study of Insulin Spherulites. *J. Mol. Biol.*, **2006**, 362(2), 327-33.
- (30) Brange, J.; Andersen, L.; Laursen, E.D.; Meyn, G.; Rasmussen, E. Toward Understanding Insulin Fibrillation. *J.Pharm. Sci.*, **1997**, 86(5), 517-525.
- (31) Hua, Q.X. Insulin: A Small Protein with a Long Journey. *Protein & Cell*, **2010**, 1(6), 537-551.
- (32) Sluzky, V.; Tamada, J.A.; Klibanov, A.M.; Langer, R. Kinetics of Insulin Aggregation in Aqueous-Solutions upon Agitation in the Presence of Hydrophobic Surface. *PNAS*, **1991**, 88(21), 9377-9381.
- (33) Ahmad, A.; Millett, I.S.; Doniach, S.; Uversky, V.N.; Fink, A.L. Partially Folded Intermediates in Insulin Fibrillation. *Biochemistry*, **2003**, 42(39), 11404-11416.

- (34) Whittingham, J.L.; Scott, D.J.; Chance, K.; Wilson, A.; Finsh, J.; Brange, J.; Dodson, G.G. Insulin at pH 2: Structural Analysis of the Conditions Promoting Insulin Fibre Formation. *J. Mol. Biol.*, **2002**, 318(2), 479-490.
- (35) Ahmad, A.; Millett, I.S.; Doniach, S.; Uversky, V.N.; Fink, A.L. Stimulation of Insulin Fibrillation by Urea-Induced Intermediates. *J. Biol. Chem.*, **2004**, 279(15), 14999-15013.
- (36) Ahmad, A.; Uversky, V.N.; Hong, D.; Fink, A.L. Early Events in the Fibrillation of Monomeric Insulin. *J. Biol. Chem.*, **2005**, 280(52), 42669-42675.
- (37) Nielsen, L.; Frokajaer, S.; Brange, J.; Uversky, V.N.; Fink, A.L. Probing the Mechanism of Insulin Fibril Formation with Insulin Mutants. *Biochemistry*, **2001**, 40(28), 8397-8409.
- (38) Weiss, M.; Steiner, D.F.; M.D., Philipson, L.H. Insulin Biosynthesis, Secretion, Structure, and Structure-Activity Relationships. *Endocrinol.*, De Groot LJ, Chrousos G, Editor. 2014, MDText.com, Inc: South Dartmouth (MA).
- (39) Hong, D.P.; Ahmad, A.; Fink, A.L. Fibrillation of Human Insulin A and B Chains. *Biochemistry*, **2006**, 45(30), 9342-9353.
- (40) 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(8), 3331-40.
- (41) Kachooei, E.; Moosavi-Movahedi, A.A.; Khodaghali, F.; Ramshini, H.; Shaerzadeh, F.; Sheibani, N. Oligomeric Forms of Insulin Amyloid Aggregation

Disrupt Outgrowth and Complexity of Neuron-Like PC12 Cells. *PLoS One*, **2012**, 7(7), e41344.

(42) Kachooei, E.; Moosavi-Movahedi, A.A.; Khodagholi, F.; Mozaffarian, F.; Sadeghi, P.; Hadi-Alijanvand, H.; Ghasemi, A.; Saboury, A.A.; Farhadi, M.; Sheibani, N. Inhibition Study on Insulin Fibrillation and Cytotoxicity by Paclitaxel. *J. Biochem*, **2014**, 155(6), 361-73.

(43) Anderson, J.L.D., M.M.; Seinfeld, J.H.; Stephanopoulos, G., Wei, J., *Advances in Chemical Engineering: Volume 19*. 1995.

(44) Dutta, C.; Yang, M.; Long, F.; Shahbazian-Yassar, R.; Tiwari, A. Preformed Seeds Modulate Native Insulin Aggregation Kinetics. *J. Phys. Chem. B.*, **2015**. 119(49), 15089–15099.

(45) Yang, M.; Dutta, C.; Tiwari, A. Disulfide-Bond Scrambling Promotes Amorphous Aggregates in Lysozyme and Bovine Serum Albumin. *J. Phys. Chem. B.*, **2015**, **119**(10), 3969-81.

(46) Winsor, C.P.; The Gompertz Curve as a Growth Curve. *PNAS*, **1932**, 18(1), 1-7.

(47) Wang, S.S.; Liu, K.N.; Han, T.C. Amyloid Fibrillation and Cytotoxicity of Insulin are Inhibited by the Amphiphilic Surfactants. *Biochim. Biophys. Acta.*, **2010**, 1802(6), 519-30.

(48) Xu, Y.; Yan, Y.; Seeman, D.; Sun, L.; Dubin, P.L. Multimerization and Aggregation of Native-State Insulin: Effect of Zinc. *Langmuir*, **2012**, 28(1), 579-86.

(49) Chang, X.; Jorgensen, A.M.; Bardrum, P.; Led, J.J. Solution Structures of the R6 Human Insulin Hexamer. *Biochemistry*, **1997**, 36, 9409-9422.

- (50) Gill, R. Protein Engineering of Insulin-like Growth Factor. *Department of Crystallography*. 1994, University of London: Birbeck College.
- (51) de Oliveira, G.A.; Silva, J.L. A Hypothesis to Reconcile the Physical and Chemical Unfolding of Proteins. *PNAS*, **2015**, 112(21), E2775-E2784.
- (52) Groenning, M.; Olsen, L.; Van de Weert, M.; Flink, J.M.; Frokjaer, S.; Jørgensen F.S. Study on the Binding of Thioflavin T to Beta-Sheet-Rich and Non-Beta-Sheet Cavities. *J. Struct. Biol.*, **2007**, 158(3), 358-369.
- (53) Nokihara, K.; Yajima, S.; Hitara, A.; Sogon, T.; Yasuhara, T. Characterization of Peptides Obtained from Digests of Bovine Brain Which Accelerate Structural Conversions of the Recombinant Bovine Prion Protein. *FEBS Lett.*, **2013**, 587(6), 673-6.
- (54) Zhong, Q.; Congdon, E.E.; Nagaraja, H.N.; Kuret, J. Tau Isoform Composition Influences Rate and Extent of Filament Formation. *J. Biol. Chem.*, **2012**, 287(24), 20711-9.
- (55) Sultana, M.J.; Ahmed, F.R.S. Phytochemical Investigations of the Medicinal Plant Swertia Chirata Ham. *Biochem. Anal. Biochem.*, **2013**, 02(145).
- (56) Moore, C.L.; Huang, M.H.; Robbenolt, S.A.; Voss, K.R.; Combs, B.; Gamblin, T.C.; Goux, W.J. Secondary Nucleating Sequences Affect Kinetics and Thermodynamics of Tau Aggregation. *Biochemistry*, **2011**, 50(50), 10876-86.
- (57) Kaye, R.; Head, E.; Sarsoza, F.; Saing, T.; Cotman, C.W.; Necula, M.; Margol, L.; Wu, J.; Breydo, L.; Thompson, J.L.; Rasool, S.; Gurlo, T.; Butler, P.; Glabe, C.G. Fibril Specific, Conformation Dependent Antibodies Recognize a Generic Epitope

Common to Amyloid Fibrils and Fibrillar Oligomers that is Absent in Prefibrillar Oligomers. *Mol. Neurodegener.*, **2007**, 2, 18.

(58) Treusch, S., Cyr, D.M.; Lindquist, S. Amyloid Deposits: Protection Against Toxic Protein Species. *Cell Cycle*, **2009**, 8(11), 1668-1674.

(59) Baker, N.A.; Sept, D.; Joseph, S.; Holst, M. J.; Mc.Cammon, J. A. Electrostatics of nanosystems: application to microtubules and the ribosome. *PNAS*, **2001**, 98(18), 10037-41.

Chapter 5. Protein aggregation in a crowded environment³

Colina Dutta, Rashmi Adhikari, and Ashutosh Tiwari*

Department of Chemistry, Michigan Technological University, Houghton, MI 49931

***Correspondence:** Ashutosh Tiwari, Department of Chemistry, Michigan Technological University, Houghton, MI, USA,

Tel:(906) 487-1840; Fax: (906) 487-2061; Email: tiwari@mtu.edu

³ The material contained in this chapter will be published in ACS Publications.

Abstract

Misfolding and aggregation of proteins have been associated with various neurodegenerative diseases. While we have intensely focused on understanding the factors that govern misfolding and aggregation of proteins *in vitro*; not enough focus has been devoted to understanding how molecular crowding may impact the process. It is important to mimic the protein-misfolding scenario within the crowded cellular environment. Therefore, we have used macromolecular crowding agent polyethylene glycol (PEG) to mimic molecular crowding in cells.

The main aim of this study is to understand the effect of macromolecular crowding agent PEG on lysozyme stability and aggregation in the presence of thermal and reducing environment stress. We investigated lysozyme unfolding between 25-95 °C using temperature dependent intrinsic and ANS fluorescence spectroscopy in the presence of both thiol and non-thiol based reducing agents dithiothreitol (DTT) and tris-(2-carboxyethyl) phosphine (TCEP). Varying concentrations of PEG-600 and PEG-2000 were used as molecular crowding agents. 200 mM of PEG-2000 exhibited stabilizing effect towards lysozyme in the presence of TCEP whereas PEG had no effect in the presence of DTT, irrespective of the molecular weight used in the study.

Introduction

Protein misfolding and aggregation have been linked to several neurodegenerative disorders and amyloid diseases¹. Protein aggregates are majorly found as deposits within the brain, liver, spleen and heart². These aggregated forms have been generated in many *in vitro* studies using both physiological and non-physiological conditions to understand the mechanism of the disease process. These studies are carried out in dilute environment, which is significantly different than the cellular environment. The interior of all cells contains high concentration of macromolecules (proteins, polysaccharides, lipids) that occupy approximately 30% of the cell volume. These macromolecules lead to volume exclusion causing significant thermodynamic and kinetic consequences on the properties of other molecules present in the cell³. Volume exclusion effect leads to specific and non-specific interactions within the cell leading to perturbations/alterations in the biochemical processes within the cell. Crowding is known to favor formation of macromolecular complex and the compact conformations of proteins as opposed to its unfolded state⁴. Hence, it is important to investigate the mechanism of protein folding and aggregation in a crowded environment in order to understand the disease process.

Lysozyme is a widely studied model protein due to its rich phase behavior that includes crystallization, liquid-liquid phase separation and formation of equilibrium clusters/gels⁵. The biophysical characteristics of this 14.7 kDa globular protein has been extensively studied by measuring the average properties of the conformational isomers that exist at thermodynamic equilibrium in denatured lysozyme⁶. Studies with lysozyme have shown that high concentrations of crowding agent (200 g/L) did not

affect refolding of oxidized lysozyme but disrupted the refolding of protein. Denatured lysozyme with intact disulfide bonds can be renatured efficiently by rapid dilution whereas renaturation of the reduced protein by dilution is inefficient unless the protein concentration is significantly low. Two-state unfolding mechanism of lysozyme was studied using disulfide-scrambled lysozyme⁷. However, very little is known how molecular crowding agents impact lysozyme stability and aggregation in the presence of disulfide reducing agents. The general mechanism of protein misfolding and aggregation of disulfide reduced protein is by scrambling of disulfide bonds^{8,9}. Hence, it is important to understand how crowding agents can impact protein aggregation under reducing conditions. For this we used lysozyme as a model protein as its aggregation properties in dilute solutions have been extensively studied^{6,7}.

A large number of molecules are known to provide volume exclusion effect in *in vitro* studies¹⁰. At high concentrations, these molecules preferentially hydrate the protein surface, thereby favoring the compact state of the protein¹¹. These molecules include amino acids, salts, sugars and polymers like polyethylene glycols (PEG). The mechanism of action in case of most amino acids, sugar and salt is through cohesive force of water that is known to stabilize the native structure of the protein. However, PEG is known to exhibit both stabilizing and destabilizing effects on the protein¹². Multiple theories have been proposed to explain the mechanism of action of PEG on protein¹³. Preferential hydration in the presence of the polymer has been hypothesized to be due to steric exclusion as well as charge on the protein¹⁴. Steric exclusion arises due to the difference in size of the water molecule and that of the additives/crowding agents. The large size of crowding agents such as PEG is key in its exclusion from the

protein surface by steric exclusion mechanism¹¹. The effect of PEG on protein thermal stability has been shown to be molecular weight and concentration dependent¹⁵. However, how does PEG impact stability of disulfide reduced proteins has not been studied.

The main aim of this study is to investigate the thermal unfolding of lysozyme in the presence of PEG and disulfide reducing agents dithiothreitol (DTT) and tris (2-carboxyethyl) phosphine (TCEP). DTT is a thiol based reducing agent unlike TCEP. Thermal unfolding of lysozyme was monitored between 25 °C to 95 °C in the presence of varying concentration of reducing agents and PEG in order to understand the effect of concentration on thermal unfolding of lysozyme. Effect of molecular weight of PEG on lysozyme unfolding was also studied using a low (600 g/mol) and medium (2000 g/mol) molecular weight of the polymer.

Materials and Methods

Unless otherwise indicated, all materials were used as supplied by the manufacturer without any further purification. TCEP was purchased from Thermo Scientific Pierce; lysozyme, DTT, 8-anilino-1-naphthalenesulfonic acid (ANS) dye, PEG-600 and PEG-2000 were purchased from Sigma.

Preparation of Lysozyme Samples - Lysozyme stock was freshly prepared by dissolving the lyophilized protein in 20 mM 4-(2 hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer at pH 7.0. Lysozyme concentration was determined using UV-Visible spectroscopy using extinction coefficient $\epsilon_{280\text{ nm}} = 38940$

M⁻¹ cm⁻¹. Working protein solution contained 40 μM protein in 20 mM HEPES buffer at pH 7.0.

Sample Preparation and Heat Treatments - Samples were prepared by incubating lysozyme with varying concentrations of reducing agents DTT (2 and 5 mM) and TCEP (0.5 and 1 mM), along with 20 and 200 mM of PEG-600 and PEG-2000. Respective controls without PEG and reducing agents were also incubated.

For heat treatment, each sample was heated to a fixed temperature (25-95 °C, 5°C intervals) for exactly 12.5 min and cooled on ice for 10 min¹⁶.

Intrinsic and extrinsic fluorescence spectroscopy - Fluorescence measurements were carried out on a Horiba Jobin Yvon spectrofluorometer (Fluoromax-4) at room temperature. Slit widths of 2 nm (entrance and exit) were used for all measurements. After heat treatments, intrinsic fluorescence spectra of the samples were collected in the 285-450 nm range after exciting at 280 nm.

For extrinsic fluorescence measurements, heat treated samples were incubated with 5 μM ANS dye for 15 min on ice. ANS stock solution was prepared in 100% ethanol. Concentration of the dye was determined using extinction coefficient $\epsilon_{350\text{ nm}} = 5000\text{ M}^{-1}\text{ cm}^{-1}$. Fluorescence spectra were collected between 400-650 nm after excitation at 380 nm.

Reducing and Non-Reducing Gel Electrophoresis - Heat treated protein samples (40μM) at 50 °C and 95 °C were mixed with fresh prepared 100 mM iodoacetamide (IAA) and incubated at room temperature for 2 hours to facilitate the blocking of free thiol groups. The reaction was terminated by adding sodium dodecyl sulfate (SDS) sample buffer containing no reducing agents and boiled for 3 minutes. The reduced

protein samples were prepared by adding SDS sample buffer containing 5% β -mercaptoethanol to protein samples (40uM) and was boiled for 3 minutes. All samples (10 μ g/lane) were run on 15% Tris-HCl polyacrylamide gel using Tris-glycine-SDS buffer (25 mM Tris, 192 mM glycine, 0.1% SDS, pH 8.3; Bio-Rad) for 3.5 hours at 80 V. The gels were stained using Coomassie blue stain and image was acquired using a scanner.

Field-Emission Scanning Electron Microscopy (FESEM) - Samples were prepared as described earlier. Briefly, heat treated samples were washed with MQ water and aliquoted on SEM stubs and dried at room temperature. The samples were coated with 10 nm platinum using a sputter coater. Images were acquired using an acceleration voltage of 10 kV.

Results & Discussion

Lysozyme is a small globular protein with 4 disulfide bonds at positions C6-C127, C30-C115, C64-C80 and C76-C94¹⁷. It is majorly α -helical protein with some β -sheet structure and a loop. The disulfide bridge between C6-C127 connects the N- and C-terminal of the protein and is partially solvent accessible, whereas other disulfide bonds are buried into the hydrophobic pocket of the protein. Under reducing conditions, lysozyme unfolds leading to formation of a random-coiled polypeptide. The protein exhibits two-stage thermal denaturation in the presence of chaotropic agents, thus serving as a model protein for studying thermal unfolding and effect of macromolecular crowding⁷. However, the denaturation process of lysozyme in the presence of reducing agent and macromolecular crowding agent in combination with thermal stress is unknown. Hence it is important to understand the unfolding of lysozyme with thermal stress in the presence of both thiol and non-thiol based reducing agents DTT and TCEP respectively. The mechanism of action of both reducing agents is distinct and it is important to investigate their effect on lysozyme denaturation in a macromolecular crowding environment¹⁸.

We monitored thermal unfolding of lysozyme between 25 to 95 °C in the presence of 20 and 200 mM of both PEG-600 and PEG-2000. Disulfide bond reduction of lysozyme was carried out using varying concentrations of both thiol and non-thiol based reducing agents DTT (2, 5 mM) and TCEP (0.5, 1 mM), respectively. Thermal unfolding of the protein was monitored using intrinsic fluorescence spectroscopy (Figures 5.1, 5.2)^{14, 17}. Conformational change and hydrophobicity was measured using ANS fluorescence (Figures 5.3, 5.4). Reducing gel electrophoresis (SDS-PAGE) was used to determine

whether the protein was completely reduced due to the heat treatment (Figure 5.5). Morphology of the aggregates in the presence of disulfide reducing agents and PEG was monitored using scanning electron microscopy (FESEM) (Figure 5.6).

Lysozyme like other proteins, contain tryptophan residues that absorb UV radiation and can be excited. The electronic transition back to the ground state is associated with photon emission observed as intrinsic fluorescence. This signal quenched in aqueous environment¹⁹. Unfolding of lysozyme due to thermal stress in the presence of TCEP and DTT was monitored. In the absence of TCEP and DTT, intrinsic fluorescence is not observed even in the absence and presence of 20 and 200 mM (Figures 5.1, 5.2). In the presence of 0.5 and 1 mM TCEP, intrinsic fluorescence is observed in the presence of PEG (Figure 5.1). Earlier studies have shown the tryptophan fluorescence of proteins get affected in the presence of PEG. Fluorescence is not observed in case of native proteins, which is in line with our observation of no fluorescence intensity in the absence of TCEP. However, PEG is known to bind to the molten globule form of proteins²⁰. Presence of 0.5 and 1 mM TCEP can partially unfold lysozyme, leading to the formation of molten globule structure, thus facilitating PEG binding with lysozyme and hence increases fluorescence²⁰. In the presence of 200 mM PEG-600 and PEG-2000 fluorescence intensity decreases suggesting stabilization of the protein structure by PEG binding. Previous studies have shown that PEG bound to protein can either be stabilizing or destabilizing depending on the size of PEG¹⁰. We observe similar trend where PEG-600 does not have any significant stabilizing effect whereas the higher molecular weight form of PEG has a greater stabilizing effect on the protein.

In the presence of both 2 and 5 mM DTT, intrinsic fluorescence is observed for both PEG-600 and PEG-2000 (Figure 5.2). This suggests that the mechanism of disulfide reduction plays an important role in the thermal unfolding of lysozyme. DTT reduced disulfide bonds are known to oxidize with time and form non-native disulfide bridges, whereas reduction of disulfide bonds is an irreversible process in the presence of non-thiol based reducing agent TCEP¹⁸. This is in line with an earlier study where authors observed that thiol based reducing agent inhibited lysozyme refolding in the presence of guanidine hydrochloride, whereas the oxidized form of the protein refolded.

Figure 5.1. Effect of PEG on the thermal unfolding of lysozyme in the presence of TCEP, monitored by intrinsic fluorescence spectroscopy. Fluorescence spectra of thermal unfolding of 40 μ M lysozyme from 25 $^{\circ}$ C to 95 $^{\circ}$ C was measured in the presence of (A) 0.5 mM TCEP, PEG-600 (B) 0.5 mM TCEP, PEG-2000 (C) 1 mM TCEP, PEG-600 and (D) 1 mM TCEP, PEG-2000. Samples were excited at 280 nm and emission spectra were collected between 285-450 nm. Peak wavelength of 365 nm was chosen based on the difference spectra.

Figure 5.2. Effect of PEG on the thermal unfolding of lysozyme in the presence of DTT, monitored by intrinsic fluorescence spectroscopy. Fluorescence spectra of thermal unfolding of 40 μ M lysozyme from 25 $^{\circ}$ C to 95 $^{\circ}$ C was measured in the presence of (A) 2 mM DTT, PEG-600 (B) 2 mM DTT, PEG-2000 (C) 5 mM DTT, PEG-600 and (D) 5 mM DTT, PEG-2000. Samples were excited at 280 nm and emission spectra was collected between 285-450 nm. Peak wavelength of 365 nm was chosen based on the difference spectra.

Exposure of hydrophobic pockets due to thermal unfolding of lysozyme was monitored using ANS dye (Figures 5.3, 5.4). Presence of PEG-600 did not affect the hydrophobicity of lysozyme when compared with samples with PEG and 0.5 mM TCEP (Figure 5.3 A). However, 200 mM of PEG-2000, lead to decrease in ANS fluorescence in the presence of 0.5 mM TCEP signifying decreased hydrophobicity.

Significant decrease in hydrophobicity and consecutive ANS fluorescence was observed in the presence of 1 mM TCEP and 200 mM PEG-2000 (Figure 5.3 C) whereas subtle decrease in fluorescence was observed with PEG-600 and 1 mM TCEP (Figure 5.3 D). Control groups without TCEP and increasing concentrations of PEG-600 and PEG-2000 did not show any fluorescence due to lack of protein unfolding and exposure of hydrophobic pockets. This is in line with the intrinsic fluorescence data where significant lysozyme stabilization was observed in the presence of 1 mM TCEP and 200 mM PEG-2000.

ANS fluorescence results with DTT are in line with intrinsic fluorescence and no change is observed in the presence of PEG with 2 and 5 mM DTT.

Figure 5.3. Effect of PEG on the hydrophobicity of lysozyme in the presence of TCEP, monitored by ANS fluorescence spectroscopy. Fluorescence spectra of thermal unfolding of 40 μ M lysozyme from 25 $^{\circ}$ C to 95 $^{\circ}$ C was measured in the presence of (A) 0.5 mM TCEP, PEG-600 (B) 0.5 mM TCEP, PEG-2000 (C) 1 mM TCEP, PEG-600 and (D) 1 mM TCEP, PEG-2000. Samples were excited at 380 nm and emission spectra was collected between 400-650 nm. Peak wavelength of 365 nm was chosen based on the difference spectra.

Figure 5.4. Effect of PEG on the thermal unfolding of lysozyme in the presence of DTT, monitored by ANS fluorescence spectroscopy. Fluorescence spectra of thermal unfolding of 40 μ M lysozyme from 25 $^{\circ}$ C to 95 $^{\circ}$ C was measured in the presence of (A) 2 mM DTT, PEG-600 (B) 2 mM DTT, PEG-2000 (C) 5 mM DTT, PEG-600 and (D) 5 mM DTT, PEG-2000. Samples were excited at 380 nm and emission spectra were collected between 400-650 nm. Peak wavelength of 365 nm was chosen based on the difference spectra.

Non-reducing SDS PAGE of heat-treated samples at 50 $^{\circ}$ C and 95 $^{\circ}$ C show formation of dimeric species at 50 $^{\circ}$ C in the presence of 1 mM TCEP whereas at 95 $^{\circ}$ C, dimers are visible in the presence of DTT only (Figure 5.5). This suggests the protective role of 1 mM TCEP towards lysozyme aggregation as the protein unfolds due to increase in

heat treatment temperature. In line with fluorescence study, lysozyme aggregates in the presence of DTT and PEG does not affect the aggregation process.

Morphology of heat-treated lysozyme samples was visualized using scanning electron microscope. Lysozyme samples that were exposed to 50 °C did not show any aggregate formation under the microscope (data not shown). However, a higher molecular weight (28 kDa) is visible in the non-reducing gel. This suggests the formation of lysozyme dimer due to intermolecular interaction but not visible aggregates under the scanning electron microscope. Samples heat-treated to 95 °C show amorphous aggregate formation both in the presence of TCEP and DTT.

Figure 5.5. SDS Gel Electrophoresis of heat-treated lysozyme samples. Lysozyme (10 µg/lane) was loaded on 15% gel. The gels were run for 3.5 hours at 80 V followed by staining with Coomassie blue stain. Reducing SDS-PAGE was run for heated treated lysozyme at (A) 50 °C and (B) 95 °C. Samples were reduced using 5% β-mercaptoethanol and boiled for 3 minutes. High molecular weight species were visualized using non-reducing SDS-PAGE for heat-treated lysozyme at (C) 50 °C and (D) 95 °C, after blocking the free sulfhydryl groups with 100 mM iodoacetamide.

Figure 5.6. Scanning electron micrographs of lysozyme aggregates. Heat-treated lysozyme samples were imaged using SEM. Lysozyme was heat treated at 95 °C with (A) 1 mM TCEP, 200 mM PEG-600 (B) 1 mM TCEP, 200 mM PEG-2000 (C) 2 mM DTT, 200 mM PEG-600 (D) 2 mM DTT, 200 mM PEG-2000

Conclusion

In summary, the role of macromolecular crowding agent PEG in lysozyme unfolding was investigated. The protein was exposed to a range of thermal stress between 25 to 95 °C, in the presence of thiol and non-thiol based reducing agents DTT and TCEP respectively. In the presence of TCEP, lower concentrations of PEG-600 and PEG-2000 (20 mM) did not have any effect on lysozyme, however 200 mM of PEG had a stabilizing effect on the protein. Stabilization was monitored using temperature-dependent intrinsic and extrinsic fluorescence and it was directly related to the molecular weight of the macromolecular crowding agent. Lysozyme in the presence of DTT did not have any effect of PEG and aggregation of the protein was observed at higher temperatures. This suggests that the mechanism of action of the reducing agent plays an important role towards the efficacy of PEG in stabilizing the native structure of lysozyme.

References

- (1) Tokuriki, N.; Kinjo, M.; Negi, S.; Hoshino, M.; Goto, Y.; Urabe, I.; Yomo, T. Protein folding by the effects of macromolecular crowding. *Protein Sci* **2004**, *13* (1), 125-33.
- (2) White, D. A. Buell., A. K.; Knowles, T. P. J.; Welland, M. E. Protein Aggregation in Crowded Environments. *J. Am. Chem. Soc.* **2010**, *132*, 5170-5175.
- (3) Uversky, V. N.; Cooper, E. M.; Bower, K. S.; Li, J.; Fink, A. L. Accelerated alpha-synuclein fibrillation in crowded milieu. *FEBS Lett.* **2002**, *515* (1-3), 99-103.
- (4) Munishkina, L. A.; Cooper, E. M.; Uversky, V. N.; Fink, A. L. The effect of macromolecular crowding on protein aggregation and amyloid fibril formation. *J. Mol. Recognit.* **2004**, *17*, 456-464.
- (5) Rozema, D.; Gellman, S. H. Artificial chaperone-assisted refolding of denatured-reduced lysozyme: Modulation of the competition between renaturation and aggregation. *Biochemistry* **1996**, *35* (49), 15760-15771.
- (6) James, S.; McManus, J. J. Thermal and solution stability of lysozyme in the presence of sucrose, glucose, and trehalose. *J. Phys. Chem. B* **2012**, *116* (34), 10182-8.
- (7) Chang, J. Y.; Li, L. The unfolding mechanism and the disulfide structures of denatured lysozyme. *FEBS Lett.* **2002**, *511* (1-3), 73-78.
- (8) Dutta, C.; Yang, M.; Long, F.; Shahbazian-Yassar, R.; Tiwari, A. Preformed Seeds Modulate Native Insulin Aggregation Kinetics. *J. Phys. Chem. B.*, **2015**. 119(49), 15089–15099.

- (9) Yang, M.; Dutta, C.; Tiwari, A. Disulfide-Bond Scrambling Promotes Amorphous Aggregates in Lysozyme and Bovine Serum Albumin. *J. Phys. Chem. B.*, **2015**, **119**(10), 3969-81.
- (10) Atha, D. H.; Ingham, K. C. Mechanism of Precipitation of Proteins by Polyethylene Glycols - Analysis in Terms of Excluded Volume. *J. Biol. Chem.* **1981**, **256** (23), 2108-2117.
- (11) Bhat, R.; Timasheff, S. N. Steric Exclusion is the Principal Source of the Preferential Hydration of Proteins in the Presence of Polyethylene Glycols. *Prot. Sci.* **1992**, **1** (9), 1133-1143.
- (12) Benton, L. A.; Smith, A. E.; Young, G. B.; Pielak, G. J. Unexpected Effects of Macromolecular Crowding on Protein Stability. *Biochemistry* **2012**, **51** (49), 9773-5.
- (13) Cleland, J. L. Builder. S. E.; Swartz, J. R.; Winkler, M.; Cheng, J. Y.; Wang, D. I. C. Polyethylene Glycol Enhanced Protein Refolding. *Nature* **1992**, **10**, 1013-1019.
- (14) Miklos, A. C.; Sarkar, M.; Wang, Y.; Pielak, G. J. Protein Crowding Tunes Protein Stability. *J Am Chem Soc* **2011**, **133** (18), 7116-20.
- (15) Wang, Y.; Sarkar, M.; Smith, A. E.; Krois, A. S.; Pielak, G. J. Macromolecular Crowding and Protein Stability. *JACS* **2012**, **134**, 16614-16618.
- (16) Manderson, G. A.; Hardman, M. J.; Creamer, L. K. Effect of Heat treatment on Bovine Beta-Lactoglobulin A, B, and C Explored using Thiol Availability and Fluorescence. *J. Agric. Food Chem.* **1999**, **47** (9), 3617-3627.
- (17) Getz, E. B.; Xiao, M.; Chakrabarty, T.; Cooke, R.; Selvin, P. R. A Comparison between the Sulfhydryl Reductants tris(2-carboxyethyl)phosphine and Dithiothreitol for use in Protein Biochemistry. *Anal. Biochem.* **1999**, **273** (1), 73-80.

- (18) Mills, O. E. Effect of Temperature on Tryptophan Fluorescence of Beta-Lactoglobulin B. *Biochim. et biophys. acta* **1976**, 434 (2), 324-332.
- (19) Cleland, J. L.; Randolph, T. W. Mechanism of Polyethylene-Glycol Interaction with the Molten Globule Folding Intermediate of Bovine Carbonic Anhydrase-B. *J. Biol. Chem.* **1992**, 267 (5), 3147-3153.
- (20) Wu, J. Z., C.; Lin, W.; Hu, R.; Wang, Q.; Chen, H.; Li, L.; Chen, S.; Zheng, J. Binding Characteristics between Polyethylene Glycol (PEG) and Proteins in Aqueous Solution. *J. Mat. Chem. B* **2014**, 2, 2983-2992.

Chapter 6. Conclusion and Future Directions

The broad aim of this work is to investigate the aggregation kinetics, mechanism and morphology of model protein insulin, in the presence and absence of disulfide-reducing agents and its relation to cellular toxicity. Elucidating protein aggregation mechanism and its implications for cellular toxicity is important in order to understand neurodegenerative disease process that are primarily due to protein misfolding.

The main aim of the first project was to investigate the aggregation kinetics and mechanism of insulin in the presence of dithiothreitol (DTT) – a thiol based reducing agent, using various spectroscopic techniques (Chapter 3). Morphology of the aggregates were characterized using scanning electron microscope. Nucleation dependent aggregation of insulin in the presence of preformed *seeds* were monitored using fluorescence spectroscopy and atomic force microscopy. Insulin formed amorphous aggregates in the absence of *seeds* whereas in the presence of *seeds*, transient fibril formation was observed that converted to amorphous forms within 24 hours.

Based on the findings from the first project, subsequent work involved correlation of insulin aggregates with their cytotoxicity using human neuroblastoma (SH-SY5Y) cells (Chapter 4). Diverse forms of insulin aggregates were prepared in the presence and absence of tris (2-carboxyethyl) phosphine (TCEP) – a non-thiol based reducing agent, using a variety of pH and temperature conditions. Insulin aggregates were prepared at acidic, neutral and basic pH at physiological (37 °C) and elevated (65 °C) temperature.

Cell viability in the presence of insulin aggregates with reduced disulfide bonds were much lower compared to their oxidized counterpart suggesting the importance of disulfide bonds in native protein fold.

The objective behind studying protein aggregation kinetics and its mechanism is to identify factors that influence protein aggregation. The goal is to characterize these distinct aggregate forms from the same protein under different condition and to identify the toxic fold/shape.

As cellular environment is highly crowded we wanted to understand how these destabilizing factors (disulfide reducing conditions, temperature) can influence protein stability and aggregation. Macromolecular crowding agents have been known to stabilize proteins against thermal stresses and against denaturants. The third part of the work involves understanding the protective role of a crowding agent polyethylene glycol (PEG) against lysozyme aggregation in the presence of thermal and reducing stress (Chapter 5). The role of concentration and molecular weight of PEG towards lysozyme aggregation at a range of temperature (25-95 °C) and reducing stress was studied. PEG was found to prevent lysozyme aggregation in a molecular weight and concentration dependent manner.

The mechanism of protein aggregation in neurodegenerative diseases has been shown to be generic in nature. Hence, information obtained from study of model proteins can be extrapolated to disease causing proteins such as TAR DNA binding protein (TDP-43), Amyloid-beta protein ($A\beta$ -42), Superoxide dismutase 1 (SOD1) etc. Characterizing various aggregate forms of model proteins, and correlating these

aggregates to their toxicity is important in order to understand neurodegenerative disease process. However, the scarcity of information about the disease related proteins increase the complexity. So, it will be important to study other model proteins and also characterize disease related proteins under similar experimental conditions in order to understand how protein misfolding and aggregation leads to these late onset neurodegenerative diseases.

Appendix

Permission for Figures 1.1, 1.2 and 2.12

Permission for Figure 1.3

Trends in biochemical sciences

Order detail ID: 69619593

Order License Id: 3811730961175

ISSN: 0968-0004

Publication Type: Journal

Volume:

Issue:

Start page:

Publisher: ELSEVIER LTD.

Author/Editor: International Union of Biochemistry

Permission Status: **Granted**

Permission type: Republish or display content

Type of use: Thesis/Dissertation

[View details](#)

Note: This item will be invoiced or charged separately through CCC's **RightsLink** service. [More info](#)

\$ 0.00

Permission for Figure 1.4

RightsLink®

[Home](#)
[Account Info](#)
[Help](#)
[Live Chat](#)

Title: Physical Stability of Proteins in Aqueous Solution: Mechanism and Driving Forces in Nonnative Protein Aggregation

Author: Eva Y. Chi

Publication: Pharmaceutical Research

Publisher: Springer

Date: Jan 1, 2003

Copyright © 2003, Plenum Publishing Corporation

Logged in as:
Colina Dutta
Account #: 3001000908

LOGOUT

Order Completed

Thank you very much for your order.

This is a License Agreement between Colina Dutta ("You") and Springer ("Springer"). The license consists of your order details, the terms and conditions provided by Springer, and the [payment terms and conditions](#).

[Get the printable license.](#)

License Number	3811990346661
License date	Feb 18, 2016
Licensed content publisher	Springer
Licensed content publication	Pharmaceutical Research
Licensed content title	Physical Stability of Proteins in Aqueous Solution: Mechanism and Driving Forces in Nonnative Protein Aggregation
Licensed content author	Eva Y. Chi
Licensed content date	Jan 1, 2003
Volume number	20
Issue number	9
Type of Use	Thesis/Dissertation
Portion	Figures/tables/illustrations
Number of figures/tables/illustrations	1
Author of this Springer article	No
Original figure numbers	1
Title of your thesis / dissertation	Effect of Disulfide Bond Scrambling on Protein Stability, Aggregation, and Cytotoxicity
Expected completion date	Apr 2016
Estimated size(pages)	200
Total	0.00 USD

Permission for Figure 1.5

Creative Commons Attribution License (CC BY)

This article is available under the terms of the [Creative Commons Attribution License \(CC BY\)](#). You may distribute and copy the article, create extracts, abstracts, and other revised versions, adaptations or derivative works of or from an article (such as a translation), to include in a collective work (such as an anthology), to text or data mine the article, including for commercial purposes without permission from Elsevier. The original work must always be appropriately credited.

Permission is not required for this type of reuse.

CLOSE WINDOW

Copyright © 2016 [Copyright Clearance Center, Inc.](#) All Rights Reserved.
Comments? We would like to hear from you. E-mail us at customercare@copyright.com

Permission for Chapter 3

[Home](#)[Create Account](#)[Help](#) Live Chat

ACS Publications
Most Trusted. Most Cited. Most Read.

Title: Preformed Seeds Modulate Native Insulin Aggregation Kinetics

Author: Colina Dutta, Mu Yang, Fei Long, et al

Publication: The Journal of Physical Chemistry B

Publisher: American Chemical Society

Date: Nov 1, 2015

Copyright © 2015, American Chemical Society

LOGIN

If you're a **copyright.com** user, you can login to RightsLink using your copyright.com credentials. Already a **RightsLink** user or want to [learn more?](#)

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.