STARS

University of Central Florida
STARS

Electronic Theses and Dissertations, 2004-2019

2011

Self-assembly Of Amyloid Aggregates Simulated With Molecular Dynamics

Workalemahu Mikre Berhanu
University of Central Florida

 Part of the Chemistry Commons
Find similar works at: https://stars.library.ucf.edu/etd
University of Central Florida Libraries http://library.ucf.edu

This Doctoral Dissertation (Open Access) is brought to you for free and open access by STARS. It has been accepted for inclusion in Electronic Theses and Dissertations, 2004-2019 by an authorized administrator of STARS. For more information, please contact STARS@ucf.edu.

STARS Citation

Berhanu, Workalemahu Mikre, "Self-assembly Of Amyloid Aggregates Simulated With Molecular Dynamics" (2011). *Electronic Theses and Dissertations, 2004-2019*. 1826.
https://stars.library.ucf.edu/etd/1826

SELF-ASSEMBLY OF AMYLOID AGGREGATES SIMULATED WITH MOLECULAR DYNAMICS

by

WORKALEMAHU MIKRE BERHANU

M.Sc. Addis Ababa University, Ethiopia. 2005

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of Chemistry
in the College of Sciences
at the University of Central Florida
Orlando, Florida

Fall Term

2011

Major Professor: Artëm E. Masunov

© 2011 Workalemahu Mikre Berhanu

ABSTRACT

Amyloids are highly ordered cross-β sheet aggregates that are associated with many diseases such as Alzheimer's, type II diabetes and prion diseases. Recently a progress has been made in structure elucidation, environmental effects and thermodynamic properties of amyloid aggregates. However, detailed understanding of how mutation, packing polymorphism and small organic molecules influence amyloid structure and dynamics is still lacking. Atomistic modeling of these phenomena with molecular dynamics (MD) simulations holds a great promise to bridge this gap. This Thesis describes the results of MD simulations, which provide insight into the effects of mutation, packing polymorphism and molecular inhibitors on amyloid peptides aggregation. Chapter 1 discusses the structure of amyloid peptides, diseases associated with amyloid aggregation, mechanism of aggregation and strategies to treat amyloid diseases. Chapter 2 describes the basic principles of molecular dynamic simulation and methods of trajectory analysis used in the Thesis. Chapter 3 presents the results of the study of several all-atom molecular dynamics simulations with explicit solvent, starting from the crystalline fragments of two to ten monomers each. Three different hexapeptides and their analogs produced with single glycine replacement were investigated to study the structural stability, aggregation behavior and thermodynamics of the amyloid oligomers. Chapter 4 presents multiple molecular dynamics (MD) simulation of a pair polymorphic form of five short segments of amyloid peptide. Chapter 5 describes MD study of single-layer oligomers of the full-length insulin with a goal to identify the structural elements that are important for insulin amyloid stability, and to suggest single glycine mutants that may improve formulation. Chapter 6 presents the investigation of the mechanism of the interaction of polyphenols molecules with the protofibrils formed by an amyloidogenic hexapeptide fragment (VQIVYK) of Tau peptide by molecular dynamics

simulations in explicit solvent. We analyzed the trajectories of the large (7×4) aggregate with and without the polyphenols.

Our MD simulations for both the short and full length amyloids revealed adding strands enhances the internal stability of wildtype aggregates. The degree of structural similarity between the oligomers in simulation and the fibril models constructed based on experimental data may explain why adding oligomers shortens the experimentally observed nucleation lag phase of amyloid aggregation. The MM-PBSA free energy calculation revealed nonpolar components of the free energy is more favorable while electrostatic solvation is unfavorable for the sheet to sheet interaction. This explains the acceleration of aggregation by adding nonpolar co-solvents (methanol, trifluoroethanol, and hexafluoroisopropanol). Free energy decomposition shows residues situated at the interface were found to make favorable contribution to the peptide -peptide association.

The results from the simulations might provide both the valuable insight for amyloid aggregation as well as assist in inhibitor design efforts. First, the simulation of the single glycine mutants at the steric zipper of the short segments of various pathological peptides indicates the intersheet steric zipper is important for amyloid stability. Mutation of the side chains at the dry steric zipper disrupts the sheet to sheet packing, making the aggregation unstable. Thus, designing new peptidomimetic inhibitors able to prevent the fibril formation based on the steric zipper motif of the oligomers, similar to the ones examined in this study may become a viable therapeutic strategy. The various steric zipper microcrystal structures of short amyloid segments could be used as a template to design aggregation inhibitor that can block growth of the aggregates. Modification of the steric zipper structure (structure based design) with a single amino acid changes, shuffling the sequences, N- methylation of peptide amide bonds to suppress hydrogen

bonding ability of NH groups or replacement with D amino acid sequence that interact with the parent steric zipper could be used in computational search for the new inhibitors.

Second, the polyphenols were found to interact with performed oligomer through hydrogen bonding and induce conformational change creating an altered aggregate. The conformational change disrupts the intermolecular amyloid contact remodeling the amyloid aggregate. The recently reported microcrystal structure of short segments of amyloid peptides with small organic molecules could serve as a pharamcophore for virtual screening of aggregation inhibitor using combined docking and MD simulation with possible enhancement of lead enrichment.

Finally, our MD simulation of short segments of amyloids with steric zipper polymorphism showed the stability depends on both sequence and packing arrangements. The hydrophilic polar GNNQQNY and NNQNTF with interface containing large polar and/or aromatic side chains (Q/N) are more stable than steric zipper interfaces made of small or hydrophobic residues (SSTNVG, VQIVYK, and MVGGVV). The larger sheet to sheet interface of the dry steric zipper through polar Q/N rich side chains was found to holds the sheets together better than non Q/N rich short amyloid segments. The packing polymorphism could influence the structure based design of aggregation inhibitor and a combination of different aggregation inhibitors might be required to bind to various morphologic forms of the amyloid peptides.

ACKNOWLEDGMENTS

I would like to express my sincerest gratitude to my mentor and research advisor, Prof. Artëm E Masunov for his guidance, mentoring and support, throughout my research work. I appreciate his countious encouragement with my area of research interest and in assisting me in doing so. I thank him for all the help he gave me throughout my study.

I am grateful to my committee members Dr. Dmitry Kolpashchikov, Dr Alfons Schulte, Dr. Jingdong Ye, Dr. Shengli Zou for their detailed review, constructive criticism and valuable advices during my candidacy and dissertation defense. I would like to thank all the former and present members of Dr Masunov group for their advices and the fruitful discussions.

I gratefully acknowledge the NSTC Graduate Fellowship bestwed by the UCF NanoScience Technology Center and its continuing funding support I received. The use of computational facilities at the National Energy Research Scientific Center (NERSC) is also gratefully acknowledged.

Finally, I would like to thank my family and friends, especially my wife, Guenet, for her unconditional support and love

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AMBER	Assisted Model Building with Energy Refinement
DSSP	Dictionary of protein secondary structure
GAFF	Generalized AMBER force field
H-bond	Hydrogen bond
MD	Molecular dynamics
MM	Molecular mechanics
MM-GBSA	Molecular mechanics-Generalized Born Surface Area
MM-PBSA	Molecular mechanics-Poisson Bolzmann Surface Area
NMA	Normal mode analysis
PDB	Protein data bank
PME	Particle Mesh Ewald
RESP	Restrained electrostatic potential
RMSD	Root mean square deviation
RMSF	Root mean square fluctation
SASA	Solvent accessible surface area
TIP3P	Transferable intermolecular potential 3 points
vdw	van der Waals

CHAPTER 1 INTRODUCTION

1.1 Protein structure and function

Proteins are the major macromolecular component of the cell that are synthesized in the ribosomes and are responsible for most of the cell's biological activity. Proteins play a variety of roles in the cell, including structural (cytoskeleton), mechanical (muscle), biochemical (enzymes), and cell signaling (hormones). Proteins consist of a linear polycondensate of amino acids linked together by peptide bonds in a specific sequence. This specific sequence is responsible for the protein's structure and function. A peptide bond formed between two amino acids is the primary element of peptide and protein structure. Protein structure has four main organizational levels: primary, secondary, tertiary and quaternary.[1]

Primary structure of proteins and peptides is defined by the linear sequence of amino acids. There are 20 naturally occurring amino acids, which differ in size, shape, charge, hydrogen bonding ability, hydrophilicity, and chemical reactivity. The sequence of the amino acid in the primary structure is unique to each protein. A one-letter or a three-letter abbreviations are used to identify sequence of amino acids in a protein.[1] The amino acids chemically combine via amide bonds to from polypeptide chains. The biosynthesis of proteins starts at the N terminus and ends at the C terminus, which carry positive and negative charges respectively at the physiological pH. The sequence of amino acids in the primary structure is unique to each protein and dictates its overall function. Depending on the number and nature of substitution, changes in the primary sequence may or may not change the overall structure.[1] If the change is due to substitution between closely related residues (such as tyrosine and phenylalanine) the three dimensional structure and its function is generally conserved. However,

changes in primary sequence which does not conserve the three dimensional structure leads to protein misfolding which is associated with protein misfolding diseases.[2]

Secondary structure is a localized spatial arrangement of amino acids. It is the first step in the folding of proteins and polypeptides. There are three types of secondary structures: helices, beta sheet and turns.

A helix is characterized by the number of residues per turn, the number of atoms making the cycle formed by the hydrogen bond and the repeat distance. The most common helix is the α-helix with a 3.6 residues per turn where hydrogen bonds are formed between C=O of residue n and N-H of residue n+4 (Figure 1-1). The other less common helices are the 3_{10}-helix and π-helix. The 3_{10}-helix has 3 residues per turn forming a hydrogen bond between C=O of residue n and N-H n+3. The π-helix is a more loosely coiled helix between C=O of residue n and N-H n+5.[3] The amino acids proline, hydroxyl proline and glycine break α-helices. Other amino acids such as glutamine, methionine, leucine tend to stabilizes helical structures.[4]

The β-sheets are the other most common structural element in protein. The β-sheet, unlike the α-helix, can be built from a combination of several regions of the polypeptide chain. The β-sheets are made up of two or more strands laterally pack together such that hydrogen bonds can form between C=O groups of one β–strands and N-H groups on an adjacent β –strands and vice versa. β –strands can arrange next to each other to form sheets. The directionality of adjacent β –strands leads to the formation of "parallel", "antiparallel" and "mixed" β-sheets.[1] Among amino acids valine, isoleucine and phenylalanine have been found to stabilize β-sheets.

Proline which cannot participate in the formation of hydrogen bonds between strands destabilizes β-sheets.

Most proteins structures based on combinations of α-helices and β-sheets are connected by turns or loops of various lengths and irregular shapes. Turns have the universal role of enabling the polypeptide backbone to change direction and are important elements that allow and drive protein compaction. Analysis of the amino acid composition in turns reveals that bulky or branched side chains occur rarely while residues with Gly, Asp, Asn, Ser, Cys and Pro with small side chains are observed predominantly in turns.[3,4]

The tertiary structure represents the three dimensionally folded polypeptide chain. The secondary structure is stabilized through hydrogen bonding while the tertiary structure stability is mainly due to hydrophobic/hydrophilic interactions as well as salt and disulfide bridges. The nonpolar residues tend to collect in the center of the protein while the hydrophilic residues are in the surface of the protein. The size and the shape of a protein is determined by the length of the amino acids and by how the secondary structure is arranged inside the protein.[1]

Many proteins are made up of more than one polypeptide chain. The structure formed upon the interaction between single tertiary structures of more than one polypeptide chain is called quaternary structure. The interactions responsible for the quaternary structure are the same as in the tertiary structures. The subunits in the quaternary structure could be based on either identical or different subunits. Quaternary structure of proteins allows greater regulation of transport function and enzyme activity.[3,1] Following their synthesis protein molecules must fold into native conformation to perform their biological functions. However under some conditions

proteins fail to fold correctly leading to protein aggregation both *in vivo* and *in vitro*. The aggregated proteins lack the biological activity of the native protein and are associated with pathological conditions[5] and are a major problem in the manufacturing of proteins.[6]

1.2 Protein aggregation

The ability of proteins to fold from the linear sequence of amino acid forming a functional conformation is one of the most important biological processes. Certain conditions (such as environmental conditions[7], and mutation[8]) initiate misfolding of protein. Protein misfolding leads to a loss of the function carried out by that protein. The correct folding of a protein involves temporary interactions with helper protein or molecular chaperones and is governed by evolutionary pressures that adjust the folding rates according to physiological requirements. Incorrect folding of a protein results in protein aggregation. Protein aggregates are insoluble molecular self-assembly that have lose of their native conformation and function. It has been shown that the protein aggregation can be influenced by a number of variables that includes: a) environmental factors which include concentration of protein, type of solvent, salt content, metal ions, pH of medium, temperature and pressure and b) structural factors involving hydrophobicity, polarity and β-sheet secondary structure content.[9] Protein aggregation that is commonly found in biopharmaceutical industries is unwanted in protein drug production.[6] The aggregates are major concern in the manufacturing process such as purification, freeze drying and storage. Protein aggregation is also associated with a number of human diseases and is gaining an increasing role in human health.[9] The mechanism explaining how amyloidoses cause cell death in neurodegenerative diseases still evades our knowledge. Plaques of aggregated

amyloid fibrils, originally suggested as the cause of the disease are no longer considered to be the pathogenic factor. Instead, the small soluble oligomers [10,11] formed at the beginning of the aggregation process are now believed to be the main cytotoxic entities. Amyloid forming proteins have been demonstrated to form pore-like(channel) structures in artificial as well as biological membranes.[12] These ion channel structures have been proposed as the reason for cell pathophysiology and degeneration in amyloid diseases.[13]

1.3 Amyloid aggregation and occurrence of amyloid disease

Amyloid and amyloid like fibrils are elongated, insoluble protein aggregates deposited *in vivo* in amyloid disease or formed *in vitro* from soluble proteins. Many amyloid proteins are now known (Table 1-1) and although they vary in primary sequence, origin and significance in normal versus aberrant physiology, they all share the property of forming water-insoluble stable aggregates with β-sheet structures. In recent years, amyloid proteins have been found throughout the human body, in a variety of different species, and playing a key functional role in some cases.[14] These discoveries have challenged the previous implication that amyloids occur through defective protein folding and cause disease, suggesting instead that amyloids may comprise evolutionally conserved folds that perhaps have important, as yet unidentified, roles in normal cellular physiology.[15] Over 30 human diseases are now associated with amyloidogenesis (Table 1-1), the formation of aggregated β-sheet structures that appear as water-insoluble deposits of "amyloid" fibrils. Amyloidosis can be classified very broadly into either localized or systemic amyloidosis, depending on the location of the amyloid fibres and the genetic (hereditary) or acquired nature of the precursor protein.

1.4 Structure of amyloid fibrils

Amyloid fibrils are linear unbranched protein aggregates associated with several degenerative diseases as well as denatured globular proteins, bacterial inclusion bodies and several normal cellular functions.[16,17] Fibrils appear to arise from the spontaneous unfolding of the proteins, the exposure of fibril-forcing segments and subsequent self-assembly.[18] These fibrils share enriched β-structure reflected in a cross-β diffraction pattern and Congo red birefringence.[19,20] The atomic-resolution structures of amyloid-forming peptides have been recently determined using X-ray diffraction, electron microscopy, atomic force microscopy, solid state NMR (ssNMR), and computational methods.[21,22,23] It is challenging to investigate the structure and dynamics of amyloid fibrils at the residue and atomic resolution because of their high molecular weight and heterogeneous properties. The one-dimensional nature of the order in the fibrils makes them poor candidates for three-dimensional crystallization. To date, the only representative crystal structures are of amyloidogenic peptides that are short enough to pack in a three-dimensional lattice.[24] These structures indicate that the most basic cross-β structure is in fact a one-dimensional crystal with single translational and rotational symmetry elements. Being neither 2D nor 3D-crystalline nor soluble, protein fibrils are difficult to investigate by X-ray diffraction or solution NMR methods, and ssNMR is the method of choice for the characterization of their structure and dynamics.[25]

As reported by Eisenberg *et al.,*[26] almost all complex proteins, even though not structurally similar, have short segments that if exposed to an appropriate environment (and sufficiently flexible) are capable of triggering amyloid formation. The modern biophysical description of amyloid is unbranched protein fiber whose repeating substructure consists of β

strands that run perpendicular to the fiber axis (Figure 1-2), forming a cross-β sheet of indefinite length.[2] Thus, amyloids are composed of an ordered arrangement of many (usually thousands) copies of a peptide or protein. They are easily identified using electron microscopy (EM) as long filaments with diameters of 6–12 nm.[27] The repeating cross-β sheet motif gives rise to characteristic X-ray fiber diffraction pattern with meridional reflections at ~4.7 Å corresponding to the inter-β strand H-bonding and equatorial reflections at 6–11 Å corresponding to the distance between stacked β sheets.[28,29]

Eisenberg *et al*. were able to grow three-dimensional microcrystals[16] and determine atomic resolution structures, using short fibril-forming peptide segments of amyloid proteins. These structures provided the atomic details of the cross-β spine architecture: a steric zipper made of two β-sheets (8.5 Å apart in this case) mating at a dry, complementary interface of interdigitated side-chains. Each sheet is built by hydrogen bonding (H-bonding) of identical segments along the fibril direction, separated by 4.9 Å.[22,30] While H-bonds hold each sheet together, van der Waals interactions bind the two sheets into the zipper spine. In contrast to the microcrystals, the fibrils show an interstrand twist angle that contributes to their stability, providing self-assembling and self-healing properties[31,32] Twist reflects the fact that successive protein chains in a β-sheet are not stacked exactly above one another but with a small angular offset.[32] This twisting optimizes the H-bonds, side-chain stacking, and electrostatic interactions.[33]

Even though amyloids were originally primarily associated with neurodegenerative diseases, in recent years, a steadily growing number of specific amyloids have been discovered that demonstrate biologically functional properties. Some of the amyloids were found to cause

amyloidosis in mammals by infection (termed Prions, from "infectious proteins"), others are not infectious and may be accumulated for internal reasons. While pathological for mammals, some Prions are functional in yeast and fungi (such as HET-s). Atomistic amyloid models have been recently classified into three types (see Figure 1-3). Type 1, characteristic of disease amyloids, is formed by lateral stacking of identical β-arches (strand-loop-strand motif). Type 2 is also pathological; it consists of superpleated β-arches with more than two strands. The functional amyloids (FA), in contrast, were found to belong to Type 3 that is called β-solenoids. The β-solenoids are built of repeated coils of two or more β-arches long with non-identical but related sequence alternates along the fibril[34]

1.5 Nucleation and template assisted self assembly of amyloid peptides

Molecular self assembly governed by noncovalent interactions is common in nature. The spontaneous organization of peptides or proteins is of biomedical significant as they are associated with various amyloid diseases (Table 1-1). Amyloid fibril formation appears to be a multistep process during which a number of intermediate aggregates are formed. The aggregation starts with the coalescence of peptide monomers to form small oligomeric aggregates such as dimers, trimers, etc. These small oligomers then grow further in size and complexity evolving into protofibrils and then mature amyloid fibrils. The full mechanism of amyloid aggregation is still unclear. The formation of amyloid can be considered to involve at least three steps (Figure 1-4) and are generally referred to as lag phase, growth phase (or elongation) phase and an equilibration phase. The lag phase represents the early stage of the amyloid fibril formation, prior to any detection of fibrils. During the fibril growth phase, the

initially soluble fibrils may become insoluble and precipitate as they exceed certain size and solubility limit. The equilibrium phase (or steady state phase) refers to a situation where the system, consisting of for example fibrils and monomers, appears to be in equilibrium. Although the first step is under kinetic control, fibril growth is under thermodynamic control and can be evaluated quantitatively in terms of equilibrium properties such as association constants.

Among amyloid formation and other nucleated processes a shared feature is that the lag phase can be accelerated on addition of aggregate which is referred to seeding.[35] Seeding involves the addition of a preformed fibrils to a monomer solution thus increasing the rate of conversion to amyloid fibrils.[36] Addition of seeds decreases the lag phase by eliminating the slow nucleation phase. Seed can be homogenous (same peptide) or heterogeneous (related or unrelated peptide) as long as it can provide its growth face as a template for the polymerization of the complement monomers. Recent studies of amyloid growth indicate that, in addition to the self-interactions mediating pathogenic self association, cross-amyloid interactions (also referred as cross seeding) may play a critical role in amyloid diseases. Examples of such interactions include the Aβ–tau, Aβ-amylin, tau–α-synuclein, and Aβ–transthyretin interaction.[37, 38, 39, 40,41]

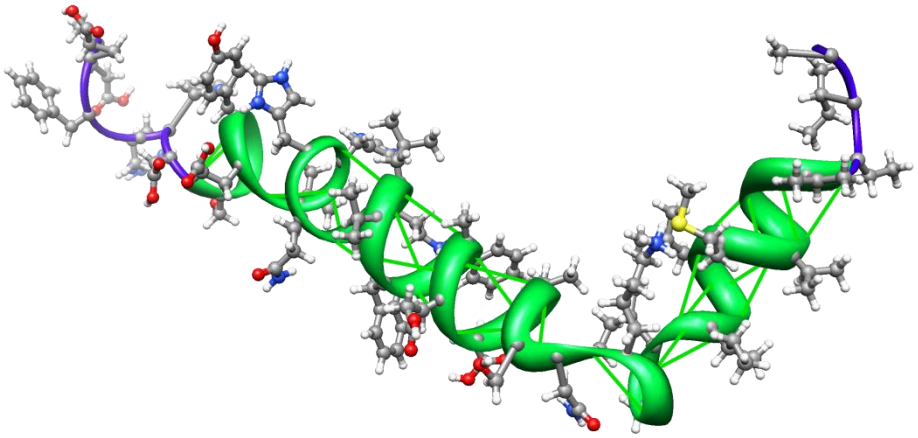
1.6 Designing of aggregation inhibitors

In the case of protein aggregation disease the development of therapeutics agent is focused on interfering with aggregation pathway or increase degradation of misfolded aggregates.[42,43] At least five different targets have been proposed to intervene against aggregation in amyloid disease (Figure 1-5) which includes: A) decrease the expression of the protein associated with the misfolding and aggregation; B) native protein stabilization; C)

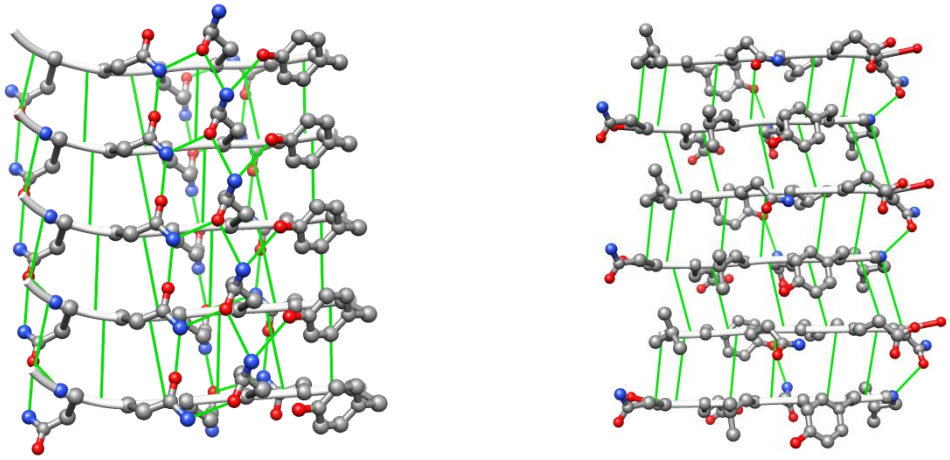
inhibition aggregation and reversal of protein conformational changes; D) increase the clearance of the misfolded protein; and E) prevent tissue degeneration induced by misfolded aggregates.

Table 1.1 Examples of the most common human protein aggregation diseases and the protein associated with the diseases.[44]

Pathological condition	**Protein associated with the diseases**
Alzheimer's disease	Aβ peptides (plaques); tau protein (tangles)
Spongiform encephalopathies	Prion proteins (full length or fragments)
Primary systemic amyloidosis	Immunoglobulin light chains (full length or fragments)
Secondary systemic amyloidosis	Serum amyloid A (full length or 76-residue fragment)
Fronto-temporal dementias	Tau (wild type or mutant)
Senile systemic amyloidosis	Transthyretin (full length or fragments)
Familial amyloid polyneuropathy I	Transthyretin (more than 45 mutants)
Hereditary cerebral amyloid angiopathy	Cystatin C (minus a 10 residue fragment)
Haemodialysis-related amyloidosis	β2-Microglobulin
Familial amyloid polyneuropathy III	Apolipoprotein A1 (fragments)
Atrial amyloidosis	Atrial natriuretic factor
Hereditary non-neuropathic systemic amyloidosis	Lysozyme (mutants)
Hereditary renal amyloidosis	Fibrinogen
Injection localised amyloidosis	Insulin
Huntington's disease	Huntingtin (intact or poly(Q) rich fragments)
Spinocerebellar ataxias	Ataxins (intact or poly(Q) rich fragments)
Medullary carcinoma of the thyroid	Calcitonin
Parkinson's disease	α-Synuclein (aa 1-100)
Type II diabetes	Amylin (aa 1-37)
Glaucoma	Aβ peptides



A



B

Figure 1.1Alpha (A) and beta sheet (B) secondary structure of amyloid peptides. A) alpha-helix structure of Aβ1-42 (Pdb code 1IYT), with helix I (residue 8-25) connected to helix II (residues 28-38) by a turn (residues 26-27). B) β-sheets structure of a short peptide segment of amyloid peptides. The GNNQQNY segment from yeast prion amyloid with parallel (left panel pdb code 1YJP) and VQIVYK segment from human tau amyloid protein (right panel pdb code 2ON9) with antiparallel conformation. Hydrogen and carbon atoms are colored gray, nitrogen blue and oxygen red. The hydrogen bonds are represented with green line. Image created with chimera.[45]

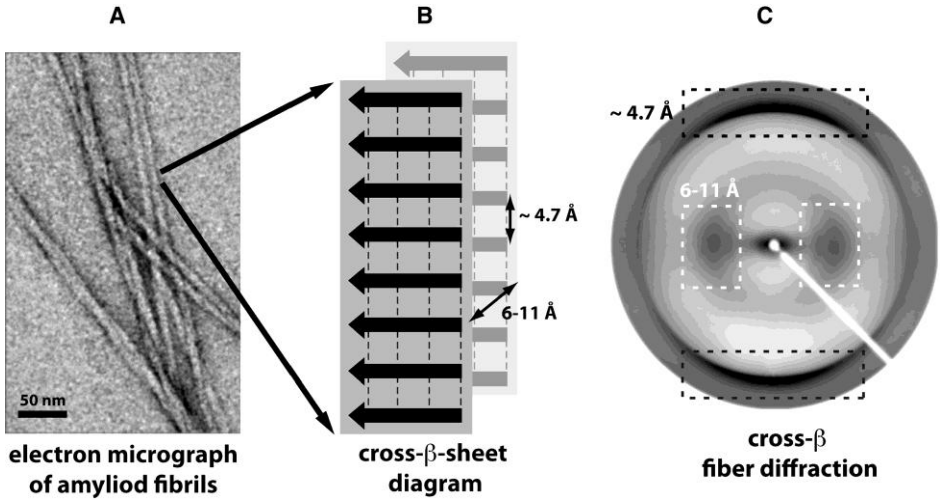


Figure 1-2 Structure of amyloid peptides (Adopted from 2) (A) Amyloid fibrils are composed of long filaments that are visible in negatively stained transmission electron micrographs; (B) Ribbon diagram of the cross-β sheets in a fibril, with the backbone hydrogen bonds represented by dashed lines; (C) the fiber diffraction pattern with a meridional reflection at 4.7 Å (black dashed box) and an equatorial reflection at 6–11 Å (white dashed box), that arise from the β-strand and β-sheet spacing respectively. Adopted from Ref [2,24]

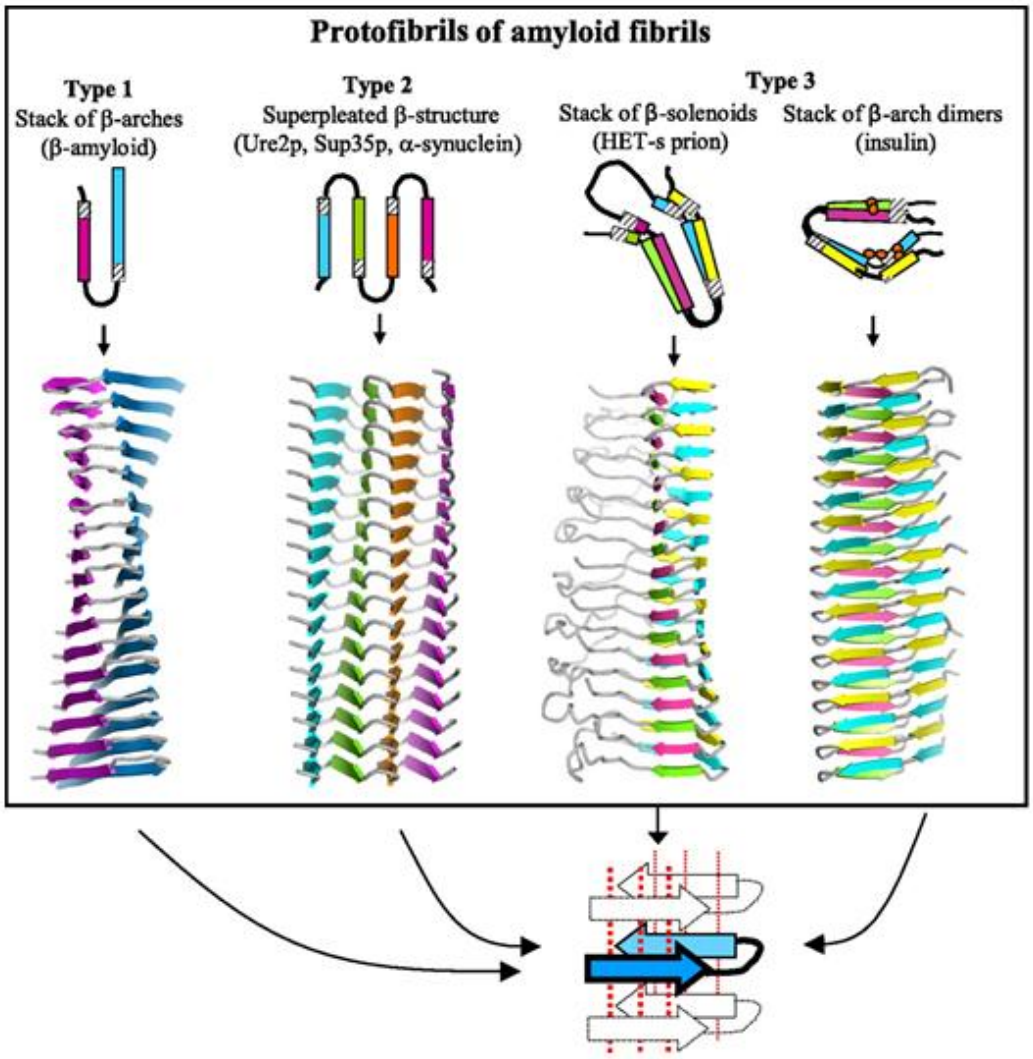
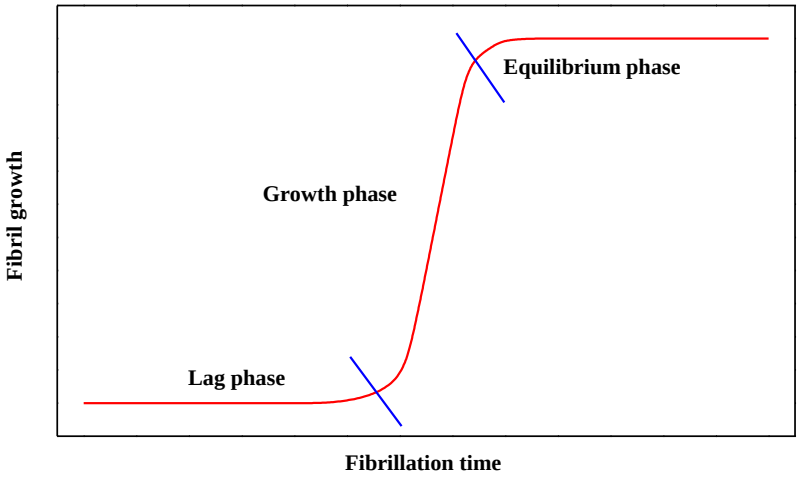
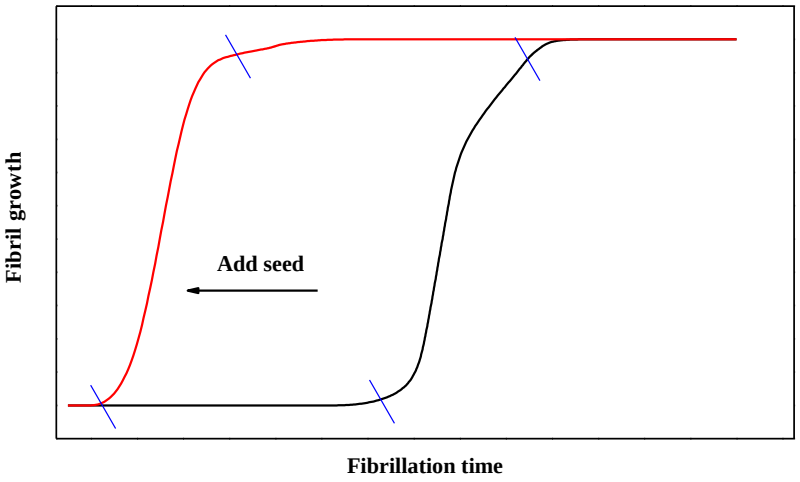


Figure 1-3 Three types of cross beta models of amyloid protofibrils; top: axial views of the repetitive structural units (rectangles represent β strands); bottom: lateral views of protofibrils formed by stacking of these repetitive units. Orange circles in the insulin model show Cys residues forming disulfide bonds. Beneath, schematic diagram of a β arcade, considered to be structural motif common to all 3 types of models. One β arch is colored in blue, with depth cuing; arrows indicate β strands; dotted lines show H bonds. Adopted from Ref. [32]



A



B

Figure 1-4 Amyloid aggregation growth curve (A) and effect of addition of seeds on the lag phase (B) Adopted from Ref. [35]

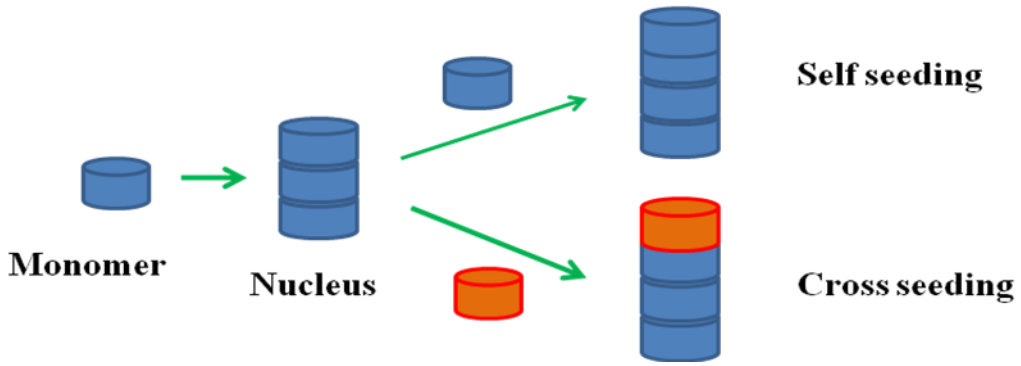


Figure 1-5 Mechanism of nucleation dependent amyloid fibril formation and cross seeding Adopted from Ref. [46,47]

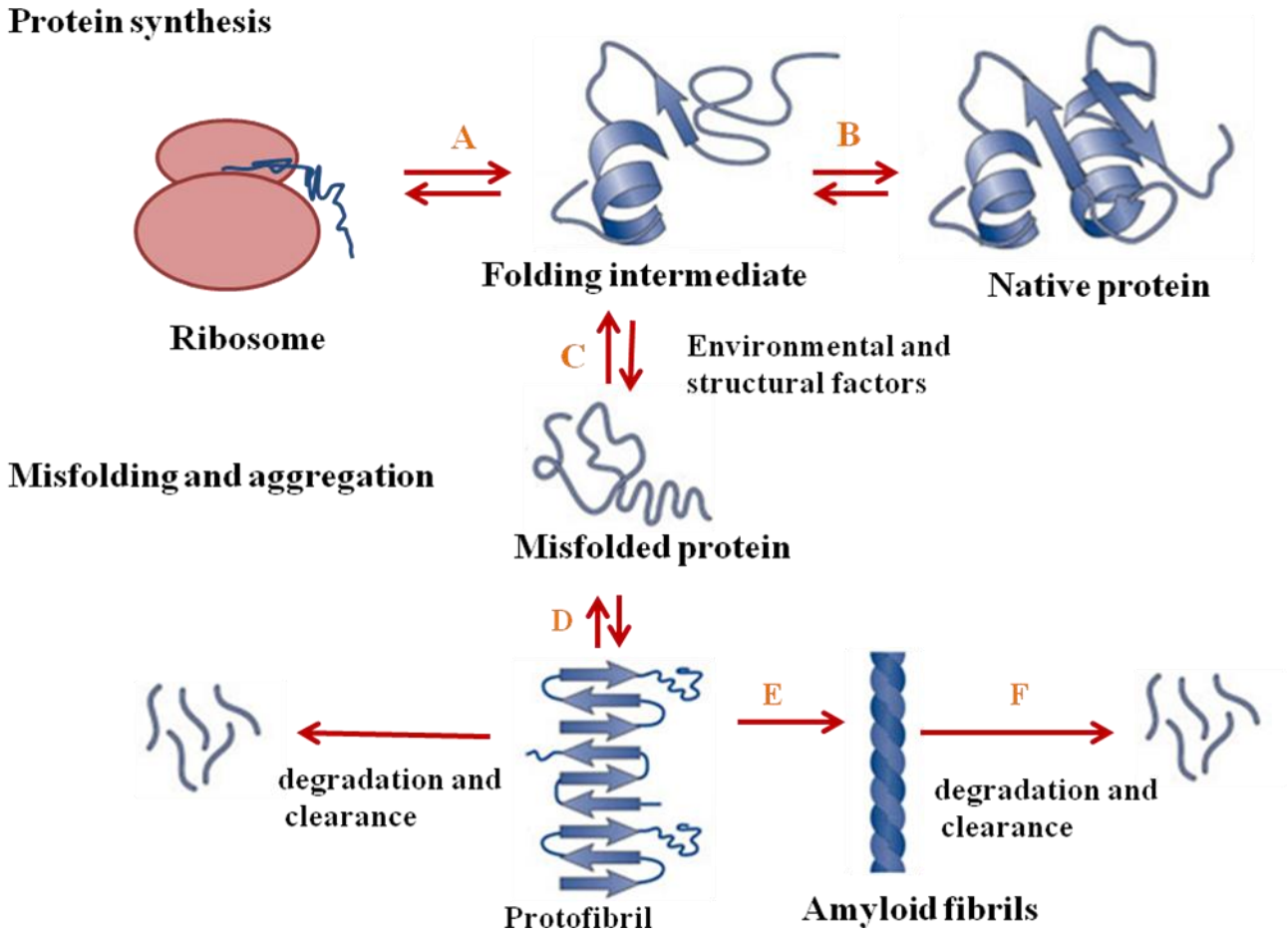


Figure 1-6 Amyloid aggregation and therapeutic intervention in amyloid diseases. The conversion of normally soluble peptides and proteins into insoluble aggregates illustrated in a schematic manner. The stages in the aggregation process where therapeutic intervention may be able to prevent or reverse aggregation are indicated. Therapeutic strategies include (A) stabilizing the native state; (B) inhibiting enzymes that process proteins into peptides with a propensity to aggregate; (C) altering protein synthesis; (D) stimulating clearance of misfolded proteins, for example, by boosting their proteasomal degradation; (E) perturbing fibril assembly; (F) neutralizing or preventing accumulation of fibril precursors. Adopted from Ref. [9, 43]

1.7 References

1. Devlin, T. M., *Text Book of Biochemistry with clinical correlations*. Seventh ed.; John Wily and Sons, Inc New Jersey, 2010.

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

3. Whitford, D., *Protein structure and functions*. John Wiley & sons, ltd: Chichester, 2005.

4. Scheibel, J. P. Z. a. T., *Protein folding-misfolding: some current concepts of protein chemistry*. Nova Science Publishers: New York, 2007.

5. Dobson, C. M., The structural basis of protein folding and its links with human disease. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **2001,** 356, (1406), 133-145.

6. Groenning, M.; Frokjaer, S.; Vestergaard, B., Formation Mechanism of Insulin Fibrils and Structural Aspects of the Insulin Fibrillation Process. *Current Protein & Peptide Science* **2009,** 10, (5), 509-528.

7. Stine, W. B.; Dahlgren, K. N.; Krafft, G. A.; LaDu, M. J., In vitro characterization of conditions for amyloid-beta peptide oligomerization and fibrillogenesis. *Journal of Biological Chemistry* **2003,** 278, (13), 11612-11622.

8. Senguen, F. T.; Doran, T. M.; Anderson, E. A.; Nilsson, B. L., Clarifying the influence of core amino acid hydrophobicity, secondary structure propensity, and molecular volume on amyloid-beta 16-22 self-assembly. *Molecular Biosystems* **2011,** 7, (2), 497-510.

9. Tyedmers, J.; Mogk, A.; Bukau, B., Cellular strategies for controlling protein aggregation. *Nature Reviews Molecular Cell Biology* **2010,** 11, (11), 777-788.

10. Dahlgren, K. N.; Manelli, A. M.; Stine, W. B.; Baker, L. K.; Krafft, G. A.; LaDu, M. J., Oligomeric and fibrillar species of amyloid-beta peptides differentially affect neuronal viability. *Journal of Biological Chemistry* **2002,** 277, (35), 32046-32053.

11. Walsh, D. M.; Selkoe, D. J., A beta Oligomers - a decade of discovery. *Journal of Neurochemistry* **2007,** 101, (5), 1172-1184.

12. Quist, A.; Doudevski, L.; Lin, H.; Azimova, R.; Ng, D.; Frangione, B.; Kagan, B.; Ghiso, J.; Lal, R., Amyloid ion channels: A common structural link for protein-misfolding disease. *Proceedings of the National Academy of Sciences of the United States of America* **2005,** 102, (30), 10427-10432.

13. Naqvi, S. H.; Wang, W. S.; Miao, J. Y.; He, R. Q., Pore-like Aggregates of Tau Protein Induced by Formaldehyde. *Progress in Biochemistry and Biophysics* **2010,** 37, (11), 1195-1203.

14. Fowler, D. M.; Koulov, A. V.; Alory-Jost, C.; Marks, M. S.; Balch, W. E.; Kelly, J. W., Functional amyloid formation within mammalian tissue. *Plos Biology* **2006,** 4, (1), 100-107.

15. Maji, S. K.; Perrin, M. H.; Sawaya, M. R.; Jessberger, S.; Vadodaria, K.; Rissman, R. A.; Singru, P. S.; Nilsson, K. P. R.; Simon, R.; Schubert, D.; Eisenberg, D.; Rivier, J.; Sawchenko, P.; Vale, W.; Riek, R., Functional Amyloids As Natural Storage of Peptide Hormones in Pituitary Secretory Granules. *Science* **2009,** 325, (5938), 328-332.

16. Wiltzius, J. J. W.; Landau, M.; Nelson, R.; Sawaya, M. R.; Apostol, M. I.; Goldschmidt, L.; Soriaga, A. B.; Cascio, D.; Rajashankar, K.; Eisenberg, D., Molecular mechanisms for protein-encoded inheritance. *Nature Structural & Molecular Biology* **2009,** 16, (9), 973-U98.

17. Sipe, J. D.; Benson, M. D.; Buxbaum, J. N.; Ikeda, S.; Merlini, G.; Saraiva, M. J. M.; Westermark, P., Amyloid fibril protein nomenclature: 2010 recommendations from the

nomenclature committee of the International Society of Amyloidosis. *Amyloid-Journal of Protein Folding Disorders* **2010,** 17, (3-4), 101-104.

18. Teng, P. K.; Eisenberg, D., Short protein segments can drive a non-fibrillizing protein into the amyloid state. *Protein Engineering Design & Selection* **2009,** 22, (8), 531-536.

19. Glenner, G. G.; Eanes, E. D.; Page, D. L., Relation of properties of congo red-stained amyloid fibrils to beta-conformation. *Journal of Histochemistry & Cytochemistry* **1972,** 20, (10), 821-828.

20. Serpell, L. C.; Berriman, J.; Jakes, R.; Goedert, M.; Crowther, R. A., Fiber diffraction of synthetic alpha-synuclein filaments shows amyloid-like cross-beta conformation. *Proceedings of the National Academy of Sciences of the United States of America* **2000,** 97, (9), 4897-4902.

21. Zheng, J.; Jang, H.; Ma, B.; Tsai, C. J.; Nussinov, R., Modeling the Alzheimer A beta(17-42) fibril architecture: Tight intermolecular sheet-sheet association and intramolecular hydrated cavities. *Biophysical Journal* **2007,** 93, (9), 3046-3057.

22. Sawaya, M. R.; Sambashivan, S.; Nelson, R.; Ivanova, M. I.; Sievers, S. A.; Apostol, M. I.; Thompson, M. J.; Balbirnie, M.; Wiltzius, J. J. W.; McFarlane, H. T.; Madsen, A. O.; Riekel, C.; Eisenberg, D., Atomic structures of amyloid cross-beta spines reveal varied steric zippers. *Nature* **2007,** 447, (7143), 453-457.

23. Luca, S.; Yau, W. M.; Leapman, R.; Tycko, R., Peptide conformation and supramolecular organization in amylin fibrils: Constraints from solid-state NMR. *Biochemistry* **2007,** 46, (47), 13505-13522.

24. Marshall, K. E.; Serpell, L. C., Fibres, crystals and polymorphism: the structural promiscuity of amyloidogenic peptides. *Soft Matter* **2010,** 6, (10), 2110-2114.

25. Bockmann, A., 3D protein structures by solid-state NMR spectroscopy: Ready for high resolution. *Angewandte Chemie-International Edition* **2008,** 47, (33), 6110-6113.

26. Goldschmidt, L.; Teng, P. K.; Riek, R.; Eisenberg, D., Identifying the amylome, proteins capable of forming amyloid-like fibrils. *Proceedings of the National Academy of Sciences of the United States of America* **2010,** 107, (8), 3487-3492.

27. Sunde, M.; Blake, C., The structure of amyloid fibrils by electron microscopy and X-ray diffraction. In *Advances in Protein Chemistry, Vol 50*, Academic Press Inc: San Diego, 1997; Vol. 50, pp 123-159.

28. Astbury, W. T., Dickinson, S, and Bailey, K., The X-ray interpretation of denaturation and the structure of the seed globulins. *Biochem. J.* **1935,** 29, 2351–0.

29. Sunde, M.; Serpell, L. C.; Bartlam, M.; Fraser, P. E.; Pepys, M. B.; Blake, C. C. F., Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *Journal of Molecular Biology* **1997,** 273, (3), 729-739.

30. Nelson, R.; Sawaya, M. R.; Balbirnie, M.; Madsen, A. O.; Riekel, C.; Grothe, R.; Eisenberg, D., Structure of the cross-beta spine of amyloid-like fibrils. *Nature* **2005,** 435, (7043), 773-778.

31. Xu, Z. P.; Paparcone, R.; Buehler, M. J., Alzheimer's A beta(1-40) Amyloid Fibrils Feature Size-Dependent Mechanical Properties. *Biophysical Journal* **2009,** 98, (10), 2053-2062.

32. Kajava, A. V.; Baxa, U.; Steven, A. C., Beta arcades: recurring motifs in naturally occurring and disease-related amyloid fibrils. *Faseb Journal* **2010,** 24, (5), 1311-1319.

33. Yu, X.; Wang, J. D.; Yang, J. C.; Wang, Q. M.; Cheng, S. Z. D.; Nussinov, R.; Zheng, J., Atomic-Scale Simulations Confirm that Soluble beta-Sheet-Rich Peptide Self-Assemblies

Provide Amyloid Mimics Presenting Similar Conformational Properties. *Biophysical Journal* **2009,** 98, (1), 27-36.

34. Fowler, D. M.; Koulov, A. V.; Balch, W. E.; Kelly, J. W., Functional amyloid - from bacteria to humans. *Trends in Biochemical Sciences* **2007,** 32, (5), 217-224.

35. Harper, J. D.; Lansbury, P. T., Models of amyloid seeding in Alzheimier's disease and scrapie: Mechanistic truths and physiological consequences of the time-dependent solubility of amyloid proteins. *Annual Review of Biochemistry* **1997,** 66, 385-407.

36. Lasagna-Reeves, C. A.; Castillo-Carranza, D. L.; Guerrero-Munoz, M. J.; Jackson, G. R.; Kayed, R., Preparation and Characterization of Neurotoxic Tau Oligomers. *Biochemistry* **2010,** 49, (47), 10039-10041.

37. Andreetto, E.; Yan, L. M.; Tatarek-Nossol, M.; Velkova, A.; Frank, R.; Kapurniotu, A., Identification of Hot Regions of the A beta-IAPP Interaction Interface as High-Affinity Binding Sites in both Cross- and Self-Association. *Angewandte Chemie-International Edition* **2010,** 49, (17), 3081-3085.

38. Guo, J. P.; Arai, T.; Miklossy, J.; McGeer, P. L., A beta and tau form soluble complexes that may promote self aggregation of both into the insoluble forms observed in Alzheimer's disease. *Proceedings of the National Academy of Sciences of the United States of America* **2006,** 103, (6), 1953-1958.

39. Nicolls, M. R., The Clinical and Biological Relationship between Type II Diabetes Mellitus and Alzheimer's Disease. *Current Alzheimer Research* **2004,** 1, (1), 47-54.

40. Wills, J.; Jones, J.; Haggerty, T.; Duka, V.; Joyce, J. N.; Sidhu, A., Elevated tauopathy and alpha-synuclein pathology in postmortem Parkinson's disease brains with and without dementia. *Experimental Neurology* **2010,** 225, (1), 210-218.

41. Du, J. L.; Murphy, R. M., Characterization of the Interaction of beta-Amyloid with Transthyretin Monomers and Tetramers. *Biochemistry* **2010,** 49, (38), 8276-8289.

42. Aguzzi, A.; O'Connor, T., Protein aggregation diseases: pathogenicity and therapeutic perspectives. *Nature Reviews Drug Discovery* **2010,** 9, (3), 237-248.

43. Bartolini, M.; Andrisano, V., Strategies for the Inhibition of Protein Aggregation in Human Diseases. *Chembiochem* **2010,** 11, (8), 1018-1035.

44. Harrison, R. S.; Sharpe, P. C.; Singh, Y.; Fairlie, D. P., Amyloid peptides and proteins in review. In *Reviews of Physiology, Biochemistry and Pharmacology, Vol 159*, Springer-Verlag Berlin: Berlin, 2007; Vol. 159, pp 1-77.

45. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E., UCSF chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **2004,** 25, (13), 1605-1612.

46. Han, H. Y.; Weinreb, P. H.; Lansbury, P. T., The core Alzheimers peptide NAC forms amyloid fibrils which seed and are seeded by beta-amyloid - is NAC a common trigger or target in neurodegenerative disease. *Chemistry & Biology* **1995,** 2, (3), 163-169.

47. Bhak, G.; Choe, Y. J.; Paik, S. R., Mechanism of amyloidogenesis: nucleation-dependent fibrillation versus double-concerted fibrillation. *Bmb Reports* **2009,** 42, (9), 541-551.

CHAPTER 2 MOLECULAR DYNAMIC SIMULATION AND TRAJECTORY ANALYSIS

2.1 Bimolecular Simulations

Molecular simulation is a way to visualize a system by generating successive configurations of the system. While experimental techniques such as x-ray crystallography can generate a snapshot of a protein (or other macromolecule), the positions of mobile elements, such as flexible loops, may remain unclear. It is possible to visualize these mobile elements with simulations. Additionally, x-ray crystallography and NMR methods are often employed under non-physiological conditions (temperature, pressure, pH, solvent, etc.), which can affect their results in unpredictable ways. Many biologically important process involves a change in the three dimensional structures of bio-molecules such as protein, RNA and DNA.[48] Protein folding, protein aggregation, enzymatic catalysis, signal transduction and other biological process involve conversion of proteins structures.[48] Bimolecular simulations can provide information to the molecular modeler about how a biological system behaves over a certain time period, under physiological conditions [49] providing continuous trajectories that can help in connecting static experimental structures. Therefore computer simulations are used to complement and extend experiment.[50]

There are two main types of simulation methods: Molecular Dynamics (MD) and Monte Carlo (MC) simulations. In MD, the configurations are produced by integrating Newton's laws of motion, resulting in a trajectory that specifies how the system behaves with time. The forces on the atoms are used with their current positions and velocities to predict new positions and velocities for the next time step. Over a given time period, a "trajectory" is generated that

describes how the system being studied changes over time. Time averages for thermodynamic properties such as internal energy, heat capacity, pressure, and temperature can be calculated.

2.2 Molecular dynamic simulation

Molecular Dynamics (MD) simulation is a computer simulation which uses molecular mechanics to describe the time evolution of a set of interacting atoms by integrating their equations of motion.[51, 52] MD simulation follows the laws of classical mechanics (Newton's law) for system constituted by *N* atoms, each atom at position r_i is treated as a point with a mass mi and a fixed charge q_i: the force F_i acting upon each atoms is determined by:

$$F_i = -\nabla U(R), \quad (1)$$

Where, U(*R*) is the potential energy of the system as a function of the atoms positions. It is a sum of energy contributed from oscillations about the equilibrium bond length (U_{bond}), oscillations of 3 atoms about an equilibrium bond angle (U_{angle}), torsional rotation of 4 atoms about a central bond ($U_{dihedral}$) and non-bonded energy terms which consists of electrostatics and Lenard-Jones ($U_{nonbond}$).[50]

$$U(R)= U_{bond} + U_{angle} + U_{dihedral} + U_{nonbond} \quad (2)$$

Once the potential energy of the system is obtained using equation 2, the force on each atom can be obtained by solving equation 1. The calculated force determines the acceleration of the atoms, knowing the positions and velocities at time t one can calculate the positions and velocities of the atoms at time (t+Δt):

$$v_i(t+\Delta t) = v_i(t) + \frac{\Delta t}{m_i} F_i(t) \quad (3)$$

$$r_i(t+\Delta t) = r_i(t) + v_i(t)\Delta t + \frac{1\Delta r^2}{2\, m_i} Fi(t) \quad (4)$$

The continuation of the steps above is a trajectory of the position and velocity of the atoms which varies with time.[50] Thus the MD simulation provides time dependent motions of the individual atoms in the system.

2.2.1 Potential energy function

MD simulations are based on knowledge of the potential energy surface which is represented by an empirical function called force fields. The force field is a collection of equations and associated constants designed to reproduce molecular geometry and selected properties of tested structures. The force fields are parameterized to approximately reproduce various experimental results from spectroscopy, calorimetry and/or quantum mechanical studies.[50] The chief advantages of force fields are the incredible reduction in computational requirements. The disadvantage of force fields is they ignore electronic effects and cannot be used to describe molecular properties that depend upon electron distribution, such as chemical reactions. The functional form for typical force field in AMBER package is given by the following equation [53]:

$$U_{\text{bonded}} = \sum K_b (b - b_0)^2 + \sum K_\theta (\theta - \theta_0)^2 + \sum k_\chi [1 + \cos(n\chi - \sigma)] \quad (5)$$

$$U_{\text{non bonded}} = \sum \left(\varepsilon_{ij} \left[\left(\frac{R_{min,ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{min,ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{r_{ij}} \right) \quad (6)$$

where U_{bonded} is the contribution to the total energy from bonded interactions and $U_{non\ bonded}$ is the contribution from non bonded interactions. The total energy is then: $U_{total} = U_{bonded} + U_{non\text{-}bonded}$. The first term in Eq. 5 is a sum over all bonded pairs of atoms and describes the stretching of bonds; *b* is the inter-atom distance (i.e., bond length); and K_b and b_0 are parameters describing the stiffness and the equilibrium length of the bond, respectively. The term has the same quadratic form as that of Hooke's law for the potential energy of a spring. The second term involves triplets of atoms, e.g., A, B, and C, where A is bonded to B and B is bonded to C, and describes the bending of angles. θ is the angle formed by the two bond vectors, K_θ and θ_0 are the parameters describing the stiffness and equilibrium geometry of the angle, and, similar to the term for bond stretching, the term is quadratic. The third and final term in Eq. 5 is a sum over quadruplets of atoms A, B, C, and D, where A is bonded to B, B to C, and C to D, and describes the energetic associated with rotation of the dihedral angle defined by those four atoms. Because such rotation is necessarily periodic in nature, a cosine function is used. χ is the value of the dihedral, K_χ is the energetic parameter that determines barrier heights, *n* is the periodicity or multiplicity, and σ is the phase. It should be noted that the bonded terms are also referred to as internal or intra-molecular interactions.[49]

The non-bonded interactions between atoms are defined as occurring either between atoms in separate molecules or between atoms separated by three or more bonds in the same molecule. Equation 6 is composed of two parts. The first, known as the Lennard-Jones (LJ) equation, is the portion in square brackets along with the prefactor ε_{ij} , and models attractive dispersion and repulsive Pauli exclusion interactions and is commonly referred to as the van der Waals term. As two atoms are brought together from infinite separation, the negative term in the

brackets, which goes as the inverse of the inter-atomic separation r_{ij} to the sixth power, dominates the interaction and the atoms feel an increasing attraction with decreasing distance as the energy becomes progressively more negative. This part of the LJ equation models dispersion, and its $(1/r)^6$ form derives from the interaction energy of an instantaneous dipole with an induced dipole, according to the definition of London's dispersion. As the atoms get progressively closer, an energy minimum is reached and, at closer distances, the $(1/r)^{12}$ term, which is positive, starts to dominate and leads to increasing energy and, hence, repulsion. Its form was originally chosen based on its computational expedience because it is simply the square of $(1/r)^{6.}$ Nonetheless, it serves as an adequate representation of the very steep repulsive energy wall that arises from Pauli exclusion as two atoms get closer than the sum of their van der Waals radii. The prefactor, ε_{ij}, is a parameter based on the types of the two interacting atoms *i* and *j*. As its value increases, the interaction minimum becomes deeper and the repulsive wall steeper. $R_{min,ij}$ is a parameter that also depends on the types of the two interacting atoms and defines the distance at which the LJ energy is a minimum. The second part of Eq. 6 is Coulomb's law and is used to model the electrostatic interaction between non-bonded pairs of atoms. As with the LJ equation, $r_{i\ j}$ is the inter-atomic distance, while q_i and $q_{\ j}$ are the parameters that describe the effective charges on atoms *i* and *j*. It is important to note that the effective charge parameters are not simply unit charges located on formally charged atoms. Rather they are partial atomic charges with non-integer values that are selected to represent the overall charge distribution of a molecule. Naturally, the sum of the partial charges in a molecule must equal the molecule's net formal charge. In addition, in the case of metal ions, the charge is typically assigned the formal charge (e.g., +1 for the sodium ion).[49] MD packages available for

commercial or academic use frequently incorporate their own force field which has already been parameterized by the developers. Some commonly used dynamics packages that include their own force fields are AMBER, CHARMM, and GROMACS.[54,55,56] There are a variety of integration methods currently employed by dynamics software packages, including the Verlet algorithm, [57] the 'leapfrog' algorithm,[58] the velocity Verlet method,[59] and Beeman's algorithm.[59] The most widely employed integration methods today are Verlet and velocity Verlet methods. The Amber package employed in our MD studies uses the velocity Verlet integration method by default.[54]

2.2.2 System setting in MD simulation

Setting up and running a molecular dynamics simulation is a complicated process which requires many considerations, such as the initial configuration of the system being studied, choice of force field and dynamics integration method, time length of the simulation and time steps, type of ensemble and energy calculations, boundary conditions, and solvation. Each consideration can influence the outcome of the simulation as well as the computational expense and time requirements. The initial configuration of the system is usually obtained from experimental data, theoretical models, or a combination of both. For example, for a protein simulation, the structure of the protein may have been obtained from x-ray crystallography, NMR, or homology modeling. Atom types for the force field being used must be defined and parameters developed if necessary. The systems are frequently minimized prior to running dynamics to eliminate high energy interactions such as steric clash.[60]

Once the initial configuration of the system has been defined and the force field and integration method (software package) selected, decisions must be made as to the length of time and the time steps that will be required for the simulation. The length of time will be determined by the nature of the system being studied, the process being studied, and the computational resources available to the modeler.. The calculation time steps is another key consideration and it depend on the integration method being used, the system studied, and the computational resources available.[60] The smaller the time step chosen, the more computational expensive will be the simulation. A standard recommendation is that the time step chosen should be one-tenth the time of the shortest motion being studied. In bio-molecular systems this is usually the C-H bond vibration which occurs on a 10 fs time scale, thus 1fs time steps would typically be chosen. If C-H bonds are held constrained during the simulation using a method known as the SHAKE algorithm, then this time step can be doubled to 2 fs. [61] The next consideration is the type of ensemble to be studied and the types of energy calculations that will be used. Molecular dynamics are traditionally performed using the NVE or micro-canonical ensemble, which holds constant the number of particles (N), the volume (V) and the energy (E). When studying bio-molecular systems, it is more practical to use the NTP, or isothermal-isobaric ensemble, which holds constant the number of particles (N), the temperature (T), and the pressure (P). This simulates physiological conditions more closely than the other types of ensembles.

2.2.3 Treatment of long range columbic force

The most time consuming part of a molecular dynamics simulation is the calculation of long range interactions and there are a variety of methods for handling this. The use of distance

cutoffs for energy calculations is one popular way to address this problem. Cutoffs present a problem with certain types of long-range interactions, such as charge-charge interactions which can still significantly contribute to the energy of the system beyond the standard cutoffs used in most dynamics simulations. Special methods have been developed to address this problem, including the Ewald summation. The version of Ewald summation method known as Particle-Mesh Ewald (PME) is employed in the Amber simulations package.[54]

2.2.4 Boundary condition and solvent models

Finally, boundary conditions and solvation methods must be decided upon. Because interactions at the boundaries of the system being studied (i.e. vacuum, wall, etc.) can influence the energy calculations, the boundaries must be defined or taken into account in some manner. For bio-molecular simulations, the most common way to do this is to employ periodic boundary conditions. Periodic boundaries involve placing the system in a cell, typically a cubic box or other geometric shape, and then surrounding the cell with mirror cells containing replicas of the system. The interactions energies can be calculated across cell boundaries overcoming the boundary effect and enabling the simulation of a much larger system. If a particle leaves one side of the cell, it subsequently enters from the other side; keeping the number of particles in the system constant. The cell size chosen must be large enough so that the actual bio-molecule being studied does not "see" itself and affect its own energy calculations. Usually, it is desirable only for solvent molecules to cross the periodic boundary.[54]

There are currently three different ways to take into account solvation: the first one involves simulating the system *in vacuum* using only a distance dependent dielectric screening

term in the force field to simulate the solvent screening effects on electrostatic charge calculations. This method is the least rigorous, and is the fastest in terms of computational expense; however it is also the least reliable and should be reserved only for simulations where solvent effects is not expected to play a key role. The second method is known as implicit solvation, or continuum solvation.[60] This method uses special energy terms in the force field to represent the solvent as a continuous medium. The two commonly used algorithms to approximate the solvent electrostatic effects are the Poisson-Boltzmann equation, and the Generalized Born model, which is a linear approximation of the Poisson-Boltzmann equation which is computationally less expensive. Both of these equations are often combined with a hydrophobic solvent accessible surface area (SA) term. Implicit solvation models still have limitations. Entropic effects are not accounted for in these models, which can be a major factor in loop movements, ligand binding, and protein folding. The effect of solvent viscosity on the motion of solutes is also not accounted for when using implicit models, although in some cases this can be desirable. Finally, although H-bonding can be generally accounted for with implicit solvation algorithms, the directionality of H-bonds cannot. The third solvation method is the explicit solvation. In this method the solvent molecules are explicitly treated by surrounding the solute or bio molecule by solvent molecules. This method is the most accurate but is also computationally the most expensive as all energy calculations must now include the many solvent molecules. For bio-molecular simulations, there are several water models that have been designed for use, the most commonly used is the TIP3P water model, a 3-site model where the water is represented by a molecule with 3 interaction sites and a rigid shape [62] 4, 5 and 6 site

models have been developed but they increase the computational expense of the simulation and are rarely used except for simulations modeling water dynamics.[49]

Once the molecular dynamics methods have been determined and the system has been set up, the simulation can be run. A typical dynamics simulation of a bio-molecular system under explicit solvation is a multi-step process. An initial solvent minimization is required, where the solvent is minimized while the solute is held under constraint. This is followed by a solvent dynamics step, where the solvent (and any counter ions added to balance the solute charge) are allowed to equilibrate. The next step would be allowing the entire system to minimize while slowly loosening the constraints on the solute, or bio-molecule. This is followed by the dynamics simulation itself which occurs in two phases, an equilibrium phase and a production phase. The equilibrium phase brings the system to equilibrium from the starting configuration. Equilibration is reached when the calculated average temperature, pressure, and energies have stabilized. Finally, the production phase of the simulation can begin, where the system is allowed to fully evolve for the desired time period. Typically only data obtained from the production phase is used to calculate the desired properties.[49]

2.3 Simulation protocol

Molecular dynamics (MD) simulations of various size short segment and full length amyloid peptides and their corresponding mutant as well as were performed using the MD simulation given below.

The molecular dynamic (MD) simulation in this thesis were performed using AMBER11[63] package with an all atom amber99SB force field and explicit TIP3P water models.

Each of the amyloid peptides, the corresponding mutants and amyloid polyphenol complexes were solvated by explicit water molecules that extends 10 Å from any edge of the octahedral box to the protein atoms. Counterions were added to the box by randomly replacing water molecules to neutralize the system. Energy of each system was initially minimized by using conjugate gradient method to remove bad contacts with the peptide atoms first constrained, and then relaxed without position constrains. The system was then subjected to 50 ps of heating procedure while constraining the backbone atoms of the protein to allow relaxation of water and ions, followed by 500 ps equilibration run without position constraints on the peptides. Constant pressure (1 atm) and temperature (300 K) on the system was maintained by isotropic Langevin barostat and a Langevin thermostat. Electrostatic interactions were calculated by using the particle mesh Ewald (PME) method. The cutoff radius for the Lennard-Jones interactions was set to 12 Å. The SHAKE algorithm [64] was used for bond constraints and the time step was 2 fs for all simulations. Each system was simulated for 20 ns and the trajectories were saved at 4.0 ps intervals for further analysis. The VMD (Visual Molecular Dynamics) program was used for the visualization of the trajectories.[65] The MM-PBSA single trajectory approach implemented as script (MMPBSA.py) in AMBER11, was used to calculate the binding energy. The length of the production simulation for the systems in chapter 3 was 10 ns. The simulation temperature was set at 330 K for the amyloid peptide models in chapter 5. The temperature 330 K was selected as a compromise so that amyloid fibrils are still experimentally stable[66,67,68] but molecular system evolves more quickly in the limited simulation time and possible kinetic traps are avoided. The force parameter for curcumin, exifone and myricetin in chapter 6 was generated by GAFF utility [69] in AMBER11 suite. Geometry optimization and partial charges were obtained using

Gaussian03 [70]. After geometry optimization at HF/6-31G* level, the partial charges were derived by fitting to the gas-phase electrostatic potential at the same theory level using the restrained electrostatic potential (RESP) method.

2.4 Analysis of MD trajectories

The trajectory analysis of the simulations was performed using the analysis tools available in VMD, and in ptraj module of AMBER software. MD simulations produce trajectories for the atomic positions and velocities. These quantities are saved at regular intervals. Usually positional information is mainly used the analysis of MD simulations. Several analyses are done once the simulation has been completed to extract structural and energetic information from the production run. The goal of the trajectory analysis is to gain structural and dynamic insights and relating structure to function. The ptraj program in AMBER soft is capable of analyzing and processing trajectory created from MD simulations.[54] The most frequently used trajectory analysis on amyloid peptide aggregates simulations are (a) RMSD and RMSF (b) Secondary structure analysis (c) interstrand distance (d) Intersheet distance (e) hydrogen bond analysis (f) Cluster analysis and (g) MMPBSA binding free energy calculation

A root mean square distance (RMSD) analysis of the amide backbone atoms is often a strong indicator of conformational changes of a protein. The root mean square distance (RMSD) between the backbone atoms of the trajectory frames of polypeptide chains and the corresponding atoms of the x-ray structure, calculated for the frame t, is given by equation 11, where x^m, y^m, z^m are the cartesian coordinates found at the X-ray structure and x^l y^l z^l are the Cartesian coordinates of trajectory frame t. N is the number of atoms.[71]

$$RMSD = \frac{1}{N}\sqrt{\sum_{i=1}^{N}\left(x_i^m - x_i^l\right)^2 + \left(y_i^m - y_i^l\right)^2 + \left(z_i^m - z_i^l\right)^2} \quad (11)$$

To obtain information on local structural flexibility, stability, and effect of mutations on the investigated amyloid peptides molecules a root mean-square fluctuations (RMSF) analysis was performed.[72] RMSF of the Cα atoms of each residue are calculated as follows: RMSF

$$RMSF = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(r_i(t) - \langle r_i\rangle\right)^2} \quad (12)$$

where T is the number of snapshots considered in the time trajectory, $r_i(t)$, the position of the Cα atom of residue i at time t, and $\langle r_i\rangle$, the time-averaged position of the Cα atom of residue I.[73]

The secondary structure dynamics shows the conformational change that occurs for a peptide or protein during the simulation. The commonly used program for analysis of the secondary structure is the Dictionary of Protein Secondary Structure (DSSP). The DSSP algorithm was written by Kabsch[74] and is based on identification of hydrogen-bonding (H-bonding) patterns and recognizes seven types of secondary structures which can be grouped into three classes: helix (α-helix, 3_{10}-helix, π-helix), β-strand (isolated β-bridge, extended β-sheet) and loop (turn, bend). We carried out secondary structure analysis using the DSSP tool in AMBER11.[54]

To examine the structural stability of the wildtype and the corresponding mutant oligomers we also analyzed the inter-strand (d_{strand}) and inter-sheet (d_{sheet}). The d_{strand} is calculated by averaging the mass center distance between each residue in one strand and its corresponding residue in adjacent strand in the same sheet, whereas d_{sheet} is calculated by

averaging the mass center distance between each strand in one sheet and its corresponding strand in the adjacent sheet.[75]

A hydrogen bond is weak electrostatic attraction. It forms when a hydrogen atom covalently binds to an electronegative atom and is electro-statically attracted to another (electronegative) atom. The atom to which the hydrogen atom (H) is covalently bound is considered the hydrogen donor (D), and the other atom is the hydrogen acceptor (A). In biological polymers, the donor and acceptor atoms are either nitrogen or oxygen., In protein for example in helices and sheets, the D–H · ·A sequence is N–H ·· O=C.[76] The strength of a hydrogen bond can be characterized by two geometric quantities which govern the hydrogen bond energy: hydrogen bond angle, D–H · A atoms and optimal hydrogen bond length, H · ·A (or D · A) distance.[76] Hydrogen bond and hydrogen bond occupancies was calculated using PTRAJ module available within AMBER. A hydrogen bond is assigned if the distance between donor D and acceptor A is ≤3.5 Å and the angle D–H · A A≥120^{o}[75] using PTRAJ module available within AMBER.

Cluster analysis ("clustering") places similar samples of data into groups called clusters, such that an ensemble of data (for example the different structures obtained from an MD trajectory) is partitioned into groups of similar objects. Structural clustering is useful for understanding the molecular motion within conformational space.[77] To identify the most populated conformations sampled, clustering was applied to all snapshots from the trajectories using the Ptraj program of AMBER11. To perform the clustering, we utilized the average linkage algorithm implemented in Ptraj.[54]

Amyloid fibrils typically exhibit twisted β-sheets and twisting of β-sheets optimize the hydrogen bonds, side chain stacking, and electrostatic interactions, thus twisted sheets are more stable than flat ones. [78] Twisting angles have been computed by using the method reported by Simone *et al* [79] and Figure 1-2 shows the average twist angle calculation for the Elk prion segment NNQNTF.

2.5 Binding free energy calculation

Free energy calculation methods have become powerful tools as they can provide quantitative measurement of protein-ligand or protein-protein interactions. The molecular mechanic Poison-Boltzman or the generalized born solvent accessible surface area (MM-PB(GB)SA) [80, 81] method as implemented in AMBER 11 was used to calculate the binding energy for non-covalent association of between the studied amyloid peptides. The calculation of the binding free energy requires three independent MD simulations of the complex and both individual protein. However an assumption was made that no significant conformational changes occur upon binding i.e. structural change is negligible and the snap shots for all three species were obtained from the single trajectory carried out on the complex by separating the complex into its constituent parts.

The free energy analyses in this thesis was done using a single trajectory approach, where the complex (C), receptor (B) oligomer aggregate), and ligand (A) snapshots were taken from the snapshot of the performed MD trajectory. According to the MM-GBSA/MM-PBSA method,[80,81] binding free is calculated using equation 11:

$$\Delta G_{bind} = \langle G_C \rangle - \langle G_A \rangle - \langle G_B \rangle \quad (11)$$

The bracket, $\langle\ \rangle$, indicates an average of these energy terms over extracted from the MD simulation.

$$\Delta G_{bind} = \langle \Delta E_{MM} \rangle + \langle \Delta G_{solv} \rangle - \langle T\Delta S \rangle \quad (12)$$

The free energy of each system X=A, B, or C was computed as a sum of the three terms:

$$\langle \Delta G_X \rangle = \langle E_{MM} \rangle + \langle \Delta G_{solv} \rangle - T\langle S \rangle \quad (13)$$

Where E_{MM} is the molecular mechanics energy of the molecule expressed as the sum of the internal energy (bonds, angles and dihedrals) (E_{int}), electrostatic energy (E_{ele}) and van der waals term (E_{vdw}):

$$E_{MM} = E_{int} + E_{ele} + E_{vdw} \quad (14)$$

ΔG_{solv} accounts for the solvation energy which can be divided into the polar and nonpolar part:

$$\Delta G_{solv} = \Delta G_{GB} + \Delta G_{SA} \quad (15A)$$

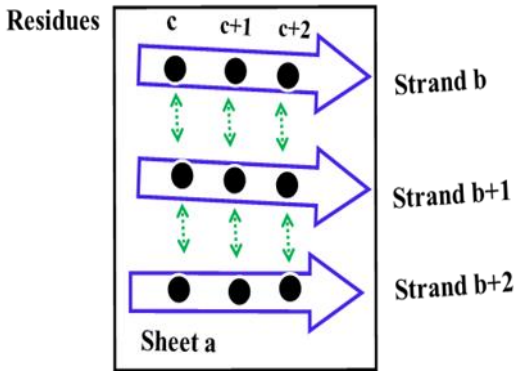
The polar part ΔG_{GB} accounts for the electrostatic contribution to solvation and is obtained from Generalized Born (GB) calculations in a continuum model of the solvent. The second term ΔG_{SA} is nonpolar contribution to solvation free energy that is linearly dependent on the solvent accessible surface area (SASA):

$$\Delta G_{SA} = \gamma SASA + b \quad (15B)$$

The ΔG_{SA} were calculated using AMBER11 default parameter for γ and b (15b). The entropic contribution was calculated in chapter 4 and 6 using the normal mode module in AMBER11.[82, 81]

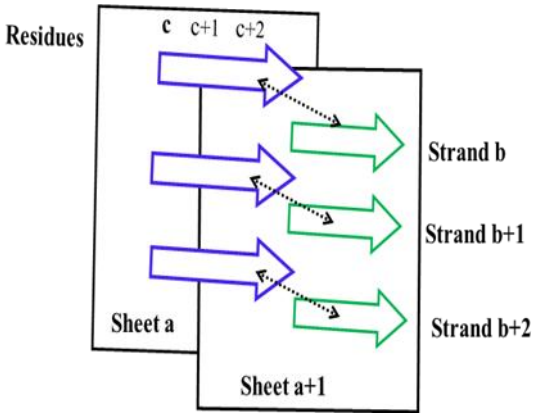
2.6 Application of MD simulation in the study of bimolecular system

Various aspects of protein structure and function have been studied by Molecular dynamics in numerous fields, including structural biochemistry, biophysics, enzymology, molecular biology, pharmaceutical chemistry, and biotechnology.[50] One notable important area of application MD simulation is structure aided drug design. Virtual compound screening using molecular docking is widely used in the discovery of new lead compounds for drug design. However, this method is not completely reliable and therefore unsatisfactory. Okimoto *et al* [83] using combined docking and molecular dynamics simulations has found improvement of 1.6 to 4.0 time in enrichment performance compared to docking method. In the study of protein aggregation MD simulation have provided insight into amyloid structure and aggregation mechanism.[84] MD simulation have been used by various researchers. [85, 86, 87] in order to understand the mechanism of aggregation inhibitor effects of small organic molecules.



$$\langle d_{strand} \rangle = \frac{\sum_{a=1}^{sheets} \sum_{b=1}^{strands} \sum_{c=1}^{residues} \left(r_{a,\ b,\ c} - r_{a,\ b+1,\ c} \right)}{sheets \times (strands - 1) \times residues} \qquad (9)$$

A



$$\langle d_{sheet} \rangle = \frac{\sum_{a=1}^{sheets-1} \sum_{b=1}^{strands} \sum_{c=1}^{residues} \left(r_{a,\ b} - r_{a,\ b+1} \right)}{(sheets - 1) \times strands} \qquad (10)$$

Figure 2-1 Schematic definition of inter-sheet and inter-strand distances Adopted from Ref. [75]

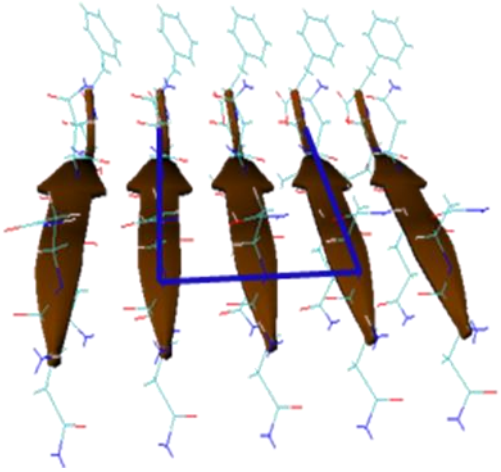


Figure 2-2 Schematic definition of the average twist angle. The twisting of SH1-ST5 of NNQNTF calculated by determining the dihedral angle from the coordinates of the 2nd and the 5th Cα-atom of the first and the last strand of the sheet. The calculated angle provides a measure of the overall twisting of each sheet. The twist angles were calculated by using the three inner strands and the average twist angles between consecutive strands were estimated by dividing the twist by three. [88]

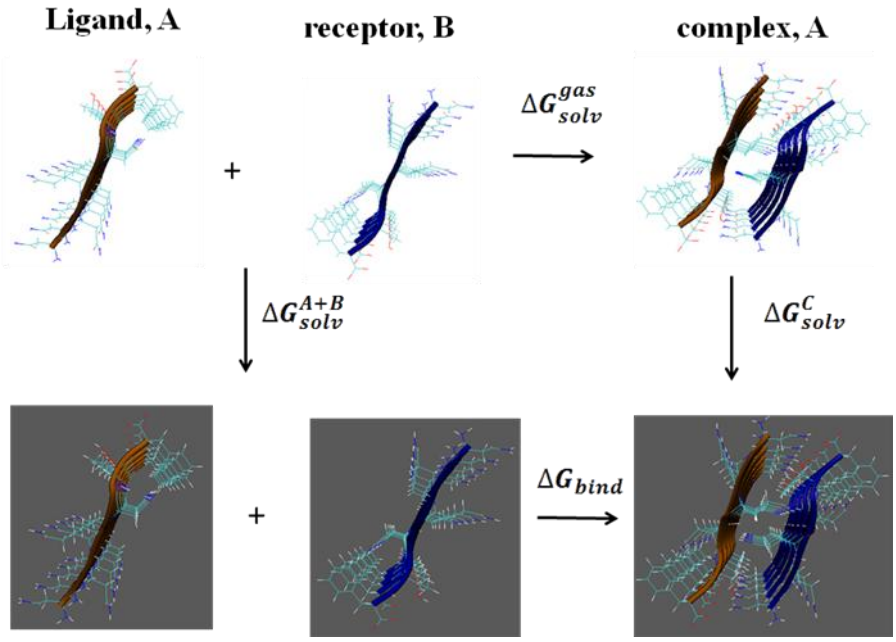


Figure 2-3 Thermodynamic cycle in MM-PB(GB)SA calculations. The gray surface represents the aqueous solvent.

2.7 References

48. Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y. B.; Wriggers, W., Atomic-Level Characterization of the Structural Dynamics of Proteins. *Science* **2010,** 330, (6002), 341-346.

49. Becker, O. M., *Computational biochemistry and biophysics* Marcel Dekker: New York 2001.

50. Adcock, S. A.; McCammon, J. A., Molecular dynamics: Survey of methods for simulating the activity of proteins. *Chemical Reviews* **2006,** 106, (5), 1589-1615.

51. Ciccotti, G. a. H., W. G. , *Molecular dynamics simulation of statistical-mechanical systems*. North-Holland: 1986.

52. Heermann, D. W., *Computer simulation methods*. Springer: 1986.

53. Ponder, J. W.; Case, D. A., Force fields for protein simulations. In *Protein Simulations*, Academic Press Inc: San Diego, 2003; Vol. 66, pp 27-87.

54. Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J., The Amber biomolecular simulation programs. *Journal of Computational Chemistry* **2005,** 26, (16), 1668-1688.

55. Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M., CHARMM: The

Biomolecular Simulation Program. *Journal of Computational Chemistry* **2009,** 30, (10), 1545-1614.

56. Van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C., GROMACS: Fast, flexible, and free. *Journal of Computational Chemistry* **2005,** 26, (16), 1701-1718.

57. Verlet, L., Computer experiments on classical fluids I Thermodynamical properties of Lennard Jones molecules. *Physical Review* **1967,** 159, (1), 98-103.

58. Hockney, R. W., The potential calculation and some applications. *Method in Computational Physics* **1970,** 9, 136-211.

59. Swope, W. C.; Andersen, H. C.; Berens, P. H.; Wilson, K. R., A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules application to small water clusters. *Journal of Chemical Physics* **1982,** 76, (1), 637-649.

60. Senn, H. M.; Thiel, W., QM/MM methods for biological systems. In *Atomistic Approaches in Modern Biology: From Quantum Chemistry to Molecular Simulations*, Springer-Verlag Berlin: Berlin, 2007; Vol. 268, pp 173-290.

61. Swope, W. C.; Andersen, H. C.; Berens, P. H.; Wilson, K. R., A computer simulation method for the calculation of equilibrium-constants for the formation of physical clusters of molecules application to small water clusters. *Journal of Chemical Physics* **1982,** 76, (1), 637-649.

62. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of simple potential functions for simulating liquid water. *Journal of Chemical Physics* **1983,** 79, (2), 926-935.

63. Case DA, D. T., Cheatham TE, Simmerling CL, Wang J, Duke RE, Luo R, Walker RC, Zhang W, Merz KM, Roberts B, Wang B, Hayik S, Roitberg A, Seabra G, Kolossváry I, Wong KF, Paesani F, Vanicek J, Liu J, Wu X, Brozell SR, Steinbrecher T, Gohlke H, Cai Q, Ye X, Wang J, Hsieh MJ, Cui G, Roe DR, Mathews DH, Seetin MG, Sagui C, Babin V, Luchko T, Gusarov S, Kovalenko A, Kollman PA (2010) AMBER 11 University of California *AMBER 11*, University of California, San Francisco: San Francisco, 2010.

64. Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C., Numerical-integration of cartesian equations of motion of a system with constraints - molecular-dynamics of n-alkanes. *Journal of Computational Physics* **1977,** 23, (3), 327-341.

65. Humphrey, W.; Dalke, A.; Schulten, K., VMD: Visual molecular dynamics. *Journal of Molecular Graphics* **1996,** 14, (1), 33-38.

66. Sasahara, K.; Naiki, H.; Goto, Y., Kinetically controlled thermal response of beta(2)-microglobulin amyloid fibrils. *Journal of Molecular Biology* **2005,** 352, (3), 700-711.

67. Meersman, F.; Dobson, C. M., Probing the pressure-temperature stability of amyloid fibrils provides new insights into their molecular properties. *Biochimica Et Biophysica Acta-Proteins and Proteomics* **2006,** 1764, (3), 452-460.

68. Mauro, M.; Craparo, E. F.; Podesta, A.; Bulone, D.; Carrotta, R.; Martorana, V.; Tiana, G.; San Biagio, P. L., Kinetics of different processes in human insulin amyloid formation. *Journal of Molecular Biology* **2007,** 366, (1), 258-274.

69. Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A., Development and testing of a general amber force field. *Journal of Computational Chemistry* **2004,** 25, (9), 1157-1174.

70. Frisch, M. J. T., G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb,; M. A.; Cheeseman, J. R. M., J. A., Jr.; Vreven, T.; Kudin,; K. N.; Burant, J. C. M., J. M.; Iyengar, S. S.; Tomasi, J.;; Barone, V. M., B.; Cossi, M.; Scalmani, G.; Rega, N.;; Petersson, G. A. N., H.; Hada, M.; Ehara, M.; Toyota, K.;; Fukuda, R. H., J.; Ishida, M.; Nakajima, T.; Honda, Y.;; Kitao, O. N., H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H.; P.; Cross, J. B. A., C.; Jaramillo, J.; Gomperts, R.; Stratmann,; R. E.; Yazyev, O. A., A. J.; Cammi, R.; Pomelli, C.; Ochterski,; J. W.; Ayala, P. Y. M., K.; Voth, G. A.; Salvador, P.; Dannenberg,; J. J.; Zakrzewski, V. G. D., S.; Daniels, A. D.;; Strain, M. C. F., O.; Malick, D. K.; Rabuck, A. D.;; Raghavachari, K. F., J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A.; G.; Clifford, S. C., J.; Stefanov, B. B.; Liu, G.; Liashenko,; A.; Piskorz, P. K., I.; Martin, R. L.; Fox, D. J.; Keith, T.;; Al-Laham, M. A. P., C. Y.; Nanayakkara, A.; Challacombe, M.;; Gill, P. M. W. J., B.; Chen, W.; Wong, M. W.; Gonzalez, C.;; Pople, J. A. Gaussian.

71. Tatsis, V. A.; Tsoulos, I. G.; Stavrakoudis, A., Molecular Dynamics Simulations of the TSSPSAD Peptide Antigen in Free and Bound with CAMPATH-1H Fab Antibody States: The Importance of the beta-Turn Conformation. *International Journal of Peptide Research and Therapeutics* **2009,** 15, (1), 1-9.

72. Kuzmanic, A.; Zagrovic, B., Determination of Ensemble-Average Pairwise Root Mean-Square Deviation from Experimental B-Factors. *Biophysical Journal* **2010,** 98, (5), 861-871.

73. Sadiq, S. K.; De Fabritiis, G., Explicit solvent dynamics and energetics of HIV-1 protease flap opening and closing. *Proteins-Structure Function and Bioinformatics* **2010,** 78, (14), 2873-2885.

74. Kabsch, W.; Sander, C., Dictionary of protein secondary structure - pattern-recognition of hydrogen-bonded and geometrical features. *Biopolymers* **1983,** 22, (12), 2577-2637.

75. Zheng, J.; Ma, B. Y.; Tsai, C. J.; Nussinov, R., Structural stability and dynamics of an amyloid-forming peptide GNNQQNY from the yeast prion sup-35. *Biophysical Journal* **2006,** 91, (3), 824-833.

76. Schlick, T., Molecular Modeling. An Interdisciplinary Guide. In [Online] Edition, S., Ed. Springer Verlag: New York, 2010.

77. Shao, J. Y.; Tanner, S. W.; Thompson, N.; Cheatham, T. E., Clustering molecular dynamics trajectories: 1. Characterizing the performance of different clustering algorithms. *Journal of Chemical Theory and Computation* **2007,** 3, (6), 2312-2334.

78. Berhanu, W. M.; Masunov, A. E., Can molecular dynamics simulations assist in design of specific inhibitors and imaging agents of amyloid aggregation? Structure, stability and free energy predictions for amyloid oligomers of VQIVYK, MVGGVV and LYQLEN. *J Mol Model, Published online : 21 Dec.2010; DOI 10.1007/s00894-010-0912-4*.

79. De Simone, A.; Esposito, L.; Pedone, C.; Vitagliano, L., Insights into stability and toxicity of amyloid-like oligomers by replica exchange molecular dynamics analyses. *Biophysical Journal* **2008,** 95, (4), 1965-1973.

80. Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S. H.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E.,

Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. *Accounts of Chemical Research* **2000,** 33, (12), 889-897.

81. Gohlke, H.; Case, D. A., Converging free energy estimates: MM-PB(GB)SA studies on the protein-protein complex Ras-Raf. *Journal of Computational Chemistry* **2004,** 25, (2), 238-250.

82. Chong, L. T.; Pitera, J. W.; Swope, W. C.; Pande, V. S., Comparison of computational approaches for predicting the effects of missense mutations on p53 function. *Journal of Molecular Graphics & Modelling* **2009,** 27, (8), 978-982.

83. Okimoto, N.; Futatsugi, N.; Fuji, H.; Suenaga, A.; Morimoto, G.; Yanai, R.; Ohno, Y.; Narumi, T.; Taiji, M., High-Performance Drug Discovery: Computational Screening by Combining Docking and Molecular Dynamics Simulations. *Plos Computational Biology* **2009,** 5, (10), 13.

84. Ma, B. Y.; Nussinov, R., Simulations as analytical tools to understand protein aggregation and predict amyloid conformation. *Current Opinion in Chemical Biology* **2006,** 10, (5), 445-452.

85. Jiang, P.; Li, W. F.; Shea, J. E.; Mu, Y. G., Resveratrol Inhibits the Formation of Multiple-Layered beta-Sheet Oligomers of the Human Islet Amyloid Polypeptide Segment 22-27. *Biophysical Journal* **2011,** 100, (6), 1550-1558.

86. Berhanu, W. M.; Masunov, A. E., Natural polyphenols as inhibitors of amyloid aggregation. Molecular dynamics study of GNNQQNY heptapeptide decamer. *Biophysical Chemistry* **2010,** 149, (1-2), 12-21.

87. Raman, E. P.; Takeda, T.; Klimov, D. K., Molecular Dynamics Simulations of Ibuprofen Binding to A beta Peptides. *Biophysical Journal* **2009,** 97, (7), 2070-2079.

88. Berhanu, W. M.; Masunov, A. E., Unique example of amyloid aggregates stabilized by main chain H-bond instead of the steric zipper: molecular dynamics study of the amyloidogenic segment of amylin wild-type and mutants. *Journal of Molecular Modeling,Published online: 28; May 2011 DOI 10.1007/s00894-011-1030-7*

CHAPTER 3 STERIC ZIPPER STABILITY IN WILDTYPE AND MUTANTS OF THREE AMYLOID FRAGMENTS

Chapter 3, in part, is a reprint of the material as it appears in Journal of Molecular Modeling,2011, Workalemahu M. Berhanu & Artem E. Masunov, published online: 21 Dec. 2010; DOI 10.1007/s00894-010-0912-4

3.1 Background

Aggregation of polypeptide chains and formation of amyloid fibrils are associated with the development of a number of disorders, including Alzheimer's, Parkinson's, type II diabetes, and Creutzfeldt-Jakob disease. [89] Amyloid deposits develop when proteins misfold out of their native conformations and aggregate into insoluble fibrils.[90] The amyloid fibrils share a sequence independent structure characterized by cross-β spine structural motif in which protein β-strands run orthogonal to the fibril axis and repetitive hydrogen bonding extends parallel to the axis.[91, 92] This cross-β spine may correspond to the global minimum energy conformation for a wide variety of proteins.[91] Identifying this structural motif in small model peptide systems and characterizing it under different conditions can yield valuable clues about the molecular-level details of amyloid formation. Recently, the microcrystal structures of several amyloidogenic peptides have been determined by x-ray crystallography.[22, 93, 94] These high resolution structures provided researchers with a unique opportunity to understand the structural details and on the factors that destabilize/stabilize the amyloid fibrils. Molecular dynamics (MD) simulations,

along with other theoretical approaches, based on these crystal structures, can often present significant contribution to this understanding.[79, 95-103] By selecting an amyloid oligomer out of the crystal structure and evaluating its conformational stability in a crystal-free environment, these investigations have provided insights into the intrinsic propensities of peptide fragments to associate in amyloid-like states, the energetic factors stabilizing these aggregates, and the possible aggregation states of oligomeric precursors or larger assemblies up to 128 β-strands.

One of the common structural features, observed in many available X-ray structures of amyloidogenic polypeptides is pairing of the β-sheets by interdigitated side chains in a dry 'steric zipper'. It is worth noting, that most of the theoretical investigations have been conducted on the systems where steric zipper interface is composed of the large polar and/or aromatic side chains. In this study we focus on aggregates stabilized by steric zipper interfaces formed by small hydrophobic residues (VQIVYK, MVGGVV) (Figure 3-1). A system with polar H-bonding side chains (LYQLEN) (Figure 3-1) is also considered for comparison. We perform all-atom MD simulations with explicit solvent on both wild type and mutant polypeptides at various degrees of aggregation. The initial structure of the aggregates is based high resolution X-ray study.[22] The MVGGVV peptide represent the fragment (residues 35-40) from the C terminal of the A$\beta_{1\text{-}40}$ peptide, associated with Alzheimer's disease.[104, 105] The VQIVYK is a fragment (residues 306-311) of the Tau protein, which is also involved in the pathogenesis of Alzheimer's disease. LYQLEN peptide is a fragment of a chain (residues 13-18) of Insulin that had been shown to form amyloid-like fibrils.[22]

In Alzheimer's disease (AD), the Tau protein forms intracellular amyloid tangles in neurons.[106, 107] The hexapeptide VQIVYK models the key amyloidogenic peptide sequence and

forms amyloid-like fibrils with the same cross-β structure found in full Tau amyloid fibrils.[22] The structural organization of VQIVYK is a parallel β-strand within the same β-sheet layer while maintaining anti-parallel organization between the adjacent β-sheet layers.[22] At the dry interface between the adjacent β-sheet layers, the shape complementarity is formed by the hydrpbobic steric zipper via the side chains of Val1, Ile3 and Tyr5 (Figure 3-2a) [22], packing against each other forming the sheet-sheet interface. Aggregation of Aβ peptides, which are the natural products of cellular proteolytic, is also linked to Alzheimer's disease (AD). The most abundant Aβ species are 40 residue peptides ($A\beta_{1\text{-}40}$). The MVGGVV peptide is a fragment (residue 35-40) from the C terminal of the $A\beta_{1\text{-}40}$ consists of parallel and anti-parallel β-strands within the same β-sheet layers. At the dry interface between the adjacent β-sheet layers, the shape complementarily is formed by the hydrophobic steric zipper via the side chains of Met1, Val2 and Val5 (Figure 3-2b, c). [22]

Fibrils of Insulin are observed extracellularly in the rare medical condition termed injection amyloidosis. These Insulin fibrils formed in vivo display the defining characteristics of amyloid aggregates such as binding the dye Congo red [30] and the cross-β X-ray diffraction pattern.[29] Both A chain and B chain can form fibrils on their own [108, 109], and seeds of A chain or B chain can nucleate the fibrillation of full length Insulin.[108] The atomic-resolution picture of the interactions between segments of Insulin which may be part of fibrillar spine came from crystal structures of the fibril forming peptide segments LYQLEN (residues A13–A18) and VEALYL (residues B12–B17).[22] The structural organization of LYQLEN is anti-parallel β-strands within the same β–sheet layer while maintaining parallel organization between the adjacent β-sheet layers.[22] At the dry interface between the adjacent β-sheet layers, the shape complementarity is

formed by the polar side chain steric zipper (Tyr2, Gln3, Leu4 and Asn6) and side chain H-bonding (Figure 3-2d).[22] Recently, serum samples from patients with Parkinson's disease have been found to display an autoimmune response to Insulin oligomers and fibrils [110], possibly indicating the presence of Insulin aggregates in this disease as well. Insulin also reported to form amyloid-like fibrils in vitro under elevated temperatures, low pH, and increased ionic strength.[111, 112] This fibril formation has been a limiting factor in long-term storage of Insulin for treatment of diabetes. Thus, better understanding of Insulin fibrillation could lead to safer handling and more cost-effective storage of Insulin.

Previous theoretical study has demonstrated the significant role of steric zipper in the structural stability of the GNNQQNY and GGVVIA oligomers stabilized with polar side chain and H-bonding.[102, 113] Park *et al.* [95] address the structural selection mechanism of different double layer peptides including GNNQQNY, NNQQ, VEALYL, KLVFFAE and STVIIE, and find that the patterns with the lowest binding free energy correspond to X-Ray structures with high accuracy. The main contribution of the binding free energy of the double layer pattern is determined by the van der Waals and hydrophobic forces. These contributions can therefore serve as a quantitative measure of shape complementarity among side chains between the *β*-sheets. The steric self-complementary (known as steric zipper) selects the most stable packing modes. It also makes parallel *β*-sheets generally preferred over anti-parallel ones. The presence of charged side chains appears to give anti-parallel *β*-sheets kinetic preference at the early stages of assembly, while the double layer formation is likely to be thermodynamically controlled. Xu *et al.* [114] investigated the β-sheets composed of seven antiparallel decapeptides, representing the 20–29 segment of human Islet amyloid polypeptide (hIAPP). The amyloid nucleus of hIAPP was

mimicked with one β-sheet of different initial separation distances between the strands. Multiple all-atom MD simulations with explicit water solvent showed that the assembly occurs not only in the lateral direction but also along the longitudinal direction. This provides a new insight into the assembly pathway at the early stage of fibril elongation. Based on the Poisson–Boltzmann free energy analysis and quasiharmonic configuration entropy estimation, the entropic contribution was found to play an important role in the longitudinal assembly. Moreover, a possible oligomeric state with cyclic form was suggested based on one assembly model found in the simulations. This evidenced the polymorphic nature of the amyloidogenic oligomerization and possible mechanism of its toxicity. The cyclic structures of amyloid oligomers have been reported to be the early intermediates in solution, capable to form ion-channel-like structures in the membrane that could be responsible for pathologic membrane permeability and destabilization of the cellular ionic homeostasis.[115, 12]

Vitagliano *et al* [116] in their molecular dynamics simulation characterizing assemblies formed by steric zipper assemblies composed of a pair of 10-stranded β-sheets of the peptides SSTSAA and VQIVYK show high fluctuations and significant distortion. The analysis of the VQIVYK crystal packing reveals two different double layers with significant interface area and surface complementarity.[22] One is characterized by nonpolar dry interface made up essentially by the side chains of V1 and I3 of the two layers, while the other is polar and involves Tyr and Gln side chains.[22] The nonpolar interface exhibits larger values of the surface area (113 vs. 89 $Å^2$), but slightly lower surface complementarity (0.76 *vs.* 0.82).[22] The stability of the nonpolar hydrophobic interfaces was studied by Vitagliano *et al.* [116] in their MD simulations. They report

high fluctuations and significant distortion (RMSD *c.a.* 10 Å within 40 ns simulation) when investigate three layer assemblies formed by steric zipper and composed of a pair of 10-stranded β-sheets of the peptides VQIVYK. In the contrast, they found RMSD below 6 Å within 40 ns simulation, when study the 10-stranded double layer with nonpolar interface. Hence, the stability of the nonpolar interface is system dependent.

However, the atomic information for the early stage of the aggregation mechanism of the VQIVYK, MVGGVV and LYQLEN peptide is still limited so far. Thus, understanding the structural stability and aggregation behavior of the VQIVYK, MVGGVV and LYQLEN peptide is expected to provide knowledge for designing an inhibitor aimed to decrease the self-aggregation into fibrils.

In this study, several all-atom MD simulations with explicit water at 300 K were conducted to investigate the structural stability, aggregation behavior and thermodynamics of the VQIVYK, MVGGVV and LYQLEN peptides with various sizes and its single glycine replacement mutations. Our aim is to elucidate: (i) the influence of the number of the peptides on the structural stability and conformational dynamics of the oligomers; (ii) the possible minimal nucleus seed for the fibril formation of the peptides; (iii) the principle driving force for the association of the peptides; and (iv) the effects of single glycine replacement mutations on the structural stability of the oligomers. The results of this study may provide insight into the possible mechanism of fibrillogenesis of the amyloid peptides. It may also be helpful for designing new or modified capping peptides capable of breaking the driving force for aggregations and preventing the fibril formation of the peptides.

3.2 Methods

The crystal structure of the VQIVYK, MVGGVV and LYQLEN had been determined by Sawaya *et al*.[22] The atomic coordinates of the multiple unit cells were taken from the website [117], and the water molecules from the crystal structure were removed. The MVGGVV have two different polymorphic forms (form 1 and 2 with resolution of 2.0 Å and 1.8 Å) both of which were used in the simulation. The Sirius visualization program from San Diego Supercomputer Center (http://sirius.sdsc.edu) was used to construct the aggregates of various sizes. The initial geometry of the largest aggregate was taken as a pair of β-sheets composed of 6 strands (5 strands for VQIVYK), it is shown on Figure 3-2. In the following we denote the aggregates ShN-StM, where N is the number of β-sheets, and M is the number of strands per β-sheet. The initial geometry of the largest wild type aggregate was taken as a pair of β-sheets composed of 6 strands (MVGGVV and LYQLEN) and 5 strands (VQIVYK), as shown on Figure 3-2. For the smaller size wild type systems, the initial structures of oligomers were obtained by removing the β-strands one by one from the Sh2-St5 (VQIVYK) or Sh2-St6 (MVGGVV and LYQLEN) models. To construct the mutant systems, several glycine replacements were made in the wild type aggregate. The mutants are denoted as XnG, where X is the replaced residue, and n is its position in the peptide sequence. Three or four mutants were designed for each peptide (**V1G**, I3G, and Y6G for VQIVYK; M1G, V2G, V5G, and V6G for MVGGVV; V2G, Q3G, L4G and V6G for LYQLEN). The simulation details for each model are summarized in Table 3-1, 3-2, 3-3 and 3-4. The MM-PBSA single trajectory approach implemented as script in Amber10. [118] was used to calculate the steric zippers binding energy for non-covalent association between the β-

sheets within the double layer. The gas phase and the solvation free energies were calculated over 500 snapshots taken at 20 ps intervals from the last 8 ns of the MD trajectories.

3.3 Results

3.3.1 Size dependent structural stability of the wild type peptides aggregates

Eight simulations of wild type VQIVYK were conducted for the aggregates build of one (models A1-A4) and two (models A5–A8) antiparallel β-sheets with parallel strands within each sheet. The relative stability of the model aggregates was measured by the backbone root mean-squared deviation (RMSD). The reference structure for calculating backbone RMSD was the energy-minimized structure. As one can see on Figure 3-3A, for the model systems of A1 (Sh1-St2) and A2 (Sh1-St3), the RMSDs remained below 2.Å for 10 ns, while for A3 (Sh1-St4) and A4 (Sh1-St5) the RMSDs increased to 4.5Å, indicating the lower relative instability of the one layer aggregate with larger number of strands. The larger two-layer model systems of A7 (Sh2-St4) and A8 (Sh2-St5), maintained RMSDs *c.a.* 4.0 Å within 10 ns, indicating relative stability of the structures compared to the smaller bilayer models A5 (Sh2-St2) and A6 (Sh2-St3), which showed large fluctuations up to 7.0 Å (Figure 3-4A)**.** The results of two-layer models suggested that the structural stability of the VQIVYK oligomers increases with increasing the numbers of β-strands, the four and five stands are more stable than two and three strands, while for one-layer models the trend is opposite.

Our simulation for 5-stranded double layers of the wildtype VQIVYK oligomers was found to have a RMSD of 4 Å, in good agreement with the result reported by Vitagliano *et al.*[116]

The comparison of the RMSD values of the nonpolar interface models (VQIVYK and MVGGVV**)** with LYQLEN that has polar residues on the dry interface indicates the nonpolar are significantly less stable. The smaller RMSD values of the polar LYQLEN is in a good agreement with the result of Zhang *et al.*[113], who found an RMSD of 2 Å by simulation of 4-stranded double layer GGNNQQNY, which has polar residues on the dry interlayer interface. Our results indicate that the polar dry interface significantly improves stability.

Another eight wild type simulations of MVGGVV1 (models C1–C8) were conducted for anti-parallel β-sheets with parallel strands within the sheets. As shown in Figure 3-3B, for the model systems of C1 (Sh1-St2) and C2 (Sh1-St3), the RMSDs were below 4.50 Å within 10 ns. For C3 (Sh1-St4) and C4 (Sh1-St5) the RMSDs were maintained below 4.5 Å and 6 Å within 10 ns respectively, the two layer model systems of C7 (Sh2-St4) and C8 (Sh2-St5), the RMSDs were below 3.0 Å within 10 ns as shown in Figure 3-4B. Aggregate C6 (Sh2-St2) maintained RMSDs below 4.0 and C5 (Sh2-St3) showed large fluctuations RMSD within the first 5 ns and then increased to 12 Å after 7 ns. Our results for one-layer models suggest that the structural stability of the MVGGVV1 oligomers increases as the number of stands decreases, while the results of two-layer models suggest that the structural stability of the MVGGVV1 oligomers increases remarkably with increasing the numbers of β-strands, the four and five stands are more stable than two and three strands.

Four wild type MVGGVV2 peptide aggregates (two layer with different number of strands), simulations were conducted for antiparallel β-sheets with parallel strands within the sheets (models E1–E4). We did not do single layer simulation, assuming the result will be the same as for polymorphic form I. As shown in Figure 3-4C, E2 (Sh2-St3) and E3 (Sh2-St4), the

RMSDs were almost identical and remained < 6.0 Å within 10 ns. While E1 (Sh2-St2) showed a large fluctuations within 1ns and remained c.a. 7.0 Å within 10 ns**.** E4 (Sh2-St5) shows the same RMSD change as E2 and E3 for the first 4ns increasing to ~5.0 Å and maintained an RMSD ≤ 5.0 Å during the 10 ns simulation. The results of two-layer models suggested that the structural stability of the MVGGVV2 oligomers increases with increasing the numbers of β-strands, the four and five stands are more stable than two and three strands.

Finally eight wild type LYQLEN peptide aggregates (models G1–G8) were considered for antiparallel β-sheets, parallel strands within the sheets. As shown in Figure 3-3C, for the model systems G1 (Sh1-St2), G2 (Sh1-St3), G3 (Sh1-St4) and G4 (Sh1-St5) consisting of one layer and different number of strands, the RMSDs remained at 2.0 Å within 10 ns, indicating exceptional stability of these structures. Figure 3-4D **s**how that for the model systems G5 (Sh2-St2) and G7 (Sh2-St4), the RMSDs showed were maintained at ~4.50 Å, for G8 (Sh2-St5) RMSDs is c.a. 2.50 Å, and for G6 (Sh2-St3) RMSD demonstrates a large fluctuation within the first 4 ns and then stabilized at 7 Å after 8 ns, which indicated that they lost their original structural organization. The results of two-layer models suggested that the structural stability of the LYQLEN oligomers increases remarkably with increasing the numbers of β-strands, with four and five stands being the most stable. Our results for one-layer models suggested that the structural stability of the LYQLEN oligomers is the same irrespective of the number of strands. One layer with two, three, four and five stands (model G1-G4) are structurally stable with RMSD *c.a* 2.0 Å. The LYQLEN oligomers is stabilized with backbone to backbone and side

chain hydrogen bonding while in the case of the VQIVYK , MVGGVV1 and MVGGVV 2 are stabilized with back bone to back to bone hydrogen bonding.

Comparing single and double layer models, our results reveal that the extra β-sheet strand contributes significantly to the structural stability of the VQIVYK, MVGGVV1 and MVGGVV2 oligomers for double layer model while in the case of single layer model it decrease. In the case of the LYQLEN our results also shows that an extra β-sheet strands contributes significantly to the structural stability of the LYQLEN oligomers for double layer models while in the case of single layer model it is almost the same irrespective of the number of strands. This is in agreement to previous studies done on different types of amyloid models. The above results all together suggest that the structural stability of the oligomers increases significantly with increasing the number of β-strands for double layer models implying extra sheet-sheet interactions are necessary for the formation of steric zipper to associate the strands, resulting in more stable oligomeric organizations. Our findings are in agreements with previous observations [102, 113] which indicate that the minimal nucleus seed for the amyloid fibril formation could be as small as three or four peptides.

The evolution of the root mean square deviations (RMSDs) between initial and the current trajectory structures indicates that the system undergoes certain rearrangement. The initial structures are taken from X-ray and may be stabilized by the crystal environment. However, the simulations are performed in the solution state. Due to this different environment, relatively large RMSDs may not always correspond to the unstable structures. To further analyze the structural stability we also performed secondary structure analysis and binding free energy calculations.

3.3.2 The effect of single-glycine mutations on structural stabilities of the aggregates

To investigate how the steric zipper interfaces influence the structural stability of the double-sheet aggregates of VQIVYK, MVGGVV and LYQLEN peptides, the side-chains participating in these interfaces were replaced by glycine (Table 3-1 to 3-4). As one can see from the Figure 3-5A, the largest aggregates composed of these mutants were less stable, compared to the respective wild type aggregates. The RMSD of I3G and Y5G are somewhat higher (c.a 5.5Å), than those of V1G (c.a. 4.5Å), indicating that I3G and Y5G exhibit higher potential to destabilize the structure of the VQIVYK aggregate. Mutations of the nonpolar side chain Ile-3 or Tyr-5 to Gly were found to result in destabilization of the oligomeric structures. Figure 3-5B indicates that none of MVGGVV1 mutants were structurally stable compared to the value of the respective wild type model. It shows that the RMSD of M1G and V6G are significantly higher (< 8.5 Å) than those of V2G and V5G (< 5.5.Å), indicating that V2G and V5G exhibit higher potential to destabilize the structural integrity of the MVGGVV1 oligomer. The result also (Figure 3-5B**)**, shows that mutation of the non polar side chain Met-1 or Val-6 to Gly negatively affects the intersheet steric zipper destabilizing the structural integrity of the MVGGVV1 oligomers to a greater extent than the V2G and V5G mutants and the wild type. Figure 3-5C shows that some of MVGGVV2 mutants (V2G and V6G) were structurally stable compared to the wild type model. It shows that the RMSD of the mutant V2G and V6G are lower (nearly 4.0 Å) than those of M1G and wild type (nearly 6.0 Å), indicating that V2G and V6G exhibit higher potential to stabilize the structural integrity of the MVGGVV2 oligomer. The result also shows (Figure 3-5C**)**, that Y5G (RMSD > 15.0 Å) destabilize the structural integrity of the MVGGVV2 oligomers to a greater extent than the other mutants and the wild type. Our finding that the

MVGGVV2 wild type aggregates are less stable compared to certain mutants is in contrast to other oligomers in our study and to conclusions of the previous work done on the hexapeptide amyloid.[102, 113] This apparent contraction could be explained based on the structural difference between this particular polymorph and other amyloid X–ray structure in that there is 90^0 bending in the upper sheet of MVGGVV form 2.[22] As can be seen in Figure 3-5D, none of the mutants of LYQLEN are as structurally stable as the wild type, indicating that the side chain interactions play an important role in determining the stability of the LYQLEN oligomers. However, the N6G mutant have small RMSD values (~2.5 Å), whereas the Y2G, Q3G and L4G mutants have large RMSD values (> 4.0Å). Comparison between the dynamics of the wild type and its mutants suggests that mutations N6G have little effect on the structural stability of the LYQLEN (low RMSD), whereas mutations Y2G, Q3G and L4G destabilize the oligomeric structures. The destabilization of the Q3G and L4G mutants is even more pronounced. As seen in Figure 3-5D mutations of the polar side chain Gln-3, or Leu-4 to Gly would affect the inter sheet steric zipper, leading to greater destabilization of the oligomeric structures.

3.3.3 Secondary Structure Assessment

We analyzed the secondary structure of the oligomers using the DSSP algorithm written by Kabsch and Sander.[74] This algorithm is based on identification of hydrogen-bonding (H-bonding) patterns and recognizes seven types of secondary structures which can be grouped into three classes: helix (α-helix, 3_{10}-helix, π-helix), β-strand (isolated β-bridge, extended β-sheet) and loop (turn, bend). The result of the secondary structure analysis for a two layers models of the amyloid peptides show that the wild type of Tau (St2-St5 VQIVYK), Insulin (St2-St6

LYQLEN) and Aβ amyloids Polymorphic form I and II (Sht2-St6 MVGGVV) appear to be stable at 300K, which is confirmed by the conservation of high anti-parallel β-sheet content throughout the whole simulation time (as shown in Figures 3-6A-D). In all systems the contents of antiparallel β sheets are much greater than the turn or parallel β-sheet indicating the conservation of the original structure. The results also indicate that antiparallel β-sheet, parallel β-sheet and turn content were preserved throughout the 10 ns simulation. The analyses of the secondary structure evolution throughout the simulation for the mutant form of the amyloids are shown in Figures 3-6 to 3.9.

The results for Tau mutant and the wild type the secondary structure is shown in Figure 3-6A. In the case of the Tau mutants the content of the secondary structure declined specially in the last 3ns of the simulation. This result is in agreement with the RMSD results discussed above. The decline is largest in I3G which has the largest RMSD value. Figure 3-7B shows the results for Aβ peptide polymorph form I for both the wild type and mutants. The result for the wild type shows that the secondary structure is preserved as indicated with its high and content of antiparallel β-sheets (~0.7). Among the mutants the content of the secondary structure was more unstable for M1G indicating the greater destabilizing effects of replacing methione with glycine. This is in agreement with the RMSD result, the highest RMSD among the mutants was from the M1G mutant (see Figure 3-6B). Figure 3-8C shows the results for Aβ peptide polymorph form II for both the wild type and mutants. The result for the wild type shows that the secondary structure is preserved as indicated with its high and content of antiparallel β-sheets (~0.7). Among the mutants the content of the secondary structure was more unstable for M1G indicating the greater destabilizing effects of replacing methione with glycine. This is in

agreement with the RMSD result, the highest RMSD among the mutants was from the M1G mutant (see Figure 3-8B).

The results for Insulin amyloid mutants (Figure 3-9D) and the wild type shows that the secondary structure is preserved as indicated with its high content antiparallel β–sheets (~0.7). In the case of the Q3G and Y2G mutants the content of antiparallel β-sheets declined specially in the last 2 ns of the simulation. This result is in agreement with the RMSD results. The RMSD for both Q3G and Y2G mutants was the largest ~5 Å.

Two trajectory snapshots (at 5 ns and 10ns) are shown in Figures 3-10 to 3-13 for each of the two layer oligomer aggregates. As the structure evolves, some of the terminal strands break the β-sheet ordering and twist relative to the remaining strands although do not dissociate from the aggregate completely. Degree of this disorder correlates with the RMSD values reported on Figures 3-3 to 3-5. Among the most disordered structures are mutant I3G of the Tau fragment (VQIVYK), mutants V6G and M1G of the Aβ fragment polymorph 1 (MVGGVV1) mutant V5G of the Aβ fragment polymorph 2 (MVGGVV2) and the mutants Y2G and Q3G of the Insulin amyloid (LYQLEN). Apparently, the mutated amino acids were involved in the steric zippers, which were not holding the β-sheets together, but also preserving them from disaggregation. Inversely, the complementarity of the amino acids sidechains would be essential for the formation of the ordered aggregate. On the other hand, disordered random aggregation may take place for any polypeptide studied in this work, as suggested by the negative values of association free energies, reported in the next section.

3.3.4 Free energy calculations

The binding free energies were calculated with the Molecular Mechanics Poisson-Boltzmann solvent accessible surface area (MM-PBSA) model [80], as implemented in AMBER. In this method, the total binding free energy in water is approximated by $\Delta G_{TOT} = \Delta E_{GAS} + \Delta G_{PB} + \Delta G_{SUR}$. The ΔE_{GAS} is the gas phase interaction energy. The $\Delta G_{PB/GB}$ is the polar part of the solvation free energy represented by Poisson-Boltzmann approaches. The ΔG_{SUR} is the surface area term, approximating the non-polar part of the solvation free energy. In this formula, the conformational entropy of the solute is not considered, while the solvent entropy is implicitly considered in the ΔG_{PB} and ΔG_{SUR}. Although the MM-PBSA calculations may overestimate the absolute binding free energy due to the missing terms (e.g., conformational entropy change of the solute upon binding), they usually give a reasonable estimate on the relative binding free energy when the conformational entropy changes of two binding modes are comparable [119].

The binding energy was calculated by MM-PBSA method and is specified in method section. The breakdown of binding energy components is listed in Table 3-4 to 3-7. The MM-PBSA analysis allows us to separate the total free energy of binding into electrostatic, van der Waals interactions and solute-solvent interactions, and thereby gain additional insights into the monomer to monomer association process in the formation of the dimer of the amyloid oligomers. As shown in Table 3-4 and 3-7**,** van der Waals interactions play a very important role in the simulation, contributing significantly more to the total interaction energy than the electrostatic interaction for the Tau (VQIVYK) and Insulin (LYQLEN) aggregates. Nonpolar solvation energies favor the binding and the polar solvation energies disfavor it. In the case of Aβ oligomers, MVGGVV, electrostatic interaction play a very important role in the simulation,

contributing significantly more to the total interaction energy than the van der Waals interactions (Table 3-5 and 3-6). Nonpolar solvation energies favor the binding and the polar solvation energies disfavor it. Note that the internal component of the molecular-mechanical energy (bond, angle, and torsional energies) has zero contribution to the binding free energy, because the structures of the monomers in its unbound and bound states were assumed to be the same (data not shown). Table 3-4 to 3-7 also reports the contributions of apolar/hydrophobic and polar/electrostatic to the free energy. We found that the predicted binding free energy is dominated by the magnitudes of the apolar components ($\Delta E_{vdw} + \Delta E_{sur}$) in all the four oligomers and their mutants. In contrast, the polar interactions ($\Delta E_{elec} + \Delta G_{PB}$) show less contribution to the binding free energy. The result of the binding free energy calculation also indicated that the wild type is the most stable structure compared to the mutants. From the negative total binding free energy of the wild types we clearly see that this is a favorable protein-protein complex in pure water. The mutants also form a stable complex based on the negative total binding free energy. However, the mutant complex is less thermodynamically favorable than the wild type complex. The calculation indicated that the mutation of bulky polar side chain from the steric zipper structure leads to the less stable dimer (see example Table 3-4 to 3-7) giving mutant with smallest binding free energy, and indicating that these residues are important for stabilizing the structure.

The widely accepted hypothesis on the amyloid disease is that soluble protein oligomers are the source for toxicity and are the primary pathogenic factor in these diseases and thus small molecules that prevent or reverse protein oligomerization may provide a mechanism to target the actual cause of the disease [120, 121]. Peptidomimetics are promising class of small molecules

capable of inhibiting oligomerization. Most fibrillogenesis inhibitors of this type were designed rationally based on molecular recognition elements found in the site of aggregation.[122] Identification of this aggregation site is often based on the mutational data. Such data helps to pinpoint the small regions on the protein interaction interface that are responsible for a disproportionate contribution to the binding energy of the two proteins.[123-125] In this work we have shown that most mutations at the aggregation site reduce the binding free energy and weaken the aggregation. Therefore, the computational studies can serve the same purpose of the rational design, as experimental mutation studies.

Another potential application of the presented approach is design of imaging agents. The progress in therapeutic agents for treatment of neurodegenerative amyloid diseases calls for development of more specific biomarkers to detect early stages of amyloid diseases.[126] Design of peptidomimetics based on the data obtained in the molecular dynamic simulation may provide the starting point for design of specific aggregation inhibitors drugs and diagnostic agents. Both structural and thermodynamic results reported in this study illustrate the higher fluctuation in RMSD values and less negative binding free energies for the mutated peptides. These mutants, therefore, may serve as aggregation inhibitors pending the experimental confirmation.

3.4 Conclusions

The major findings of this study can be summarized as follows:

1. The stability of the VQIVYK, MVGGVV and LYQLEN peptides oligomers increases with increasing the number of β-strand;

2. The Sh2-St4 model acts as a stable seed in prompting amyloid fibril formation for all the cases considered;

3. The binding energy calculated by MM-PBSA method and the analysis of individual contributions to the binding energy shows the hydrophobic interactions play an important role in stabilizing the structural organizations between β-sheet layers in the oligomers. The result of the binding free energy calculation also indicated that the wild type is the most stable structure compared to the mutants;

4. The hydrophobic steric zipper on the intersheet interface contributes significantly to the stability of the entire aggregate structures. Mutations of the side-chains participating in the steric zipper interfaces of the oligomeric (VQIVYK, MVGGVV1 and LYQLEN) peptides to Gly resulted in decline of secondary structure content compared to corresponding wild type indicating that the role of the replaced amino acid in stabilizing the structure;

5. A single glycine substitution at the steric zipper interface disrupts the hydrophobic steric zipper remarkably, indicating that the hydrophobic attraction is a major driving force for stabilizing and aggregation of oligomers. Consequently, the substantial reduction in the van der Waals intersheet interactions leads to destabilization of the oligomers. Overall, aggregation of both wild type and mutant peptides is driven by nonpolar interaction.

Some evidence from the experimental work suggests that short peptides may share similar intermolecular interactions to their parent proteins while forming amyloid fibril.[127] Thus, exploring the structural stability and aggregation behavior of the short peptides may gain insights into the self-assembly process at the early stage of fibril formation and provide a clue to understand the possible aggregation mechanism of their parent proteins. The hexapeptide

NFGAIL, a fragment truncated from human islet amyloid polypeptide (hIAPP, residues 22–27), is one of the shortest fragments that have been shown to form amyloid fibrils similar to those formed by the full polypeptide [127] and the fibrils are cytotoxic toward the pancreatic cell line. Therefore, this hIAPP ‘‘amyloid-core’’ peptide has been used as a simplified model system to facilitate the discovery of key factors underlying amyloid fibril formation and the development of anti-amyloid agents. Porat *et al* [128] showed that whereas the NFGAIL was a minimal fibril forming fragment from hIAPP with Tyr substituted for Phe (*i.e.* ^{22}NFGAILSS^{29} to ^{22}NYGAILSS^{29} did not form fibrils by itself and even inhibited fibril formation. Along these lines one can envision a possible strategy to inhibit the formation of early aggregates that includes the design of specific inhibitor, breaking the hydrophobic steric zipper observed in the structure of hydrophobic region of the amyloid aggregate. Proof of principle for such a strategy has been published recently.[129,130] Thus, designing new peptidomimetic inhibitors able to prevent the fibril formation based on the steric zipper motif of the oligomers, similar to the ones examined in this study may become a viable therapeutic strategy. The peptidomimetic approach can also be implemented in designing specific biomarkers for early stage detection of aggregate formation.[130,131]

Table 3-1 Summary of the VQIVYK oligomer models and simulation system

Model	Systems	Sheet/strand organization	Simulation box size (Å)	Simulation time (ns)	T(K)
Wilde type					
A1 *(Sh1-St2)*	sheet1, strands2	----/parallel	49.31× 49.31 ×49.31	10	300
A2 (*Sh1-St3*)	sheet1, strands3	----/parallel	50.84×50.84×50.84	10	300
A3 (*Sh1-St4*)	sheet1, strands4	----/parallel	51.67×51.67×51.67	10	300
A4 (*Sh1-St5*)	sheet1, strands5	----/parallel	54.06×54.06×54.06	10	300
A5 (*Sh2-St2*)	sheet2, strands2	Antiparallel / Parallel	60.97×60.97×60.97	10	300
A6 (*Sh2-St3*)	sheet2, strands3	Antiparallel / Parallel	63.36×63.36×63.36	10	300
A7 (*Sh2-St4*)	sheet2, strands4	Antiparallel / Parallel	64.84×64.84×64.84	10	300
A8 (*Sh2-St5)*	sheet2, strands5	Antiparallel / Parallel	65.88 ×65.88×65.88	10	300
VQIVYK (*Sh2-St5*)	sheet2, strands5	Antiparallel / Parallel	65.88 ×65.88×65.88	10	300
Mutants					
B1	sheet2, strands5, V1G	Antiparallel / Parallel	65.74×65.74×65.74	10	300
B2	sheet2, strands5, I3G	Antiparallel / Parallel	65.71×65.71×65.71	10	300
B3	sheet2, strands5, Y5G	Antiparallel / Parallel	65.26×65.26×65.26	10	300

Table 3-2 Summary of MVGGVV1 oligomer models and simulation system

Model	Systems	Sheet/strand organization	Simulation box size (Å)	Simulation time (ns)	T(K)
Wilde type					
C1 (*Sh1-St2*)	sheet1, strands2	----/antiparallel	52.59×52.59 ×52.59	10	300
C2 (*Sh1-St3*)	sheet1, strands3	----/antiparallel	53.46×53.46×53.46	10	300
C3 (*Sh1-St4*)	sheet1, strands4	----/antiparallel	54.14×54.14×54.14	10	300
C4 (*Sh1-St5*)	sheet1, strands 5	----/antiparallel	56.05×56.05×56.05	10	300
C5 (*Sh2-St2*)	sheet2, strands2	Antiparallel / Antiparallel	63.02×63.02×63.02	10	300
C6 (*Sh2-St3*)	sheet2, strands3	Antiparallel /Antiparallel	63.17×63.17×63.17	10	300
C7 *Sh2-St4*)	sheet2, strands4	Antiparallel /Antiparallel	66.36×66.36×66.36	10	300
C8 (*Sh2-St5*)	sheet2, strands5	Antiparallel / Antiparallel	67.13 ×67.13×67.13	10	300
MVGGVV1 (*Sh2-St6*)	sheet2, strands6	Antiparallel / Antiparallel	69.17 ×69.17×69.17	10	300
Mutants					
D1	sheet2, strands6, M1G	Antiparallel / Antiparallel	69.16 ×69.16×69.16	10	300
D2	sheet2, strands6, V2G	Antiparallel / Antiparallel	69.02×69.02×69.02	10	300
D3	sheet2, strands6, V5G	Antiparallel / Antiparallel	69.45×69.45×69.45	10	300
D4	sheet2, strands6, V6G	Antiparallel / Antiparallel	68.78 ×68.78×68.78	10	300

Table 3-3 Summary of the MVGGVV2 models and simulation system

Model	Systems	Sheet/strand organization	Simulation box size (Å)	Simulation time (ns)	T(K)
Wilde type					
E1 (*Sh2-St2*)	sheet2, strands2	Antiparallel /Antiparallel	60.58 ×60.58×60.58	10	300
E2 (*Sh2-St3*)	sheet2, strands3	Antiparallel /Antiparallel	61.70 ×61.70×61.70	10	300
E3 (*Sh2-St3*)	sheet2, strands4	Antiparallel /Antiparallel	63.24 ×63.24×63.24	10	300
E4 (*Sh2-St5*)	sheet2, strands5	Antiparallel /Antiparallel	64.50 ×64.50×64.50	10	300
MVGGVV2 (*Sh2-St6*)	sheet2, strands6	Antiparallel /Antiparallel	66.83 ×66.83×68.83	10	300
Mutants					
F1	sheet2, strands6, M1G	Antiparallel /Antiparallel	64.15 ×64.15×64.15	10	300
F2	sheet2, strands6, V2G	Antiparallel /Antiparallel	68.86×68.86×68.86	10	300
F3	sheet2, strands6, V5G	Antiparallel /Antiparallel	66.40×66.40×66.40	10	300
F4	sheet2, strands6, V6G	Antiparallel /Antiparallel	66.25 ×66.25×66.25	10	300

Table 3-4 Summary of the LYQLEN oligomer models and simulation system

Model	Systems	Sheet/strand organization	Simulation box size (Å)	Simulation time (ns)	T(K)
Wilde type					
G1(*Sh1-St2*)	sheet1, strands2	----/antiparallel	50.43×50.43 ×50.43	10	300
G2(*Sh1-St3*)	sheet1, strands3	----/antiparallel	51.95×51.95×51.95	10	300
G3(*Sh1-St42*)	sheet1,strands 4	----/antiparallel	51.93×51.93×51.93	10	300
G4(*Sh1-St5*)	sheet1, strands5	----/antiparallel	55.75×55.75×55.75	10	300
G5(*Sh2-St2*)	sheet2, strands2	Antiparallel /Antiparallel	65.67×65.67×65.67	10	300
G6(*Sh2-St3*)	sheet2, strands3	Antiparallel /Antiparallel	66.97×66.97×66.97	10	300
G7(*Sh2-St4*)	sheet2, strands4	Antiparallel /Antiparallel	68.59×68.59×68.59	10	300
G8(*Sh2-St5*)	sheet2, strands5	Antiparallel /Antiparallel	69.82 ×69.82×69.82	10	300
LYQLEN(*Sh2-St6*)	sheet2, strands6	Antiparallel /Antiparallel	70.46×70.46×70.46	10	300
Mutants					
Y2G	sheet2, strands6, Y2G	Antiparallel /Antiparallel	70.04 ×70.04×70.04	10	300
Q3G	sheet2, strands6, Q3G	Antiparallel /Antiparallel	70.35×70.35×70.35	10	300
L4G	sheet2, strands6, L4G	Antiparallel /Antiparallel	70.23×70.23×70.23	10	300
N6G	sheet2, strands6, V6G	Antiparallel /Antiparallel	69.35×69.35×69.35	10	300

Table 3-5 Binding free energy components calculated with MM-PBSA for the wildtype and mutants of VQIVYK oligomer of tau peptide (Sh2-St5)

Energy (Kcal/mol)	Wild type Mean ±std	Mutant -V1G Mean ±std	Mutant -I3G Mean±std	Mutant –Y5G Mean±std
$\Delta E_{elec.}$	446.47±24.29	396.40±49.64	379.36±22.97	396.29±39.45
ΔE_{vdw}	-101.84±4.74	-85.05±4.29	-89.53±4.29	-80.01±4.47
ΔE_{gas}	344.62±24.30	311.35±50.17	289.83±22.37	316.28±38.48
ΔG_{PB}	-404.42 ±23.24	-359.20±45.04	-331.30±19.30	-370.80±38.51
ΔG_{sur}	-15.65±0.34	-14.05±0.48	-14.78±0.43	-13.68±0.33
ΔG_{polar}	42.05±7.58	37.20±10.03	48.06±9.13	25.48±8.44
$\Delta G_{non-polar}$	-117.49±5.08	-99.1±4.77	-104.3±4.72	-93.69±4.80
ΔG_{TOT}	-75.44±6.22	-61.91 ± 9.29	-56.25±8.34	-68.20±7.04

* E_{vdw}, and E_{elec} are the van der Waals and electrostatic binding terms. ΔG_{PB} and ΔG_{sur} are the solvation energies of polar and nonpolar residues, calculated by Amber 10 using the Generalized Born model. ΔG_{polar} and $\Delta G_{non-polar}$ are the sums of polar energy (ΔE_{elec}+ ΔG_{PB}) and non-polar energy components (ΔE_{vdw}+ ΔG_{sur}), respectively. ΔG_{TOT} is the sum of ΔG_{polar} and $\Delta G_{non-polar.}$ ΔG_{TOT} (the binding free energy can also be obtained using the equation, $\Delta G_{TOT} = \Delta E_{GAS} + \Delta G_{PB} + \Delta G_{SUR}$.

Table 3-6 Binding free energy components calculated with MM-PBSA for wildtype and mutants of MVGGVV1 oligomer of Abeta peptide (Sh2-St6)

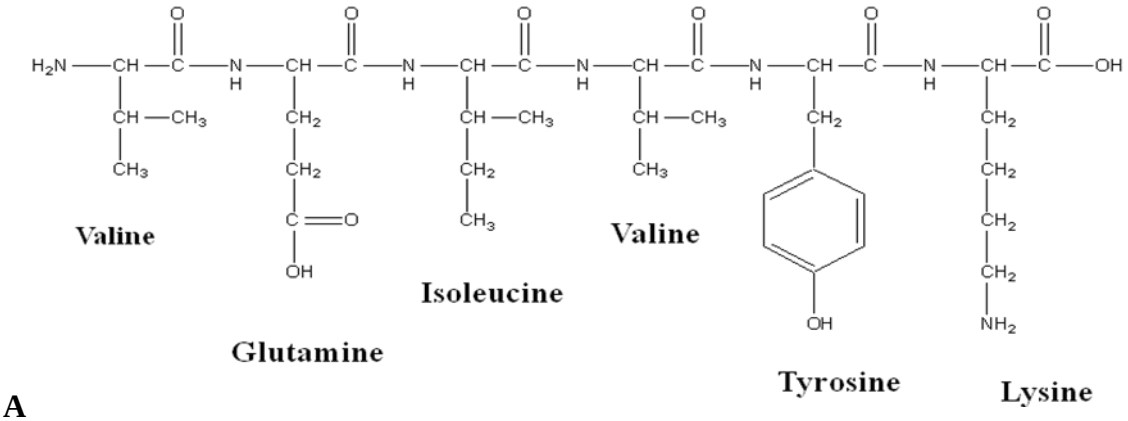
Energy (Kcal/mol)	Wild type Mean ±std	Mutant- M1G Mean ±std	Mutant -V2G Mean ±std	Mutant -V5G Mean±std	Mutant -V6G Mean±std
$\Delta E_{elec.}$	-214.20±16.40	-200.00±25.12	-262.67±30.24	-187.40±21.12	-239.17±29.63
ΔE_{vdw}	-94.88±5.58	-74.93±6.31	-76.93±4.85	-78.13±4.56	-81.06±5.58
ΔE_{gas}	-309.09±15.72	-275.62±24.88	-339.61±28.93	-265.53±20.19	-320.23±29.43
ΔG_{PB}	230.47± 55.68	222.67± 23.68	265.34±26.56	199.03±18.34	244.04±28.33
ΔG_{sur}	-14.97±0.45	-12.97±0.76	-13.20±0.47	-13.19±0.49	-13.65±0.63
ΔG_{polar}	16.27±52.64	21.98±8.22	2.66±8.03	11.63±6.42	4.87±7.37
$\Delta G_{non\text{-}polar}$	-109.85±6.03	-87.90±7.07	-90.13±5.32	-91.32±5.05	-94.70±6.21
ΔG_{TOT}	-93.58±53.51	-65.91±6.36	-87.47±6.04	-79.69±5.17	-89.83±7.18

Table 3-7 Binding free energy calculated with MM-PBSA for the wildtype and mutants of MVGGVV2 of Abeta peptide (Sh2-St6)

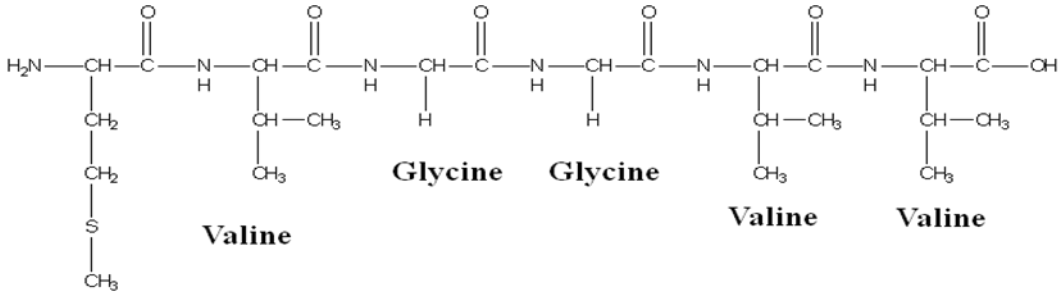
Energy (Kcal/mol)	Wild type Mean ±std	Mutant -M1G Mean ±std	Mutant- V2G Mean ±std	Mutant -V5G Mean±std	Mutant- V6G Mean±std
ΔE_{elec}	-641.67±38.91	-419.66± 37.36	-592.76±33.31	-527.91± 54.10	-544.54±26.74
ΔE_{vdw}	-132.12±7.26	-102.91±5.88	-98.90±6.73	-106.59±6.44	-117.32± 6.79
ΔE_{gas}	-773.80±35.40	-522.57± 35.91	-691.66±33.23	-634.50± 54.88	-661.86±24.16
ΔG_{PB}	631.05± 33.17	432.44± 32.78	580.15±27.02	530.17±49.25	538.05±22.21
ΔG_{sur}	-22.81± 0.32	-19.16± 0.38	-19.83±0.63	-19.51±0.38	-20.52±0.36
ΔG_{polar}	-10.62 ±10.21	12.78±9.91	-12.61±10. 51	2.26±10.51	-6.50±10.27
$\Delta G_{non\text{-}polar}$	-154.93±7.58	-122.07±6.26	-118±7.36	-126.10±6.82	-137.84±7.15
ΔG_{TOT}	-165.56 ± 7.57	-109.29±7.85	-131.34±9.80	-123.84±9.09	-144.33 ±7.47

Table 3-8 Binding free energy components calculated with MM-PBSA for wildtype and mutants of LYQLEN of insulin (Sh2-St6)

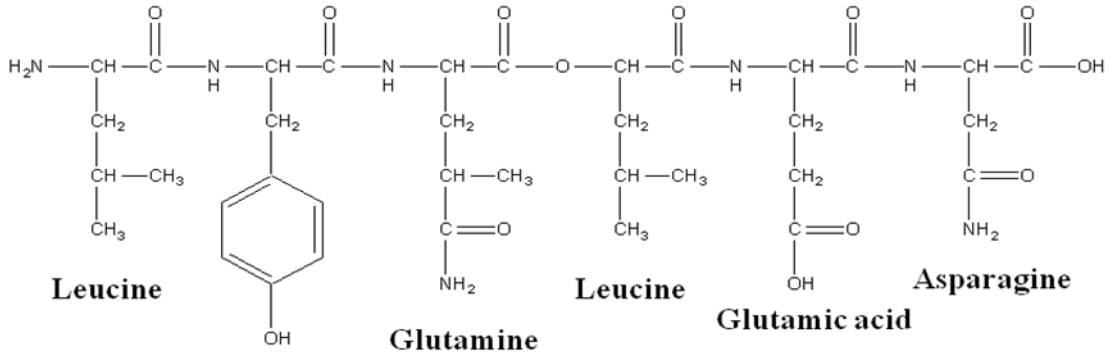
Energy (Kcal/ mol)	Wild type Mean ±std	Mutant – Y2G Mean±std	Mutant – Q3G Mean±std	Mutant –L4 G Mean±std	Mutant – N6G Mean±std
ΔE_{elec}	298.79±36.86	374.35±30.91	359.94± 33.49	327.50±36.68	304.01±36.57
ΔE_{vdw}	-105.15±4.97	-80.86±4.71	-87.77±5.63	-86.63±4.87	-91.68±5.30
ΔE_{gas}	193.65±35.14	293.49±30.22	272.16± 33.19	240.87±36.64	212.32±34.90
ΔG_{PB}	-280.31± 32.94	-362.93± 26.10	-345.09±33.94	-307.99±33.68	-289.87±34.98
ΔG_{sur}	-16.58±0.32	-14.11±0.48	-15.08±0.43	-15.27±0.42	-15.36±0.48
ΔG_{polar}	18.48±8.39	11.41±8.26	14.85±6.86	19.50±8.27	14.14±7.59
$\Delta G_{non\text{-}polar}$	-121.73±5.29	-94.97±5.19	-102.85±6.06	-101.90±5.29	-107.85±5.78
ΔG_{TOT}	-103.24 ±6.30	-83.56±7.25	-88.00±6.07	-82.40 ±7.42	-92.91±6.26



A



B **Methionine**



C

Figure 3-1 Chemical structure of VQIVYK (A), MVGGVV (B) and LYQLEN (C)

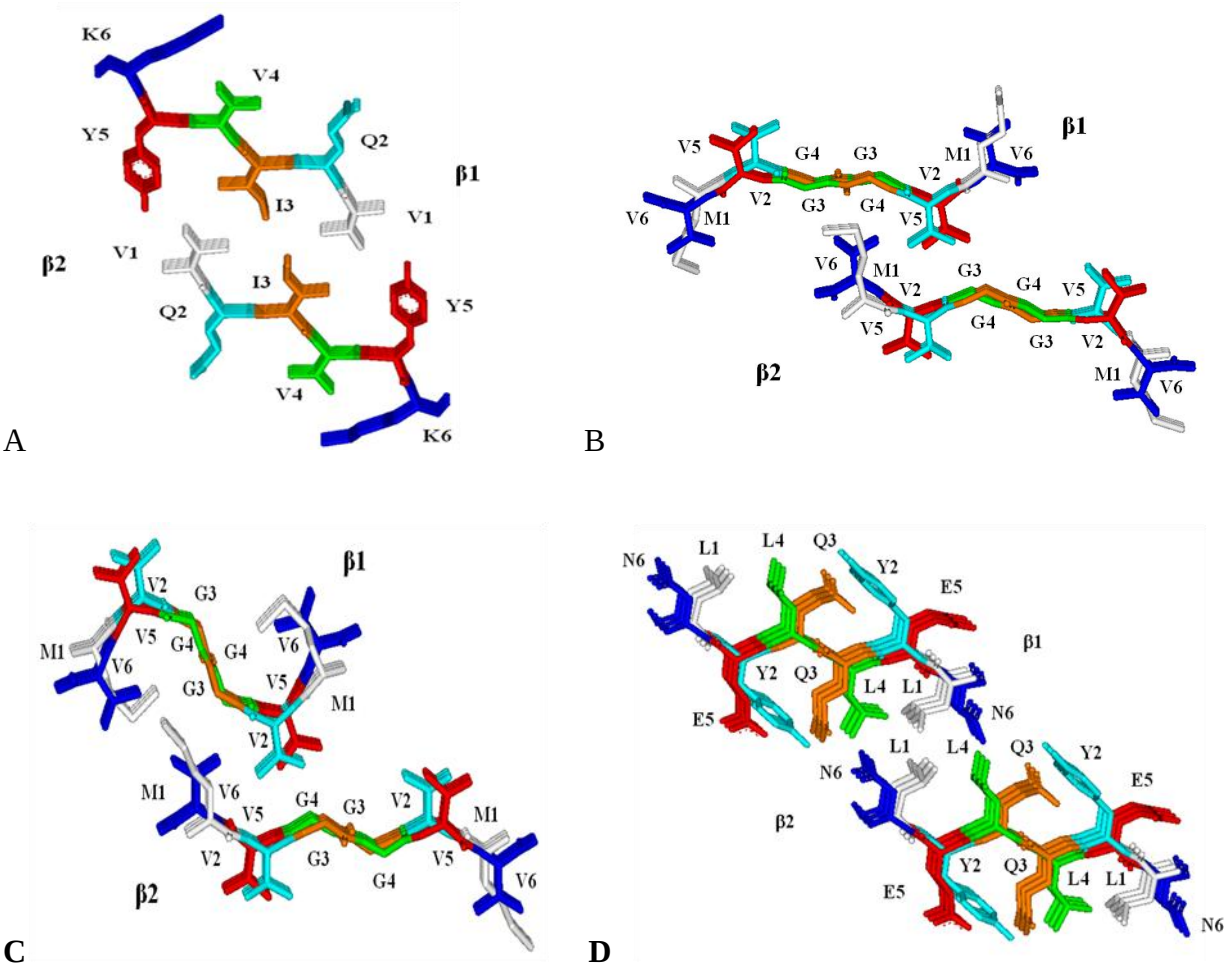
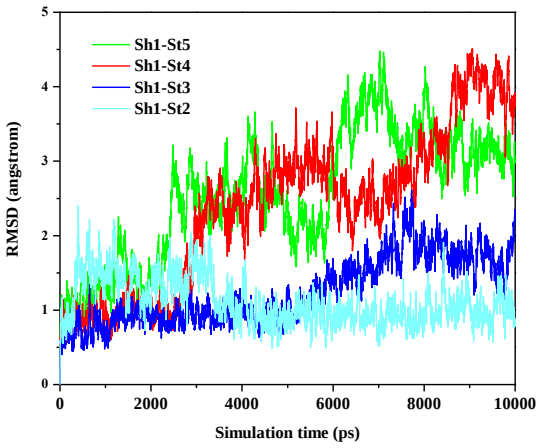
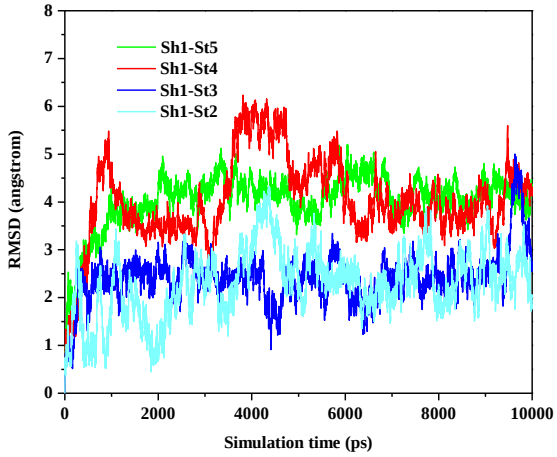


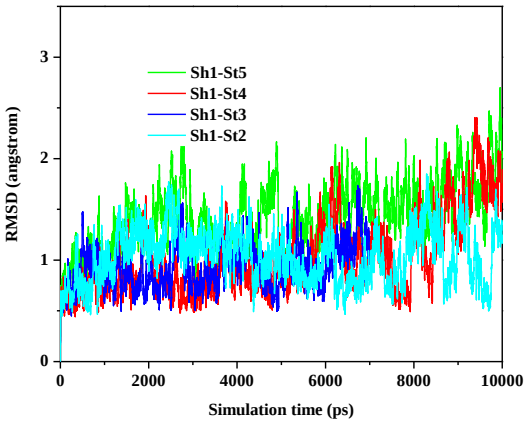
Figure 3-2 The atomic representation of the aggregates of VQIVYK (A), MVGGVV1 B), MVGGVV2 (C) and LYQLEN (D)



A

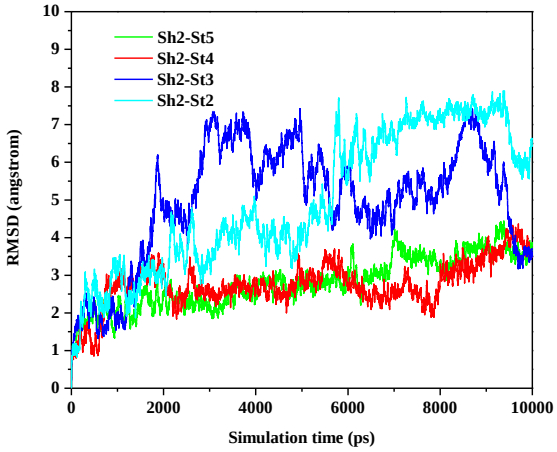


B

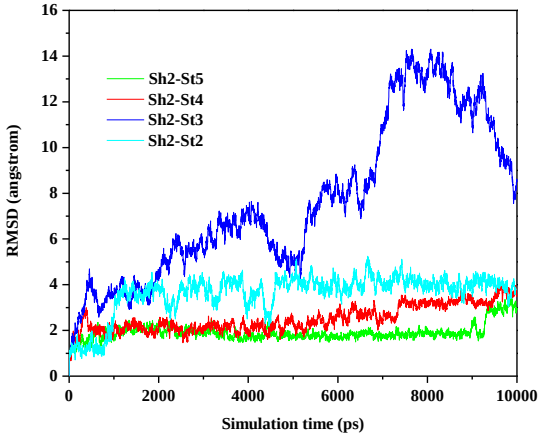


C

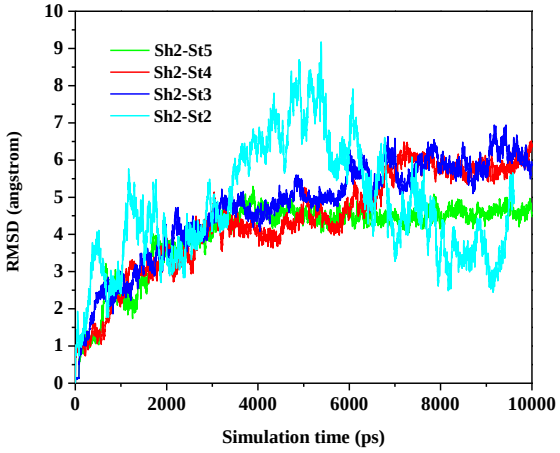
Figure 3-3 Time evolution of the RMSD of the wildtype and mutants of single sheet with different number of strands:(A) tau oligomer (VQIVYK), (B) Abeta amyloid (MVGGVV1) and (C) insulin amyloid (LYQLEN)



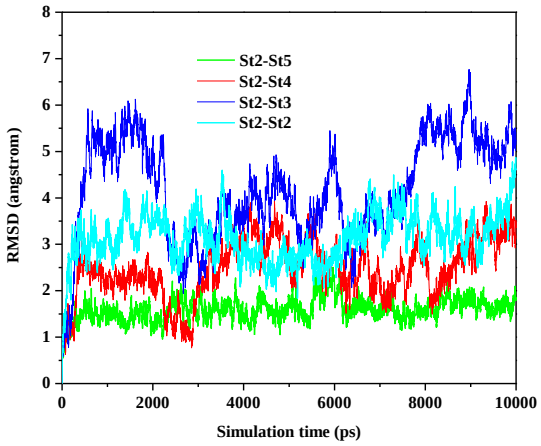
A



B

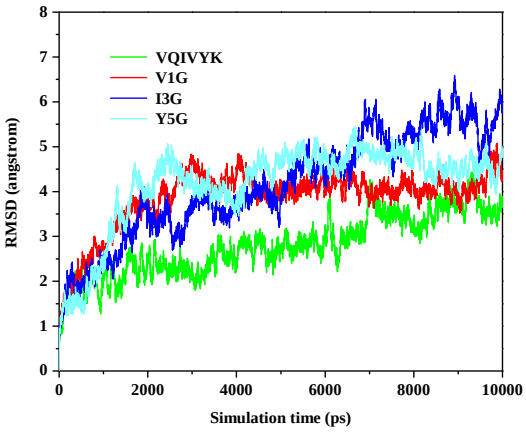


C

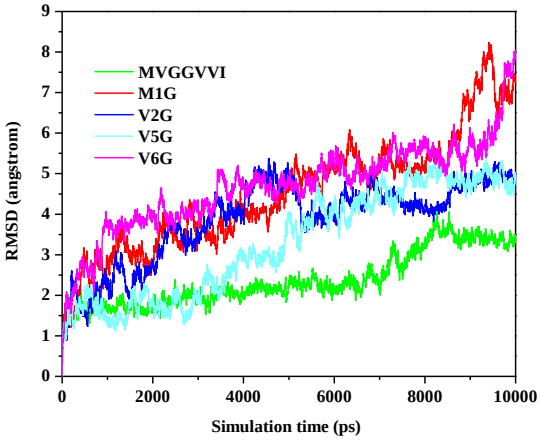


D

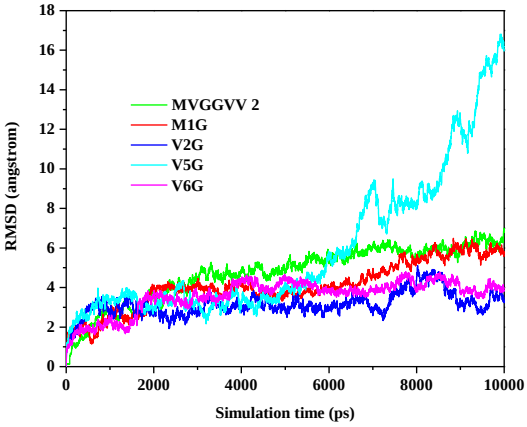
Figure 3-4 Time evolution of the RMSD values of wildtype aggregates of two sheets with different number of strands: tau oligomer VQIVYK (A), Abeta amyloid MVGGVV1 (B) , Abeta amyloid MVGGVV2(C) and insulin amyloid LYQLEN (D)



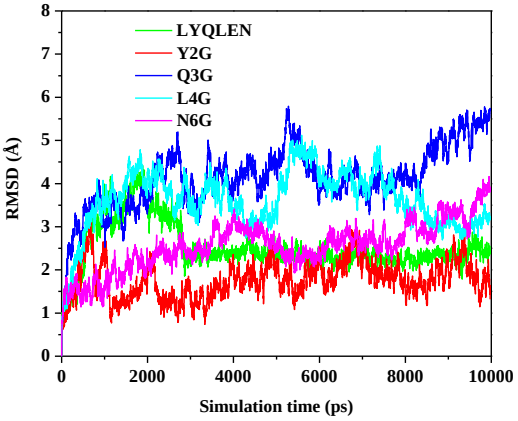
A



B

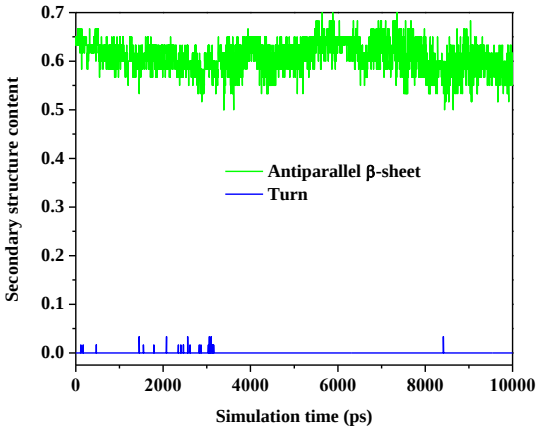


C

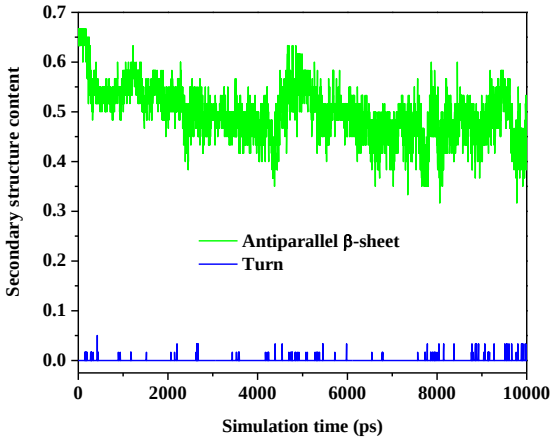


D

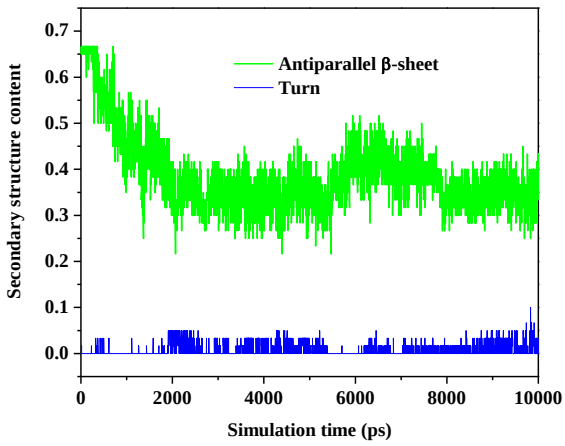
Figure 3-5 Time evolution of the RMSD of wildtype and mutants of (A) VQIVYK oligomer of (Sh2-St6) tau peptide, (B) MVGGVV1 oligomer of (Sh2-St6) Abeta peptide, (C) MVGGVV2 oligomer (Sh2-St6) of Abeta petide and (D) LYQLEN oligomer of (Sh2-St6) insulin



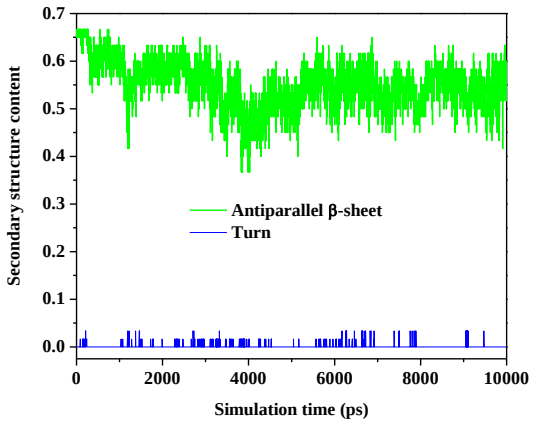
A



B

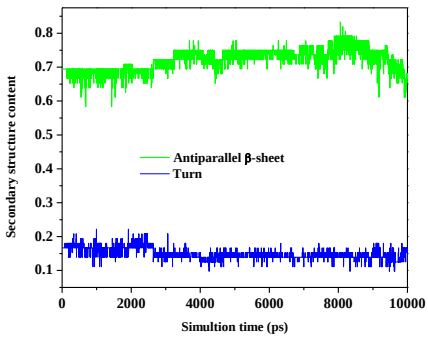


C

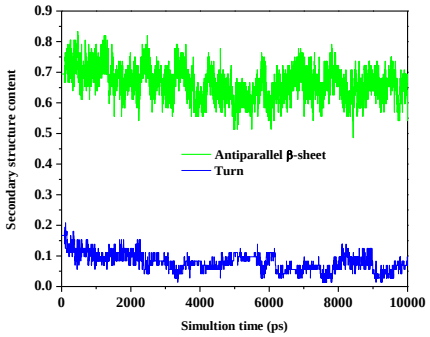


D

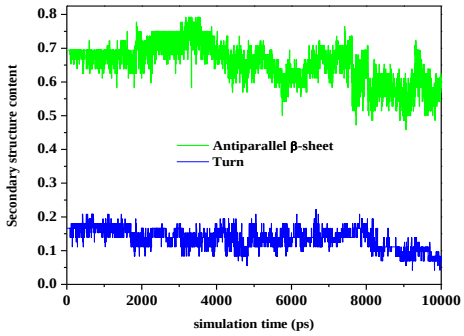
Figure 3-6 Time evolution of the secondary structure contents for Sh2-St5 aggregate of VQIVYK and its mutants (A) Wild type, (B) Y5G, (C) I3G and (D) V1G



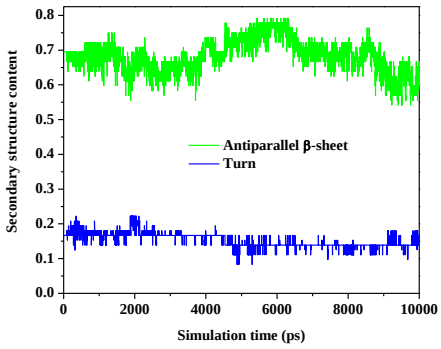
A



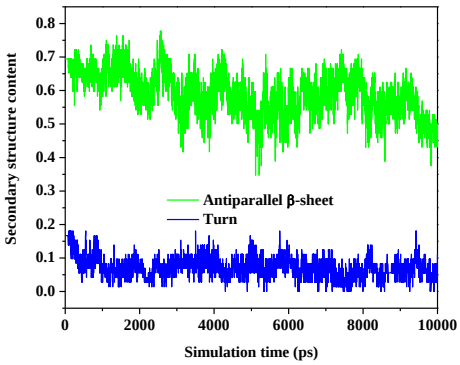
B



C

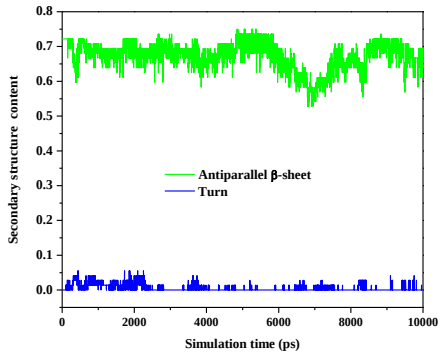


D

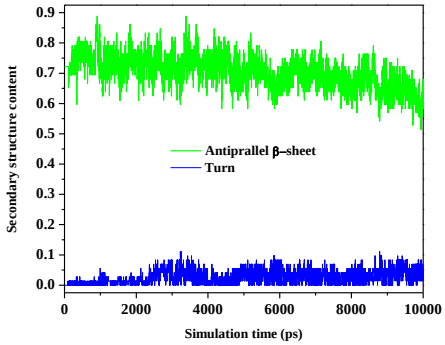


E

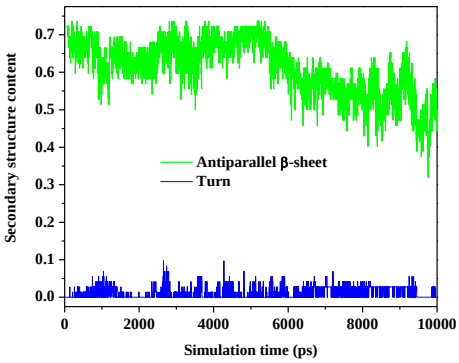
Figure 3-7 Time evolution of the secondary structure contents for Sh2-St6 aggregates of MVGGVV1 Wild type and mutants (A) Wild type, (B) V6G, (C) V5G, (D) V2G and (E) M1G



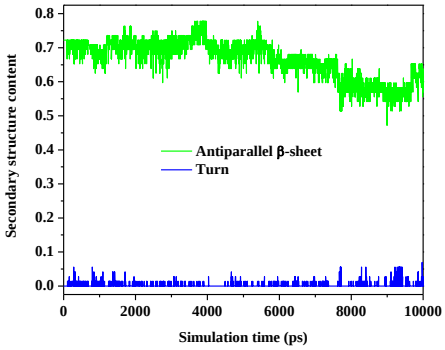
A



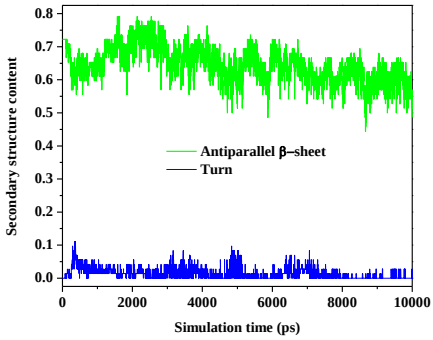
B



C

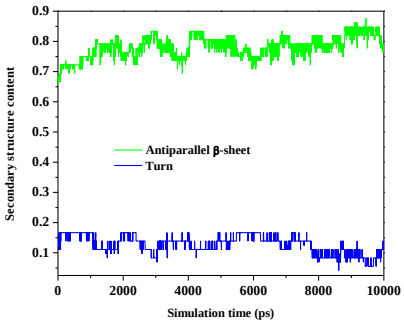


D

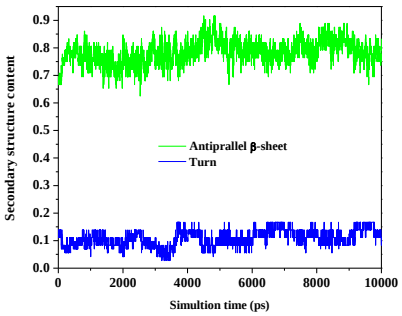


E

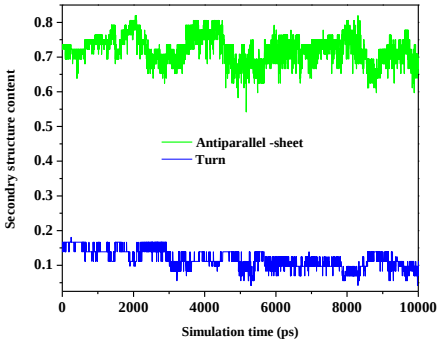
Figure 3-8 Time evolution of the secondary structure contents for Sh2-St6 aggregate of MVGGVV2 wild type and mutants (A) Wild type, (B) V6G, (C) V5G, (D) V2G and (E) M1G



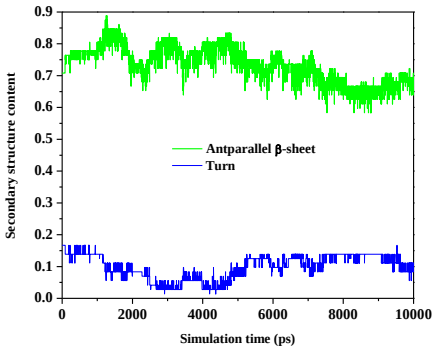
A



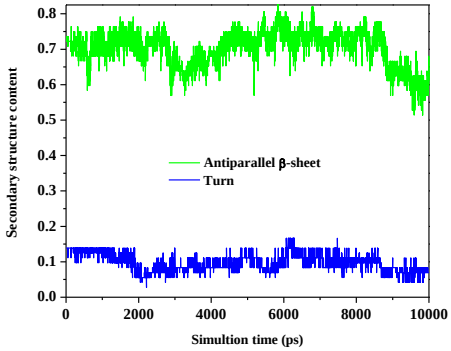
B



C



D



E

Figure 3-9 Time evolution of the secondary structure contents for Sh2-St6 aggregate of LYQLEN wildtype and mutants. (A) Wildtype (B) N6G, (C) L4G, (D) Q3G and (E) Y2G

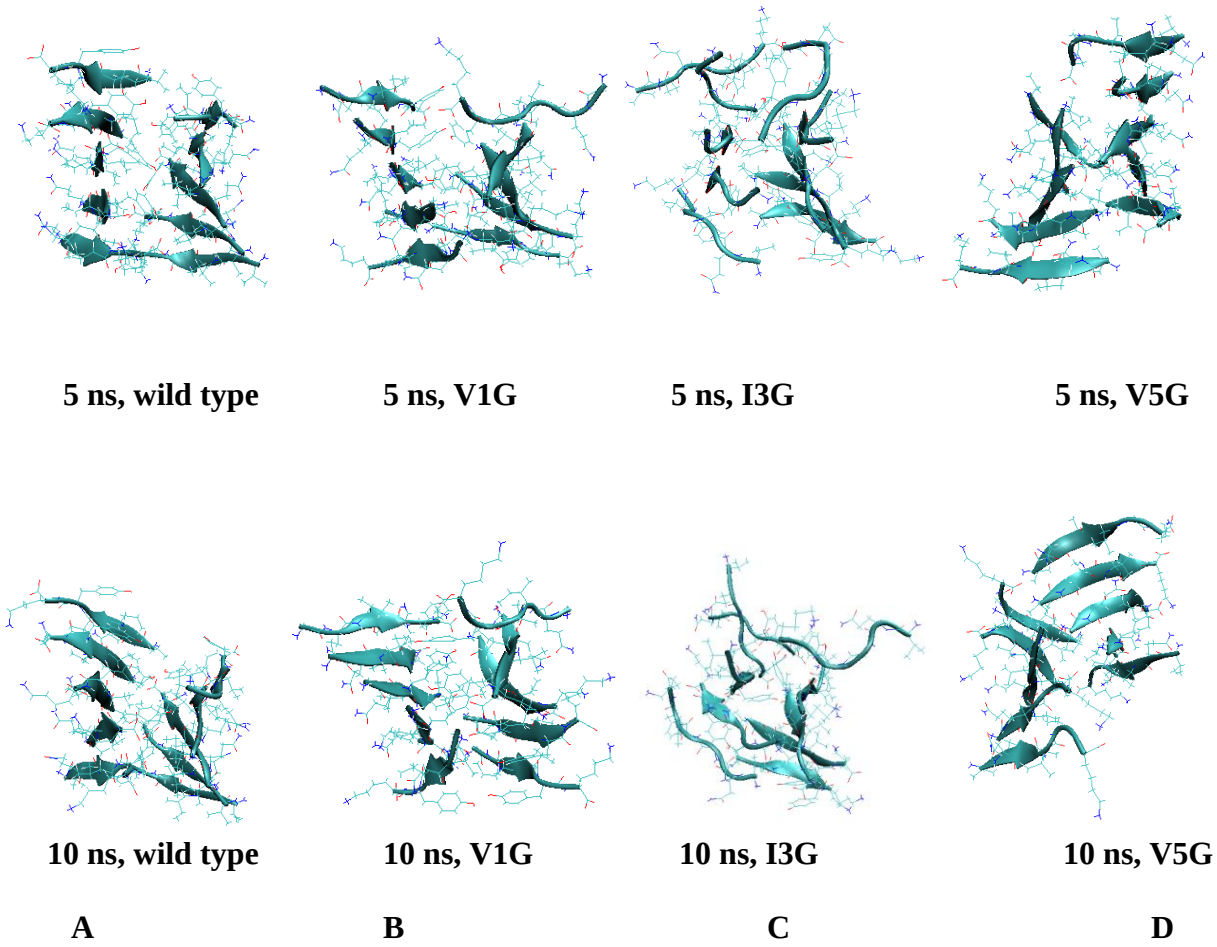
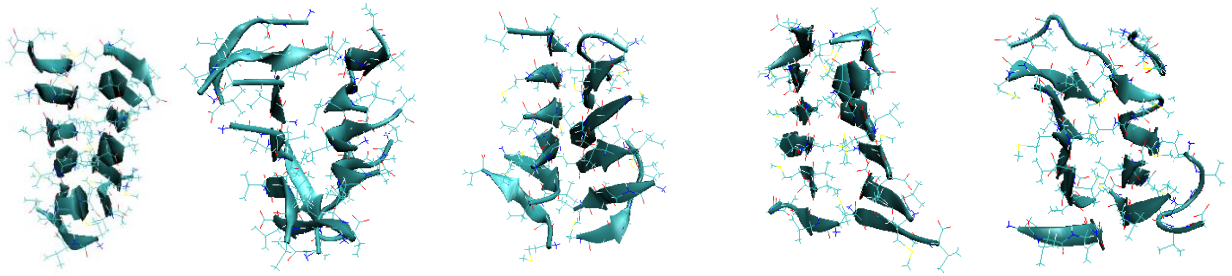
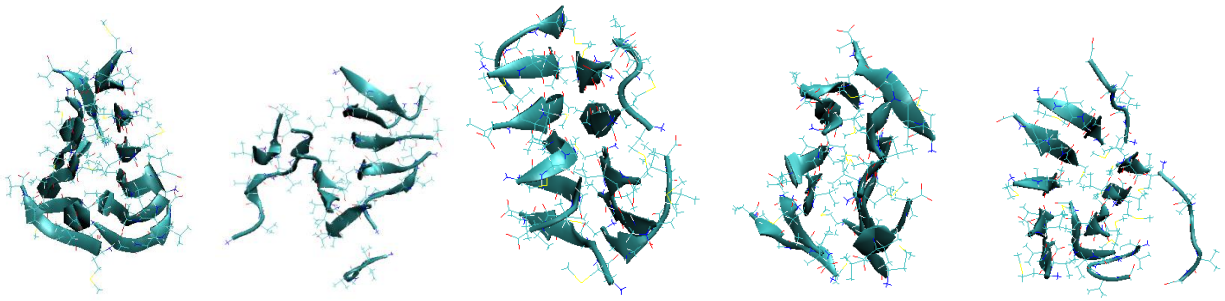


Figure 3-10 Snapshots of Sh2-St5 aggregates of VQIVYK wild type and mutants at 5 ns (top) and 10ns (bottom). (A) Wild type, (B) V1G, (C) I3G and (D) I5G. While the wild type is the most stable, the mutant I3G is the least stable.



5 ns, wild type **5 ns, M1G** **5 ns, V2G** **5 ns, V5G** **5 ns, V6G**



10 ns, wild type **10 ns, M1G** **10 ns, V2G** **10 ns, V5G** **10 ns, V6G**

A **B** **C** **D** **E**

Figure 3-11 Snapshot of Sh2-St6 aggregate of MVGGVV1 and mutants at 5ns (top) and 10ns (bottom) (A) Wild type, (B) M1G, (C) V2G, (D) V5G and (E) V6G. The wild type is the most stable, while the V6G and M1g mutants are the least stable.

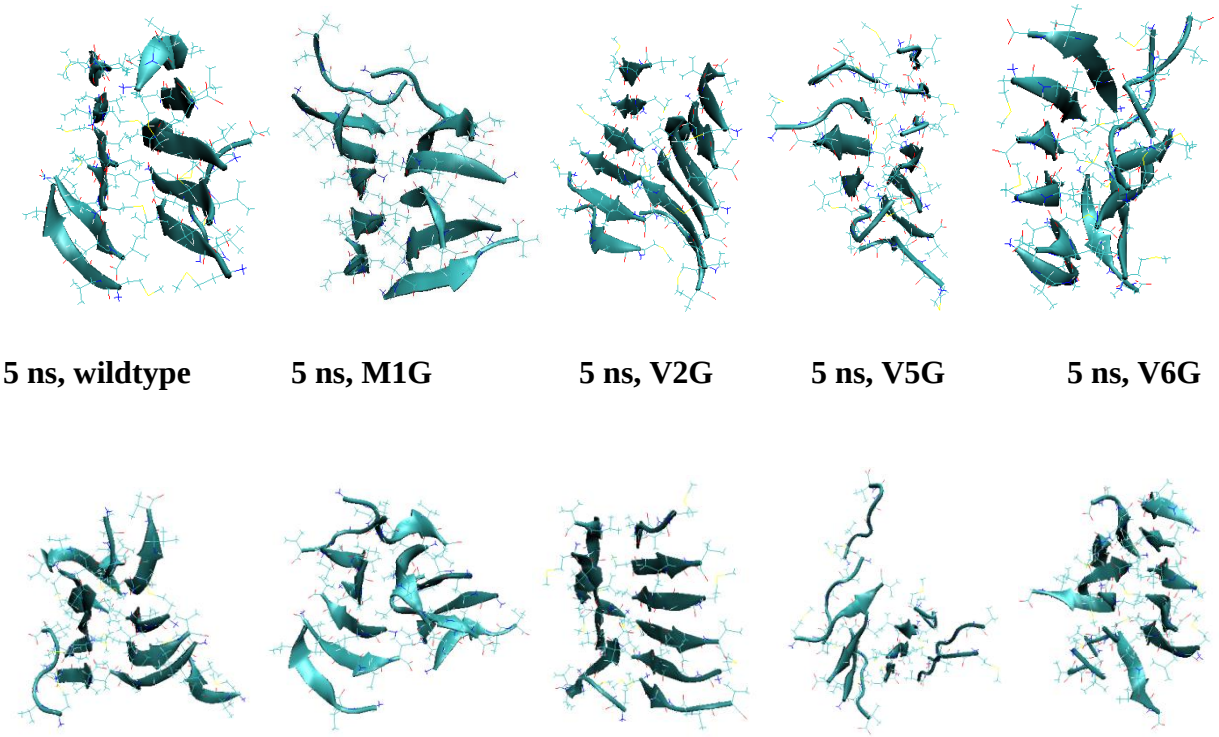


Figure 3-12 Snapshots of Sh2-St6 aggregates of MVGGVV2 and mutants at 5ns (top) and 10ns (bottom) (A) Wild type, (B) M1G, (C) V2G, (D) V5G and (E) V6G. the wild type is the most stable while the mutant V5G is the least stable.

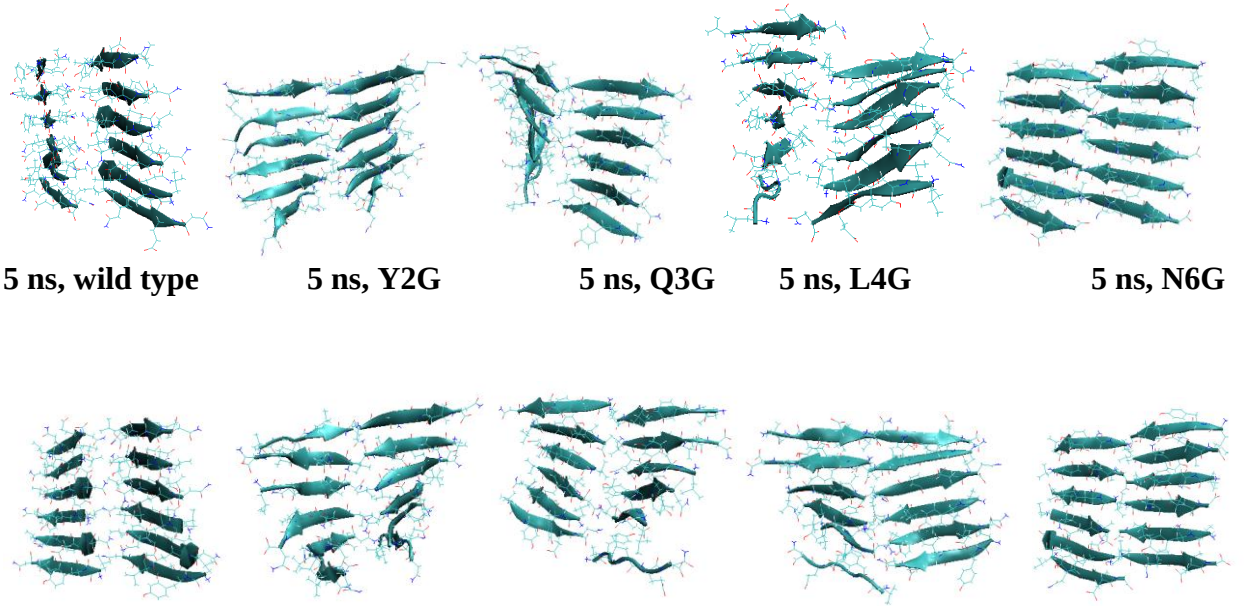


Figure 3-13 Snapshots of Sh2-St6 aggregate of LYQLEN and mutants at 5ns (top) and 10ns (bottom) (A) Wild type, (B) Y2G, (C) Q3G, (D) L4G and (E) N6G.While the wild type is the most stable, the mutants Y2G and Q3G are the least stable.

3.5 References

89. Selkoe, D. J., Folding proteins in fatal ways. *Nature* **2003,** 426, (6968), 900-904.

90. Hamley, I. W., Peptide fibrillization. *Angewandte Chemie-International Edition* **2007,** 46, (43), 8128-8147.

91. Makin, O. S.; Serpell, L. C., Structures for amyloid fibrils. *Febs Journal* **2005,** 272, (23), 5950-5961.

92. Nelson, R.; Eisenberg, D., Recent atomic models of amyloid fibril structure. *Current Opinion in Structural Biology* **2006,** 16, (2), 260-265.

93. Wiltzius, J. J. W.; Sievers, S. A.; Sawaya, M. R.; Cascio, D.; Popov, D.; Riekel, C.; Eisenberg, D., Atomic structure of the cross-beta spine of islet amyloid polypeptide (amylin). *Protein Science* **2008,** 17, (9), 1467-1474.

94. Ivanova, M. I.; Sievers, S. A.; Sawaya, M. R.; Wall, J. S.; Eisenberg, D., Molecular basis for insulin fibril assembly. *Proceedings of the National Academy of Sciences of the United States of America* **2009,** 106, (45), 18990-18995.

95. Park, J.; Kahng, B.; Hwang, W., Thermodynamic Selection of Steric Zipper Patterns in the Amyloid Cross-beta Spine. *Plos Computational Biology* **2009,** 5, (9), 17.

96. Gsponer, J.; Haberthur, U.; Caflisch, A., The role of side-chain interactions in the early steps of aggregation: Molecular dynamics simulations of an amyloid-forming peptide from the yeast prion Sup35. *Proceedings of the National Academy of Sciences of the United States of America* **2003,** 100, (9), 5154-5159.

97. Esposito, L.; Paladino, A.; Pedone, C.; Vitagliano, L., Insights into structure, stability, and toxicity of monomeric and aggregated polyglutamine models from molecular dynamics simulations. *Biophysical Journal* **2008,** 94, (10), 4031-4040.

98. Esposito, L.; Pedone, C.; Vitagliano, L., Molecular dynamics analyses of cross-beta-spine steric zipper models: beta-sheet twisting and aggregation. *Proceedings of the National Academy of Sciences of the United States of America* **2006,** 103, (31), 11533-11538.

99. Wei, G. H.; Song, W.; Derreumaux, P.; Mousseau, N., Self-assembly of amyloid-forming peptides by molecular dynamics simulations. *Frontiers in Bioscience* **2008,** 13, 5681-5692.

100. Wu, C.; Wang, Z. X.; Lei, H. X.; Zhang, W.; Duan, Y., Dual binding modes of Congo red to amyloid protofibril surface observed in molecular dynamics simulations. *Journal of the American Chemical Society* **2007,** 129, (5), 1225-1232.

101. Berryman, J. T.; Radford, S. E.; Harris, S. A., Thermodynamic Description of Polymorphism in Q- and N-Rich Peptide Aggregates Revealed by Atomistic Simulation. *Biophysical Journal* **2009,** 97, (1), 1-11.

102. Chang, L. K.; Zhao, J. H.; Liu, H. L.; Liu, K. T.; Chen, J. T.; Tsai, W. B.; Ho, Y., Molecular Dynamics Simulations to Investigate the Structural Stability and Aggregation Behavior of the GGVVIA Oligomers Derived from Amyloid beta Peptide. *Journal of Biomolecular Structure & Dynamics* **2009,** 26, (6), 731-740.

103. Nerelius, C.; Johansson, J.; Sandegren, A., Amyloid beta-peptide aggregation. What does it result in and how can it be prevented? *Frontiers in Bioscience* **2009,** 14, 1716-U3856.

104. Rauk, A., Why is the amyloid beta peptide of Alzheimer's disease neurotoxic? *Dalton Transactions* **2008**, (10), 1273-1282.

105. Rauk, A., The chemistry of Alzheimer's disease. *Chemical Society Reviews* **2009,** 38, (9), 2698-2715.

106. Teplow, D. B.; Lazo, N. D.; Bitan, G.; Bernstein, S.; Wyttenbach, T.; Bowers, M. T.; Baumketner, A.; Shea, J. E.; Urbanc, B.; Cruz, L.; Borreguero, J.; Stanley, H. E., Elucidating amyloid beta-protein folding and assembly: A multidisciplinary approach. *Accounts of Chemical Research* **2006,** 39, (9), 635-645.

107. Berriman, J.; Serpell, L. C.; Oberg, K. A.; Fink, A. L.; Goedert, M.; Crowther, R. A., Tau filaments from human brain and from in vitro assembly of recombinant protein show cross-beta structure. *Proceedings of the National Academy of Sciences of the United States of America* **2003,** 100, (15), 9034-9038.

108. Devlin, G. L.; Knowles, T. P. J.; Squires, A.; McCammon, M. G.; Gras, S. L.; Nilsson, M. R.; Robinson, C. V.; Dobson, C. M.; MacPhee, C. E., The component polypeptide chains of bovine insulin nucleate or inhibit aggregation of the parent protein in a conformation-dependent manner. *Journal of Molecular Biology* **2006,** 360, (2), 497-509.

109. Hong, D. P.; Fink, A. L., Independent heterologous fibrillation of insulin and its B-chain peptide. *Biochemistry* **2005,** 44, (50), 16701-16709.

110. Wilhelm, K. R.; Yanamandra, K.; Gruden, M. A.; Zamotin, V.; Malisauskas, M.; Casaite, V.; Darinskas, A.; Forsgren, L.; Morozova-Roche, L. A., Immune reactivity towards insulin, its

amyloid and protein S100B in blood sera of Parkinson's disease patients. *European Journal of Neurology* **2007,** 14, (3), 327-334.

111. Brange, J.; Andersen, L.; Laursen, E. D.; Meyn, G.; Rasmussen, E., Toward understanding insulin fibrillation. *Journal of Pharmaceutical Sciences* **1997,** 86, (5), 517-525.

112. Ahmad, A.; Uversky, V. N.; Hong, D.; Fink, A. L., Early events in the fibrillation of monomeric insulin. *Journal of Biological Chemistry* **2005,** 280, (52), 42669-42675.

113. Zhang, Z. Q.; Chen, H.; Bai, H. J.; Lai, L. H., Molecular dynamics Simulations on the oligomer-formation process of the GNNQQNY peptide from yeast prion protein sup35. *Biophysical Journal* **2007,** 93, (5), 1484-1492.

114. Xu, W. X.; Ping, J.; Li, W. F.; Mu, Y. G., Assembly dynamics of two-beta sheets revealed by molecular dynamics simulations. *Journal of Chemical Physics* **2009,** 130, (16), 8.

115. Kayed, R.; Pensalfini, A.; Margol, L.; Sokolov, Y.; Sarsoza, F.; Head, E.; Hall, J.; Glabe, C., Annular Protofibrils Are a Structurally and Functionally Distinct Type of Amyloid Oligomer. *Journal of Biological Chemistry* **2009,** 284, (7), 4230-4237.

116. Vitagliano, L.; Stanzione, F.; De Simone, A.; Esposito, L., Dynamics and Stability of Amyloid-Like Steric Zipper Assemblies with Hydrophobic Dry Interfaces. *Biopolymers* **2009,** 91, (12), 1161-1171.

117. http://www.doe-mbi.ucla.edu/~sawaya/chime/xtalpept (September, 2009),

118. Case DA, D. T., Cheatham TE, Simmerling CL, Wang J, Duke RE, Luo R, Walker RC, Zhang W, Merz KM, Roberts B, Wang B, Hayik S, Roitberg A, Seabra G, Kolossváry I, Wong KF, Paesani F, Vanicek J, Liu J, Wu X, Brozell SR, Steinbrecher T, Gohlke H, Cai Q, Ye X,

Wang J, Hsieh MJ, Cui G, Roe DR, Mathews DH, Seetin MG, Sagui C, Babin V, Luchko T, Gusarov S, Kovalenko A, Kollman PA *Amber 10*, University of California, : San Francisco,, 2008.

119. Chong, L. T.; Duan, Y.; Wang, L.; Massova, I.; Kollman, P. A., Molecular dynamics and free-energy calculations applied to affinity maturation in antibody 48G7. *Proceedings of the National Academy of Sciences of the United States of America* **1999,** 96, (25), 14330-14335.

120. Hardy, J.; Selkoe, D. J., Medicine - The amyloid hypothesis of Alzheimer's disease: Progress and problems on the road to therapeutics. *Science* **2002,** 297, (5580), 353-356.

121. Lesne, S.; Koh, M. T.; Kotilinek, L.; Kayed, R.; Glabe, C. G.; Yang, A.; Gallagher, M.; Ashe, K. H., A specific amyloid-beta protein assembly in the brain impairs memory. *Nature* **2006,** 440, (7082), 352-357.

122. Sciarretta, K. L.; Gordon, D. J.; Meredith, S. C., Peptide-based inhibitors of amyloid assembly. In *Amyloid, Prions, and Other Protein Aggregates, Part C*, Elsevier Academic Press Inc: San Diego, 2006; Vol. 413, pp 273-312.

123. Bogan, A. A.; Thorn, K. S., Anatomy of hot spots in protein interfaces. *Journal of Molecular Biology* **1998,** 280, (1), 1-9.

124. Haydar, S. N.; Yun, H. E. D.; Staal, R. G. W.; Hirst, W. D., Small-Molecule Protein-Protein Interaction Inhibitors as Therapeutic Agents for Neurodegenerative Diseases: Recent Progress and Future Directions. In *Annual Reports in Medicinal Chemistry, Vol 44*, Elsevier Academic Press Inc: San Diego, 2009; Vol. 44, pp 51-69.

125. Blazer, L. L.; Neubig, R. R., Small Molecule Protein-Protein Interaction Inhibitors as CNS Therapeutic Agents: Current Progress and Future Hurdles. *Neuropsychopharmacology* **2009,** 34, (1), 126-141.

126. Otto, M.; Lewczuk, P.; Wiltfang, J., Neurochemical approaches of cerebrospinal fluid diagnostics in neurodegenerative diseases. *Methods* **2008,** 44, (4), 289-298.

127. Tenidis, K.; Waldner, M.; Bernhagen, J.; Fischle, W.; Bergmann, M.; Weber, M.; Merkle, M. L.; Voelter, W.; Brunner, H.; Kapurniotu, A., Identification of a penta- and hexapeptide of islet amyloid polypeptide (IAPP) with amyloidogenic and cytotoxic properties. *Journal of Molecular Biology* **2000,** 295, (4), 1055-1071.

128. Porat, Y.; Mazor, Y.; Efrat, S.; Gazit, E., Inhibition of islet amyloid polypeptide fibril formation: A potential role for heteroaromatic interactions. *Biochemistry* **2004,** 43, (45), 14454-14462.

129. Sato, T.; Kienlen-Campard, P.; Ahmed, M.; Liu, W.; Li, H. L.; Elliott, J. I.; Aimoto, S.; Constantinescu, S. N.; Octave, J. N.; Smith, S. O., Inhibitors of amyloid toxicity based on beta-sheet packing of A beta 40 and A beta 42. *Biochemistry* **2006,** 45, (17), 5503-5516.

130. Takahashi, T.; Ohta, K.; Mihara, H., Rational design of amyloid beta peptide-binding proteins: Pseudo-A beta beta-sheet surface presented in green fluorescent protein binds tightly and preferentially to structured A beta. *Proteins-Structure Function and Bioinformatics* **2010,** 78, (2), 336-347.

131. Kim, Y. S.; Lee, J. H.; Ryu, J.; Kim, D. J., Multivalent & Multifunctional Ligands to beta-Amyloid. *Current Pharmaceutical Design* **2009,** 15, (6), 637-658.

CHAPTER 4 ALTERNATIVE PACKING MODES AS BASIS FOR AMYLOID POLYMORPHISM IN FIVE FRAGMENTS

4.1 Background

X-ray diffraction of amyloid fibrils indicated the structure is highly conserved along the fibril axis, with variation in the plane of the fibril cross-section.[132] There are three major structural features that may decide the overall amyloid fibril morphologies: (1) backbone orientation; (2) backbone conformation; and (3) the way in which the oligomers associate. The combination of these three factors can give rise to an enormous variation in conformational detail and fibril morphology.[133]

Eisenberg laboratory was able to grow three-dimensional microcrystals [16,22,94,93,30,134] and determine atomic resolution structures of about 50 short fibril-forming peptide segments of hexapeptide abd hepatpetide fragments of several amyloid proteins (inclusing insulin, Aβ, tau, prion and amylin). These studies have identified examples for most out of eight possible packing classes (parallel and antiparallel β-sheetsstacked face to face or face to back in parallel or antiparallel manner). These common motifs called steric zippers, in which interdigitated side chain hold together pair of β-sheets.[22,18] Further, some peptides are capable of forming different types of steric zipper, that offers a possible explanation for amyloid polymorphism.[22] Polymorphism of amyloid fibrils by a range of proteins including Aβ, prion, glucagon and amylin has been observed and is influenced by the environment in which the fibrillogenes occurs. The polymorphic nature of Aβ peptide fibrils has been suggested to alter their pathogenic action. Polymorphism of amyloid lead to difference in terms protofilament backbone regions,

secondary structure, chromophore alignment along the fibril axis, and fibril superstructure.[135] While the selection of the filament structure depends on the growth condition, which can be purely mechanical agitation, once a stable filament is formed, it continues to grow, keeping the atomistic order even if the growth condition changes.[136] A fundamental question remains regarding how these structures are formed. The importance of oligomer polymorphism is increasingly recognized, explaining several observations, from propagation of prion strain infectivity and other protein polymorphism [136, 137, 138, 139] to the variable cytotoxicity of amyloids differently grown from the same peptides and proteins [139, 140, 141, 142], the appearance of in-path and off-path intermediates of fibril growth [143, 142] and the structural heterogeneity of amyloid fibrils and their precursors grown from the same peptide/protein under different environmental conditions.[144, 145]

Hydrogen-bonding interactions are inherent to β-sheet stability.[146] The stacking of multiple β strands in amyloid fibrils has been, in part, ascribed to cooperative hydrogen bonding.[147] Molecular modeling has shown that polar residues β-sheets are stabilized by hydrogen bonds between polar side chains, such as those between glutamine and asparagine. The glutamine- and asparagine-rich regions are commonly found in the N-termini of both mammalian and yeast prion proteins. Glutamine/asparagine (Q/N) rich domains show a greater tendency to form self propagating amyloid fibrils. Gln and Asn rich sequence are prone to assemble into different amyloid structures, since they are able to form diverse array of molecular interactions. They can be donor or acceptor of both main-chain/side-chain and side-chain/main-chain hydrogen bonds.[148] Our previous study comparing the aggregation behaviors of short segments

of amyloid peptides with small hydrophobic residues (VQIVYK, MVGGVV) and Q/N rich residues (LYQLEN) indicated there is a greater stability in LYQLEN segment from insulin.[78] The extra structural stability of the Q/N residues LYQLEN aggregates may be related to the number of hydrogen bonds formed between the backbone and the side chains.[149]

The significant role of steric zipper in the structural stability of short amyloidogenic peptides have been demonstrated by molecular dynamic studies. [88, 98,150] Park *et al*. [95] address the structural selection mechanism of different double layer peptides including GNNQQNY, NNQQ, VEALYL, KLVFFAE and STVIIE, and found that the patterns with the lowest binding free energy correspond to X-Ray structures. Wu *et al* [151] using MD simulation in combination with ssNMR data proposed two polymorphs of $A\beta_{9-40}$ peptide fibril. Papacone *et al* [152] presented a systemic study between two polymorphic forms of $A\beta_{9-40}$ suggesting double layer morphology is more stable than the three fold morphology. Computational investigation on the structure, energy, and solvent interaction of four classes of Aβ dodecamers by Ma *et al* [153] indicate that β-sheets packed orthogonally could be the most stable species for Aβ dodecamers. Berryman *et al* [154, 101] investigated the thermodynamic stability of various possible polymorphic models of short segments of amyloid peptides.

Here we report on multiple all-atomic MD simulations with explicit water at 300K,conducted on five polymorphic pairs short amyloidogenic peptides oligomers aggregates in a crystal-free context. We performed the simulation starting with experimentally determined microcrystal structure. Our study focuses on investigating the stability of various polymorphic pairs. This study aims to answer the following questions:

1) What is the effect of side chain (polar versus nonpolar) on the stability of pair of steric zipper polymorphs forms of small amyloid segments?

2) Among the studied pair of polymorphs which one of them are more stable in crystal free context under physiological conditions?

3) Is there a relationship between hydrogen bond content and structural stability?

4) What is the driving force for the association of the polymorph aggregates?

5) How does a single point mutation of the N/Q side chains in the steric zipper of NNQNTF and GNNQQNY influence the stability of the aggregates?

4.2 Methods

The polymorphic pairs of five small peptides (VQIVYK, SSTNVG, MVGGVV, GNNQQNY and NNQNTF) were investigated. The aggregates were derived from the crystal structure with different packing polymorphs (Figure 4-1). The microcrystal structure and coordinate of the five peptides polymorphic forms I and II (VQIVYK, SSTNVG, MVGGVV, GNNQQNY and NNQNTF) assembled with two layers of β sheets was determined by Eisenberg group.[22, 93,16] The detailed summery of the simulation conditions of each the peptides and single point mutation of N/Q residues with glycine in the steric zipper for are GNNQQNY and NNQNTF are shown in Table 4-1. The microcrystal structure chosen in this study represent three major groups: polar steric zipper rich with Q/N residues (GNNQQNY and NNQNTF), polar steric zipper not rich in Q/N residues (SSTNVG and VQIVYK) and nonpolar steric zipper

(MVGGVV). All of the studied structures were hexapeptide (except GNNQQNY, which consists of seven residues).

The hexapeptide SSTNVG is a segment of amylin (residue 28-31) that forms fibrillar amyloid deposits among the pancreatic β-islet cells of type II diabetes. Microcrystals of SSTNVG grown from different solutions revealed two polymorphs. The SSTNVG form I features a pair of serine residues at the center. The structure of SSTNVG form II packing has the center of the interface Asn31, rather than Ser29 (Figure 4-1A).[16] The hexapeptide VQIVYK, residue 306-311 of the tau, forms intracellular amyloid fibrils in Alzheimer's disease.[22] The structural organization of VQIVYK is a parallel β-strand within the same β-sheet layer while maintaining anti-parallel organization between the adjacent β-sheet layers.[22] The VQIVYK from tau protein shows polymorphism. The VQIVYK form I is characterized by an apolar dry interface made essentially by the side chains of V1 and I3 of the two sheets and polar interface involving Tyr and Gln side chains (Figure 4-1B). The VQIVYK form II is characterized by an apolar dry interface made essentially by the side chains of V1 and I3 of the two sheets (Figure 4-1B).[16] VQIVYK form I can be transformed to the VQIVYK form II by moving the top sheet to the right. The hexameric peptide segment 170-175 from elk prion with amino acid sequence NNQNTF forms two facial polymorphs. Both of them are found in the same crystal structure. One of them is with a face to face packing, with N1, Q3 and T5 of both sheets forming the interdigitated interface while the other one is a back to back, with side chains N2, N4 and F6 interdigitated (Figure 4-1C).[16] The microcrystals formed from Aβ (35-40) peptide under different incubation conditions gives two different crystal structures showing different conformations and

arrangements of the peptides MVGGVV from Aβ protein. The MVGGVV peptide consists of anti-parallel β-strands within the same β-sheet layers. The shape complementarity involves the side chains of Met1, Val2, Val5 and Val6 β-sheet. The MVGGVV form II polymorph is different from the MVGGVV form I in that that there is 90^0 bending in the upper sheet of MVGGVV form II (Figure 4-1D). The heptameric peptide with amino acid sequence GNNQQNY from the yeast prion Sup35 forms two crystalline forms.[30,22] The β-strands are stacked parallel and are oriented upwards within each β-sheet, and the β-sheets are face-to-face i.e. the same residues on each β-strand interdigitate. Both of them share many similarities. The polar side chains of the dry interface (Asn 2, Gln 4, and Asn 6) are tightly interdigitated with the same three side chains of the mating sheet (Figure 4-1E). These opposing side chains do not form hydrogen bonds with each other; rather, their shapes complement each other closely, forming van der Waals interactions. They also exhibit some differences, particularly around the tyrosine residue, which appears to play a stabilizing role across sheets in the monoclinic form but not in the orthorhombic form.[30, 22]

Computational mutation of the Q/N residues involved in the dry steric-zipper between the sheets for NNQNTF and GNNQQNY was preformed to examine the effect of a single point glycine mutation All the starting structures of the mutants were built from the wild type [93] by replacing the side chains of the targeted residues with glycine, but without changing the backbone conformations of side-chain orientations using VMD.[155] The simulation details for each model are summarized in Tables 4-1.

The trajectories obtained on the production stages were analyzed to examine the structural change of the oligomers aggregates. The root mean square deviation (RMSD), root mean square fluctuation, inter-sheet distance and inter-strand distance were calculated. To gain an insight into the driving force for association of the preformed oligomer we used a molecular mechanics/generalized Born-surface area method and normal mode analysis (NMA) to calculate various energy terms in binding free energy (ΔG) .[156] We also examined the stability of the oligomers by following the changes in the number of hydrogen bonds and their occupancy.

4.3 Result

4.3.1 Structural stability of wild type steric zipper packing polymorphs: RMSD and RMSF analysis.

The conformational stabilities of the oligomers were monitored by the time evolution of the backbone root mean square (RMSD) and root mean square fluctuation (RMSF) relative to their initial energy minimized structure as shown in Table 4-2 and Figures 4-2. The RMSDs provide useful information on relative stability of the oligomers, and were previously used in stability analyses of amyloid oligomers with β-sheet structure.[21, 98, 157, 158]

The RMSD values for the studied oligomers were calculated for backbone heavy atoms against the corresponding energy minimized structures. To examine the effects of the polar versus nonpolar side chain steric zipper on the stability of amyloid oligomers the mean RMSD (<RMSD>) averaged over the five trajectories was determined. The <RMSD> for both nonpolar and polar dry interface of the five polymorphs are shown in Table 4-2 and Figure 2. The

<RMSD> of the polar Q/N oligomers (GNNQQNY and NNQNTF) were less than 2.0 Å suggesting the Q/N rich peptides maintained the initial structure (see Figure 4-3). The hepta-peptide GNNQQNY with one additional amino acids was found to be relatively more stable with an <RMSD> less than 1 Å. The average RMSD for the other non Q/N rich systems was in the range of 2 to 5.0 Å within the 20 ns simulations. The larger <RMSD> indicates the oligomers that are not rich in Q/N residues are more flexible than those with Q/N side chain at the steric zipper. The oligomers with largest <RMSD>, those with an RMSD ≥ 4 Å, were the two polymorphic forms of MVGGVV and polymorph form II of VQIVYK. A closer look at the structure of these three models indicates that they are having the smallest sheet-to-sheet interface, consequently, their final solution conformations moves further away from the conformation of the initial model (Figure 4-3). The qualitative structural stability comparison of the oligomers based on RMSD of the may be used to predict which of the one of the polymorphs pair might exist outside of the crystal environment.[159] The RMSD during the 20 ns simulation for the Q/N rich oligomer (NNQNTF and GNNQQNY) indicates both forms could exist with similar probability under physiologic condition. Among the other three pairs it might be possible to find the both polymorph form with similar probability for SSNTVG and MVGGVV. The polymorph form I of VQIVYK based on the <RMSD> is most likely to be a dominate structure under physiologic condition. A recent structure based design of amyloid aggregation for tau peptide inhibitor used the VQIVYK polymorph form I as a template[160] indicating the significant of the determining the difference in stability of the packing polymorph of the steric zipper segments of amyloid peptides.

The residue-based root mean square fluctuations (RMSF) of the backbones were used to assess the local dynamics and flexibility of each residue for the five polymorphic forms using ptraj tool in AMBER11. Figure 4-2 shows the RMSF profiles of the different oligomers of the wildtype of various studied amyloid oligomers. The RMSF values for all the five pairs of the amyloid peptide models from our simulations indicates both N- and C-terminal regions residues have a larger RMSF as they are exposed to the solvent water molecules with greater mobility. The models GNNQQNY and NNQNTF have the smallest RMSF compared to other models (see Table 4-2 and Figure 4-2). The RMSF values for the model VQIVYK-II, MVGGVV-I and MVGGVV-II are generally larger than other models. The RMSF analysis is consistence with the RMSD in that the structural stable Q/N rich oligomers have both smaller RMSF and RMSD while the structurally unstable models (VQIVYK-II, MVGGVV-I and MVGGVV-II) have a larger RMSD and RMSF.

4.3.2 Inter-strand (d_{strand}) and inter-sheet (d_{sheet}) distances

To examine the structural stability of the wildtype of the five polymorphic forms oligomers we analyzed the inter-strand and inter-sheet. The d_{strand} is calculated by averaging the mass center distance between each residue in one strand and the respective in-register residue in the adjacent strand, whereas d_{sheet} is calculated by averaging the mass center distance between sheets.[75] The inter-sheet and inter-strand distances for wild type are shown in Table 4-2 and Figures 4-4. The inter-strand and inter-sheet distance from the simulations for the Q/N rich

models GNNQQNY and NNQNTF were found to be almost the same as the initial structures (see Table 4-2 and Figure 4-4). These results indicate the GNNQQNY and NNQNTF models have a greater structural stability (see Figure 4-3). The polar side chains of Asn and Gln with larger sheet-to-sheet interface keeps together the two neighboring sheets. Significant changes in inter-strand and inter-sheet distances were observed for the models with smaller sheet-to-sheet interface for the nonpolar small size side chain at the steric zipper (VQIVYK and MVGGVV).

4.3.3 Sheet-to-sheet association energy

To further quantify the driving force underlying the β-sheet association of the studied wildtype amyloid polymorph models, we calculated the interaction energy between β-sheets for the three pairs (SSNTVG, NNQNTF and GNNQQNY) using the MM-PBSA module [80] in the AMBER package. The VQIVYK and MVGGVV pairs showed a larger RMSD values with significant conformational changes making it difficult to apply the MM-PBSA single trajectory approach for calculating sheet to sheet association energy.[80] The entropy calculation is and thus the -TΔS was averaged over 100 frames of the MD trajectory (1 frame taken at an interval of 50 frames from the total of 5000 frames).

The non covalent association of the β-sheets of amyloid aggregates takes place spontaneous only if it is associated with a negative binding free energy.[88] Detailed characterization of individual energy terms of the calculated binding free energy are shown in Table 4-3. The individual energy terms may be of similar or opposite in sign. An inspection of the free energy components for the wild types investigated in this study reveals that the electrostatic component of the solvation free energy ΔG_{PB} is destabilizing (positive), while the

nonpolar component G_{SA} is stabilizing (negative). This is expected, since the complex formation desolvates the monomers, and reduces solvent-accessible surface area. Entropy component was found to contribute unfavorably to binding, since complexation reduces freedom of motion for the monomers. The electrostatic interaction between sheets is stabilizing. These observations are consistent with previous calculations of the components of the free energy of solvation.[88] However, the less favorable electrostatics in each case is compensated by highly favorable nonpolar component of the free energy. In each case, favorable nature of the nonpolar interaction mostly originates from the van der Waals interaction energy ΔG_{vdW}, as opposed to the nonpolar component of solvation ΔG_{SA}. Despite the fact that the β-sheets association for the oligomers (see Table 4-3) shows a more favorable binding enthalpy, there is an entropy penalty (-5.27 to -9.91). The order water molecules around the amino acid residues in the peptide upon the β-sheets association are released and to an increase in entropy.

In order to identify the residues that contribute the most to the calculated overall binding energy, we used a residue-by-residue decomposition protocol. Binding free energy decomposition at the atomic level allows evaluating the contribution of each residue to the total binding free energy, as well as the contributions of its side-chain and backbone. The MMPBSA.py script in AMBER11 implements per-residue decomposition with both PB and GB implicit solvent models.[54] The PB non-polar solvation component is currently not decomposable. However, the non-polar solvation remains constant about -1.50Kcal/mol (see Table 4-3) and is much smaller than the other energy terms. Thus, we used the MM-PBSA decomposition and the results are show in Figure 4-5 to 4-7. The residues making the most favorable contributions to

the binding free energy between the two sheets are the residues situated at the interface between the two sheets and form stable hydrogen bonds between their backbone atoms and van der Waals interactions between their side-chains (Figure 4-5 to 4-7). The contribution of the N/Q side-chains to the association of the 5 stranded double layer oligomers is larger than the other nonpolar and small size amino acids at the interface, underlining the importance of Q/N amino acid in stabilizing the short segment amyloid peptides in crystal free context.

4.3.4 Structural stability of the wildtype oligomers: hydrogen bond analysis

The amyloid configuration and properties primarily depend on the density of hydrogen bonds involving the backbone of the polypeptides, while the side chains hydrogen bonds are involved in the geometrical details and extension of the disordered parts of the structure.[161,162] To further characterize the structural stability of the studied oligomer models we determined the number of the hydrogen-bonds as the function of the MD simulation. The numbers of hydrogen bonds in the GNNQQNY and NNQNTF systems are larger than the corresponding hydrogen in the other remaining models (see Figure 4-8). The amyloid configuration and properties primarily depend on the density of hydrogen bonds involving the backbone of the polypeptides, while the side chains hydrogen bonds are involved in the geometrical details and extension of the disordered parts of the structure.[162] Figure 4-8 shows the number of hydrogen bonds throughout the simulation for the various models studied. The results of the analysis of total, main chain and side chain hydrogen bonds indicate the Q/N rich models content more hydrogen bonds. The hydrogen bond analysis indicates the GNNQQNNY and NNQNTF are more stable with an

average of 8 and 5.5 hydrogen bonds per strand. The remaining other models with hydrogen bond of less than 3.5 are less stable. The hydrogen bond analysis in combination with the geometric stability analysis (RMSF, RMSD etc.) were found to support that an increase in side chain and main chain hydrogen bonds increase the stability of the short peptide oligomers.

We also performed statistical analysis on the number of hydrogen bonds to determine their occupancy. Hydrogen bond occupancy is defined as ratio of times where the hydrogen bond is present relative to the total time length of the considered trajectory. Hence, hydrogen bonds that are never disrupted correspond to unit occupancy. Hydrogen bonds were characterized according to their temporal occupancy during the 20 ns simulation using criteria based on angle and distance. The strength of a hydrogen bond can be characterized by two geometric quantities which govern the hydrogen bond energy: hydrogen bond angle, D–H · ·A atoms and optimal hydrogen bond length, H · ·A (or D · ·A) distance [76]. Hydrogen bond occupancies and structure RMSDs was calculated using PTRAJ module available within AMBER. A hydrogen bond is assigned if the distance between donor D and acceptor A is ≤3.5 Å and the angle D-H …A≥120^{o}.[75] Those hydrogen bonds in the backbone with a fractional occupancy greater than 50% (considered to be strong hydrogen bonds) are detailed in Figure 4-9. The hydrogen bond occupancy analysis revealed that a large value of average occupancy of hydrogen bonds is in general associated with small average structural fluctuations and greater stability (see Figure 4-9 and Figure 4-3). The stability of the Q/N rich oligomers was found be supported by large values of average hydrogen bond occupancy compare to the oligomer lacking Q/N amino acid side chains. The MVGGVV segment pairs with antiparallel arrangement of strand per sheets lacks

side chain hydrogen bonds. The absence of the side chain hydrogen bond, despite a similar number of total hydrogen bond contents and occupancy as in the slightly stable SSNTV forms , might be the main reason for it instability. The highly stable Q/N rich models contents a greater number of side chain hydrogen bond which responsible for retaining the initial microcystal structure in crystalline free physiological environment.

4.3.5 Effect of mutation of Q/N residues at steric zipper on the stability of NNQNTF and GNNQQNT

Side chain mutagenesis has proven to be a very effective means of identifying energetically important backbone H-bonds and side chain interactions (including hydrophobic and electrostatic interactions) in peptides.[163] To test the hypothesis that the stability of the Q/N rich oligomers is due to the presence of the extra side chain hydrogen bonding we did Gly-mutations of the Q and N side chains involved in the steric zipper for GNNQQNY and NNQNTF wildtype oligomers with Hydrogen (performed *in silico* mutations) The summary of the structural stability of the mutants are shown in Table 4-2. The simulation revealed that mutation of the Q and N residues at the steric zipper for the GNNQQNY oligomer affect the structural stability. In the case of mutants Q4G and N6G after the 20 ns simulation one of the β-strands starts to separate from the remaining aggregates (see Figure 4-10). The mutation of Asn-2, Gln-4, or Asn-6 by Gly could disrupt the steric zipper, leading to unstable oligomers. The hydrogen bond analysis showed the mutants had a smaller backbone-backbone and side chain-side chain hydrogen bonds compared to the wildtype. The simulation of the NNQNTF mutants (N1G and Q3G) showed mutant Q4G is having a larger RMSD, RMSF and smaller hydrogen bond contents with significant structural

instability compared to the wildtype (Table2 and Figure 4-10). The mutation of the polar Asn and Gln side chains at the steric zipper with Gly revealed the polar side chains of Asn and Gln are important for the stability the oligomers of GNNQQNY and NNQNTF.

4.4 Conclusions

The major findings of our 20 ns multiple MD simulation suggests the following qualitative conclusions:

1. The short segments amyloid peptide rich in Q/N amino acid are have greater structural and the packing polymorphs are stable and under crystal free contents and physiological environments they might have similar probability of occurrence.

2. The short segments amyloid peptide lacking the Q/N amino acid have been found to be structurally unstable and the packing polymorphs show different stability with greater chance different probability of occurrence under crystal free contents and physiological environments.

3. The simulations of Q/N→G mutants disrupted the steric zipper, leading to unstable oligomers. The comparison of the structural stability of the wildtype and mutants stability of the Q/N rich oligomers was found be supported by large values of average hydrogen bond occupancy of the wild type compare to mutants.

4. The Q/N residue rich short amyloid segments have larger hydrogen bond contents and hydrogen bond occupancy. The overall increase of hydrogen bond in the Q/N residue rich peptides with smaller RMSD, RMSF and greater stability suggesting the stability of oligomer models is associated with an increase in side chain and total hydrogen bond contents.

5. The MM-PBSA binding free energy method was applied to the study of the β-sheet association. The nonpolar component of free energy is more favorable, while the electrostatic solvation is unfavorable for sheet to sheet interaction. The decomposition of the binding energy per residue showed the contribution of the N/Q side-chains to the association of the 5 stranded double layer oligomers is larger than the other nonpolar and small size amino acids at the interface, underlining the importance of Q/N amino acid in stabilizing the short segment amyloid peptides in crystal free context.

Our simulations provide detailed insight into the structural stability of various short segment amyloid oligomer aggregates Exploring the structural stability and aggregation behavior of the short peptides may gain insights into the self-assembly process at the early stage of fibril formation and provide a clue to structure based design of amyloid aggregation inhibitors. The rational design of successful therapeutic strategies requires detailed characterization of amyloid formation. Polymorphism in amyloid peptides with the same sequence is due to difference in β-sheet packing (steric zipper).[22] Landau *et al* [16] found that different aggregation inhibitor molecules bind to different polymorphs of amyloid peptides and suggested a combination of aggregation inhibitors might be required to bind to the various morphologic form of a given amyloid peptide. Sievers *et al* [164] using known atomic structures of segments of amyloid fibrils as templates have designed amyloid aggregation inhibitors. Our simulations indicate that there is a difference in the stability polymorphs of a given sequence and certain amino acids are significantly important for stability. Results from this work might be useful in designing

peptidomimetic aggregation inhibitors by single amino acid change or shuffling the sequence so as to disrupt the steric zipper and prevent amyloid aggregation.

Table 4-1 Summary of simulation system and condition of double layer models of amyloid peptide segments with packing polymorphism

Model name	Sheet organization	# peptide/ water mol.	PDB code	Length (ns) simulation	Simulation box size (Å)	T., K	Number of simulation
SSTNVG 1 (*SH2-ST5*)	Parallel/parallel	760/2633	3DG1	20	58.01×58.01×58.01	300	5
SSTNVG 2 (*SH2-ST5*)	Parallel/parallel	760/2689	3FTR	20	58.41×58.41×58.41	300	5
VQIVYK 1 (*SH2-ST5*)	Parallel/parallel	1140/4575	2ON9	20	69.00×69.00×69.00	300	5
VQIVYK 2 (*SH2-ST5*)	Parallel/parallel	1140/3890	3FQP	20	65.88×65.88×65.88	300	5
NNQNTF 1 (*SH2-ST5*)	Parallel/parallel	960/2967	3FVA	20	60.50×60.50×60.50	300	5
N1G *(SH2-ST5*)	Parallel/parallel	890/3006	-	20	60.64×60.64×60.64	300	1
Q3G*(SH2-ST5*)	Parallel/parallel	860/3021	-	20	60.74×60.74×60.74	300	1
NNQNTF 2 (*SH2-ST5*)	Parallel/parallel	960/3482	3FVA	20	63.35×63.35×63.35	300	5
GNNQQNY 1 (*SH2-ST5*)	Parallel/parallel	1070/3284	1YJP	20	62.42×62.42×62.42	300	5
N2G*(SH2-ST5*)	Parallel/parallel	1000/3351	-	20	62.88×62.88×62.88	300	1
Q4G*(SH2-ST5*)	Parallel/parallel	970/3397	-	20	63.09×63.09×63.09	300	1
N6G*(SH2-ST5*)	Parallel/parallel	1000/3340	-	20	62.76×62.76×62.76	300	1
GNNQQNY 2 (*SH2-ST5*)	Parallel/parallel	1070/3241	2OMM	20	62.25×62.25×62.25	300	5
MVGGVV 1 (*SH2-ST5*)	Parallel/antiparallel	820/4200	2ONA	20	67.13×67.13×67.13	300	5
MVGGVV 2 (*SH2-ST5*)	Parallel/antiparallel	820/3785	2OKZ	20	64.98×64.98×64.98	300	5

Table 4-2 Summary of structural analysis of double layer amyloid peptide segments with packing polymorphism

Model name	<RMSD>	<RMSF>	Simulation (Å)		Crystal (Ås)		Average number of hydrogen bonds		
			$<d_{sheet}>$	$<d_{strand}>$	$<d_{sheet}>$*	$<d_{strand}>$‡	<Total $_{H\text{-}bond}$>	Main-chain	Side-chain
SSNTVG-I	3.52(1.22)	1.76(0.52)	7.75(1.04)	6.33(1.01)	6.22	4.79	39.09(1.92)	24.70(1.21)	16.84(0.88)
SSNTVG-II	3.53(0.46)	1.64(0.36)	10.06(0.22)	5.69(0.37)	8.96	4.79	36.68(0.85)	21.35(1.38)	13.15(1.21)
VQIVYK-I	2.45(0.55)	1.34(0.2)	13.64(0.5)	5.57(0.27)	11.03	4.86	35.41(1.81)	29.67(1.43)	4.7(0.77)
VQIVYK-II	5.49(1.12)	1.99(0.22)	15.61(2.65)	5.81(0.16)	9.16	4.86	30.54(1)	26.79(1.11)	3.59(1.11)
NNQNTF-I	1.73(0.45)	0.79(0.09)	8.85(0.19)	5.18(0.08)	8.50	4.84	58.41(2.11)	24.63(0.90)	29.34(3.25)
N1G#	1.78(0.41)	0.78(0.50)	7.85(0.29)	5.17(0.13)	8.50	4.84	49.22(4.08)	26.9(1.93)	20.86(2.87)
Q3G#	2.39(0.53)	0.80(0.58)	6.83(0.18)	5.08(0.20)	8.50	4.84	37.49(4.23)	20.53(2.38)	18.48(2.34)
NNQNTF-II	2.06(1.43)	0.85(0.28)	10.97(0.37)	5.08(0.08)	10.52	4.84	51.95(0.52)	25.90(1.60)	13.19(1.10)
GNNQQNY-I	0.93(0.07)	0.55(0.03)	9.74(0.04)	4.96(0.01)	9.61	4.87	79.93(0.94)	29.38(0.68)	48.80(1.46)
N2G#	0.84(0.19)	0.55(0.41)	8.62(0.14)	6.98(0.09)	8.50	4.84	73.48(2.89)	31.06(0.98)	41.64(2.70)
Q4G#	2.52(0.77)	1.41(0.97)	8.74(0.55)	5.35(0.22)	8.50	4.84	37.49(4.23)	20.46(3.79)	13.01(3.64)
N6G#	1.77(0.21)	0.79(0.57)	8.91(0.23)	5.48(0.240	8.50	4.84	68.71(2.94)	29.21(1.42)	37.57(2.45)
GNNQQNY-II	0.86(0.05)	0.53(0.02)	9.66(0.06)	4.93(0.01)	10.04	4.93	76.43(0.86)	21.45(0.67)	51.11(1.60)
MVGGVV-I	4.06(0.94)	2.09(0.34)	8.53(0.89)	5.36(0.29)	7.50	4.84	33.14 (2.7)	32.27(1.94)	0.7(0.40)
MVGGVV-II	4.49(0.95)	2.13(0.25)	8.86(0.86)	5.53(0.26)	8.01	4.76	33.17(2.94)	32.07(2.65)	0.6(0.41)

* The inter-sheet distance (d_{sheet}) was calculated by averaging pair-wise residue Cα-Cα distances between the adjacent β-sheet as the shortest distance between two the main chain atoms of two opposing strands in the dry interface: (**SSNTVG-I**: S-N,S-T,T-S and N-S; **SSNTVG-II**: S-V,T-N, N-T and V-S; **VQIVYK-I**: V-Y, I-I and Y-V, **VQIVYK-II** : V-I and I-V; **NNQNTF-I** and **NNQNTF-II**:N-F, N-T ,Q-N , N-Q, T-N , and F-N; **GNNQQNY-I** and G**NNQQNY-II**: G-Y, N-N, N-Q , Q-Q , Q-N, N-N and Y-G; **MVGGVV-I**: V-G,G-V and G-M , and; **MVGGVV-II**: ,G-M,V-G , and G-V)

‡ Angular brackets < > indicate time averaging and the mean values and standard deviation (SD) were calculated by using the five trajectories for each models. Mutations were done using the initial structure of the on the most stable polymorphic forms and the analysis was done on a single trajectory

Table 4-3 Summary of the MM-PBSA components of the double layer amyloid peptide segments with packing polymorphism

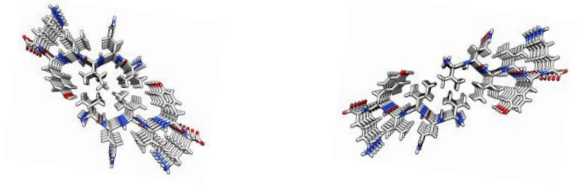
Contribution	**Model name**					
	SSNTVG-I	**SSNTVG-II**	**NNQNTF-I**	**NNQNTF-II**	**GNNQQNY-I**	**GNNQQNY-II**
$\langle\Delta E_{vdw}\rangle$	-14.83(1.80)	-14.14(0.65)	-23.58(0.53)	-19.76(0.42)	-20.54(0.13)	-20.62(0.05)
$\langle\Delta E_{ele}\rangle$	-228.34(17.87)	-209.97(13.82)	-187.71(2.91)	-97.40(4.50)	-174.86(1.98)	-176.04(1.18)
$\langle\Delta G_{PB}\rangle$	230.13(16.81	210.46(12.04)	193.04(4.80)	104.50(4.59)	178.55(2.24)	179.95(1.34)
$\langle\Delta G_{SA}\rangle$	-1.40(0.10)	-1.39(0.07)	-1.95(0.05)	-1.53(0.04)	-1.56(0.01)	-1.57(0.01)
$\langle\Delta G_{solv}\rangle$	228.73(16.72)	209.06(12.08)	191.08(4.76)	102.97(4.55)	176.99(2.24)	178.38(1.13)
$\langle\Delta G_{subtotal}\rangle$	-14.44(2.33)	-15.05(2.06)	-20.204(1.89)	-14.19(0.37)	-18.41(0.18)	-18.28(0.10)
$\langle -T\Delta S\rangle$	-9.16(0.58)	-9.28(0.50)	-10.29(0.48)	-7.52(0.28)	-8.97(0.32)	-9.12(0.13)
$\langle\Delta G_{(per\ stand)}\rangle$	-5.27(1.77)	-5.77(1.77)	-9.91(2.14)	-6.68(0.48)	-9.44(0.31)	-9.16(0.06)

Binding free energy components (kcal mol^{-1}) and standard deviations calculated with MM-PBSA for oligomer double-layers (SH2-ST5) models: Average over 5000 snapshots of the trajectory. [b] Entropy calculations were based on normal modes analysis using 100 snapshots (1 frame taken at an interval of 50 frams from the total of 5000 frames). Δ $\mathbf{E_{vdw}}$, non-bonded van der Waals energy; ΔE_{ele}, non-solvent electrostatic potential energy; ΔG_{PB}, electrostatic contributions to the solvation free energy calculated with Poisson-Boltzmann equation; ΔG_{SA} , ΔG_{Solv} are nonpolar and total solvation energies; $-T\Delta S$; Entropic contributions to binding. ΔS; sum of rotational, translational and vibrational entropies; $\Delta G_{(per\ stand)}$, per strand binding energy of the system. All energies are in kcal/mol.
$\Delta G_{subtotal} = \Delta E_{vdw} + \Delta E_{ele} + \Delta G_{sol}$; $\Delta G_{sol} = \Delta G_{PB} + \Delta G_{SA}$; $\Delta E_{gas} = \Delta E_{elect} + \Delta E_{vdW}$ and $\Delta G_{(per\ stand)} = \Delta G_{subtotal} - T\Delta S$

* Mean values were calculated by using the five trajectories for each model from the multiple simulations. Standard deviation (SD) was also calculated by using the five trajectories.

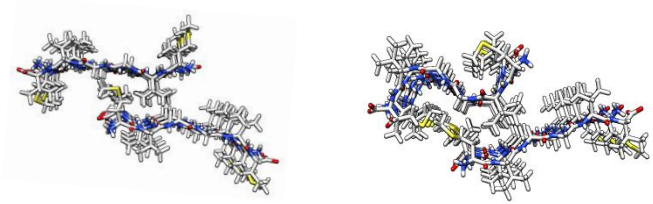
A

SSTNVG form I **SSTNVG form II**



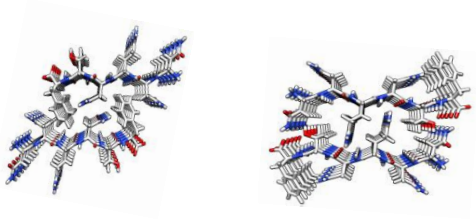
B

VQIVYK form I **VQIVYK form I**



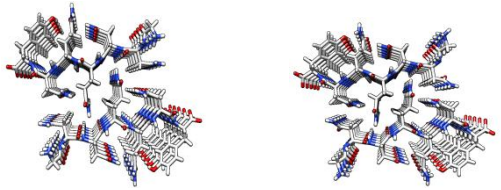
C

MVVGGVV form I **MVVGGVV form II**



D

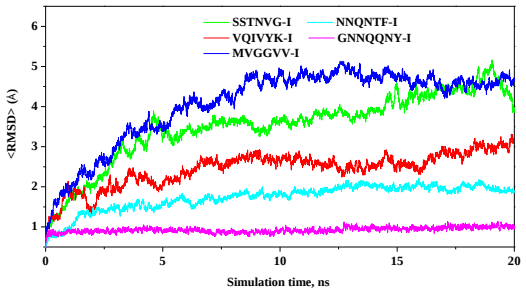
NNQNTF form I **NNQNTF form II**



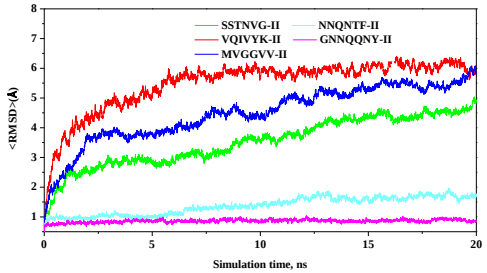
E

GNNQQNY form I **GNNQQNY form I**

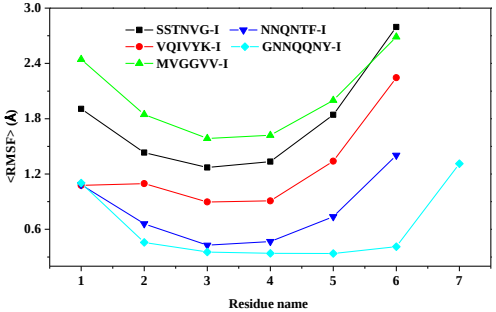
Figure 4-1 Structure of polymorphism of steric zippers of amyloid peptide segments studied (A) SSTNVG from IAPP, (B) VQIVYK from tau protein, (C) NNQNTF from elk prion,) GNNQQNY form yeast prion Sup35 and (D) MVGGVV from Abeta.



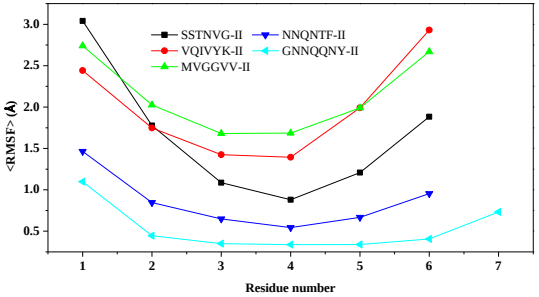
A



B



C



D

Figure 4-2 RMSD and RMSF values as a function of time. The RMSD and RMSF were calculated by averaging over five trajectories compared to the starting structure in each models. (A) RMSD for polymorph form I, (B) RMSD for polymorph form II, (C) RMSF for polymorph form I and (D) RMSF for polymorph form II. The RMSF values are average for 10 β-strands.

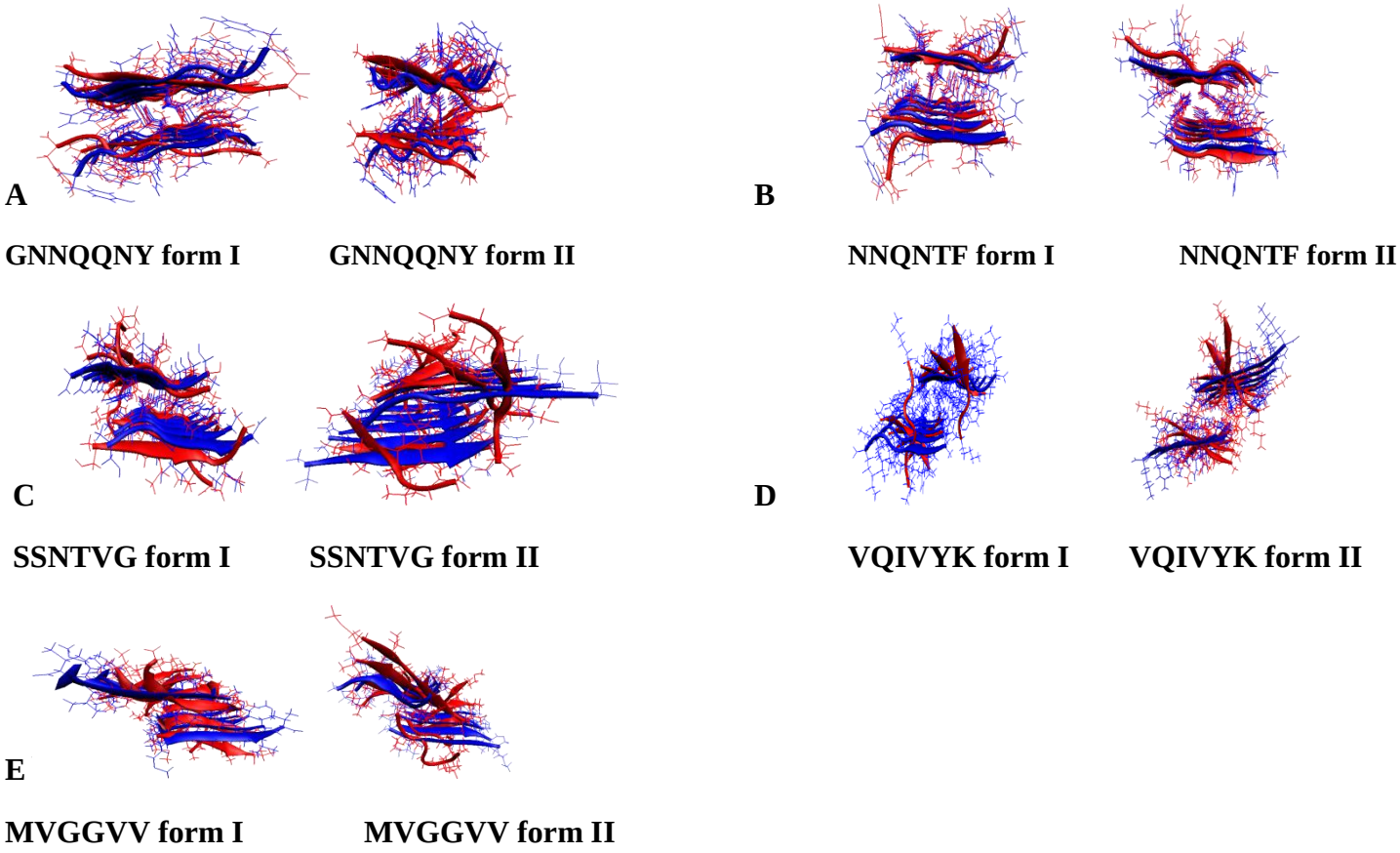
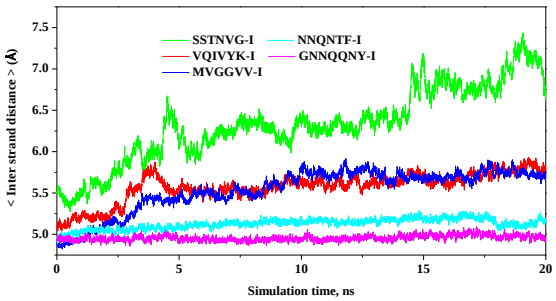
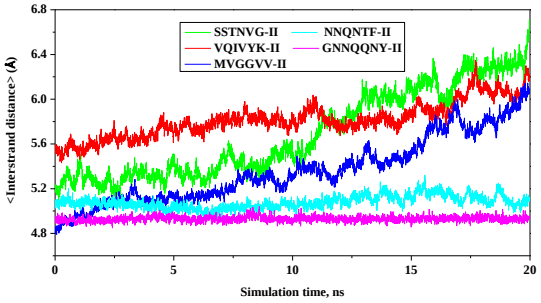


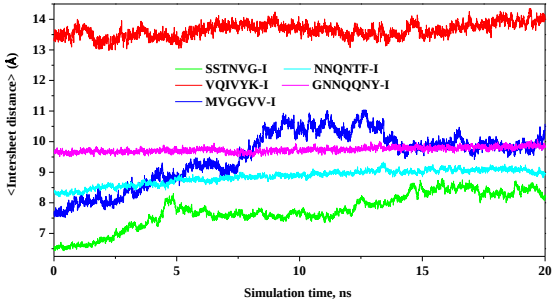
Figure 4-3 Superposition of the backbone atoms of the staring conformation with the conformation at 20 ns. The initial structures are colored in blue while the structure after 20 ns simulation is shown in red. The structure for the conformation after 20 ns was taken from the trajectory with the smallest RMSD and RMSF values out of the five trajectories for each system.



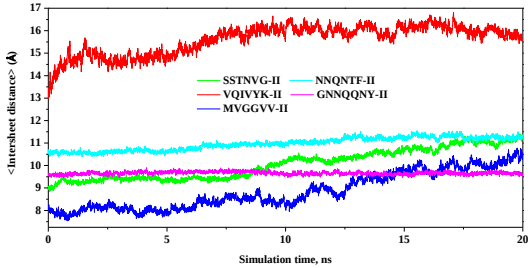
A



B

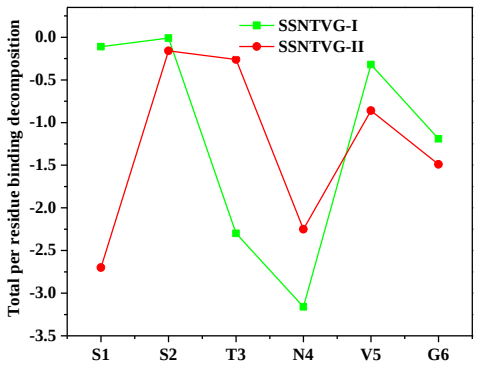


C

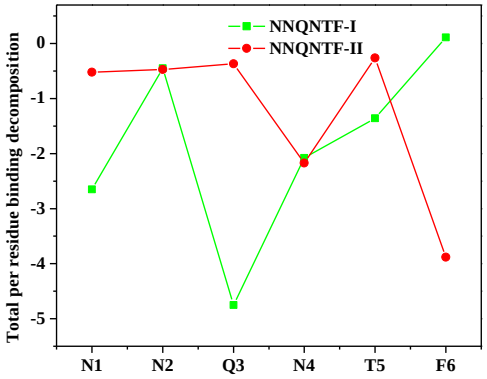


D

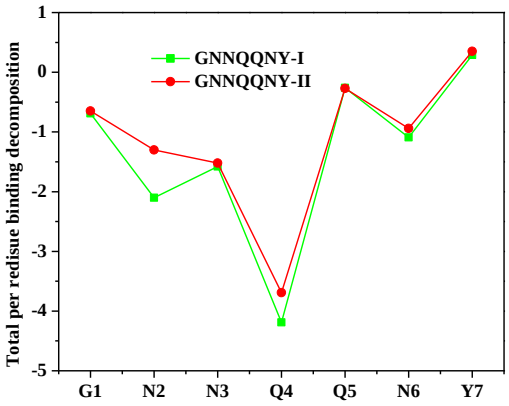
Figure 4-4 The averaged inter-strand distances (A and B) and inter-sheet distances (C and D) calculated by averaging over five trajectories. The distances were measured in comparison with the corresponding initial structure in each model.



A

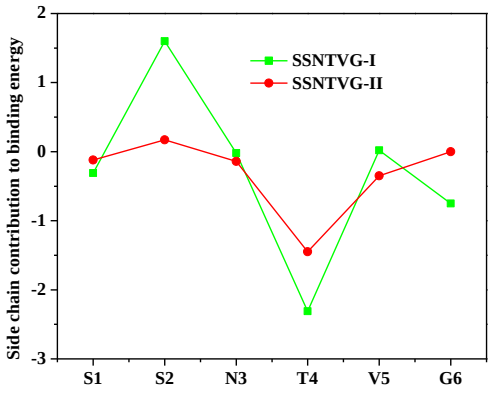


B

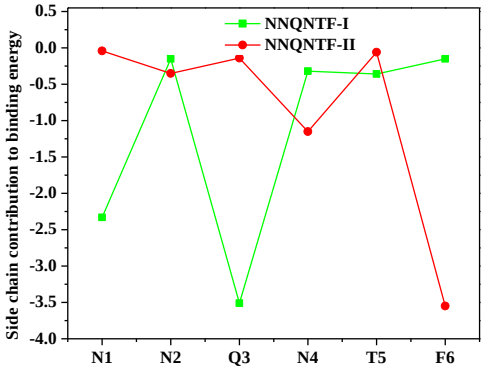


C

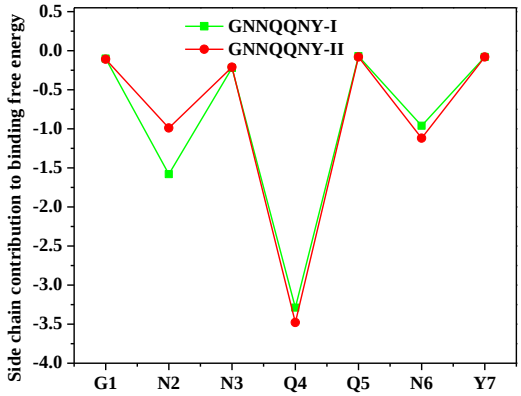
Figure 4-5 MM-PBSA per residue decomposition of total binding free energy. (A) SSNTVG polymorph form I and II , (B) NNQNTF polymorph form I and II and (C) GNNQQNY polymorph form I and II.



A

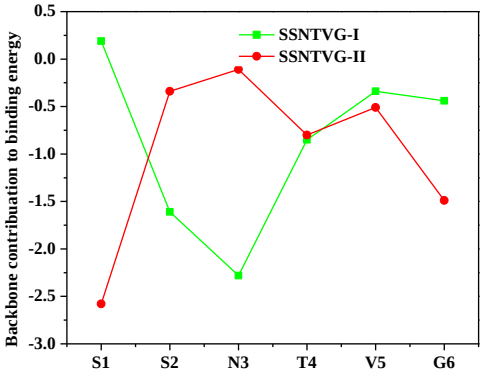


B

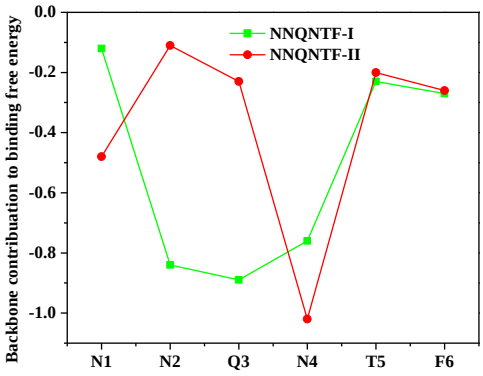


C

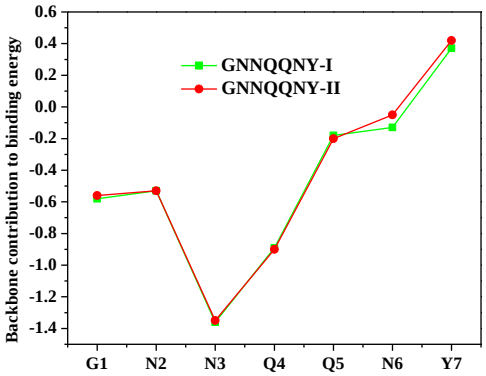
Figure 4-6 MM-PBSA per residue decomposition of side chain contribution to binding free energy. (A) SSNTVG polymorph form I and II , (B) NNQNTF polymorph form I and II and (C) GNNQQNY polymorph form I and II.



A

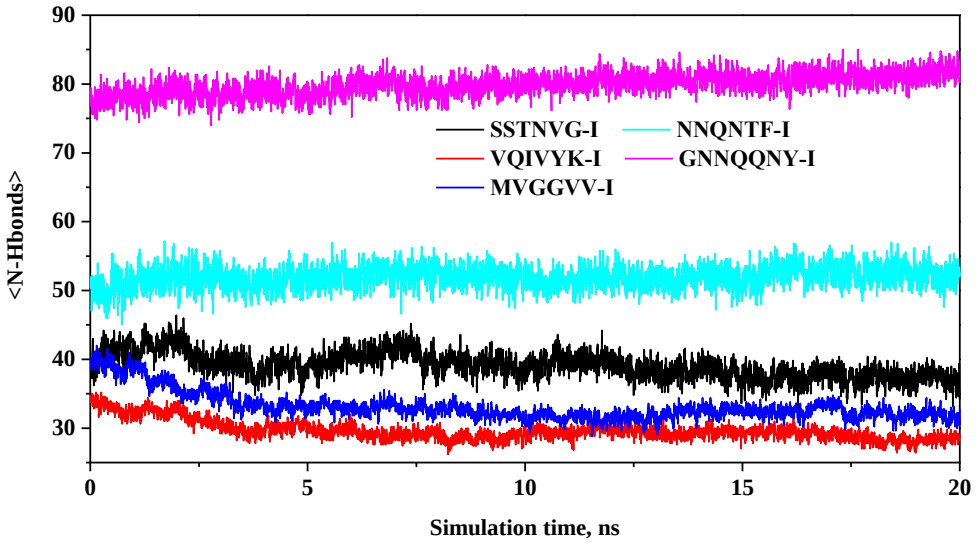


B

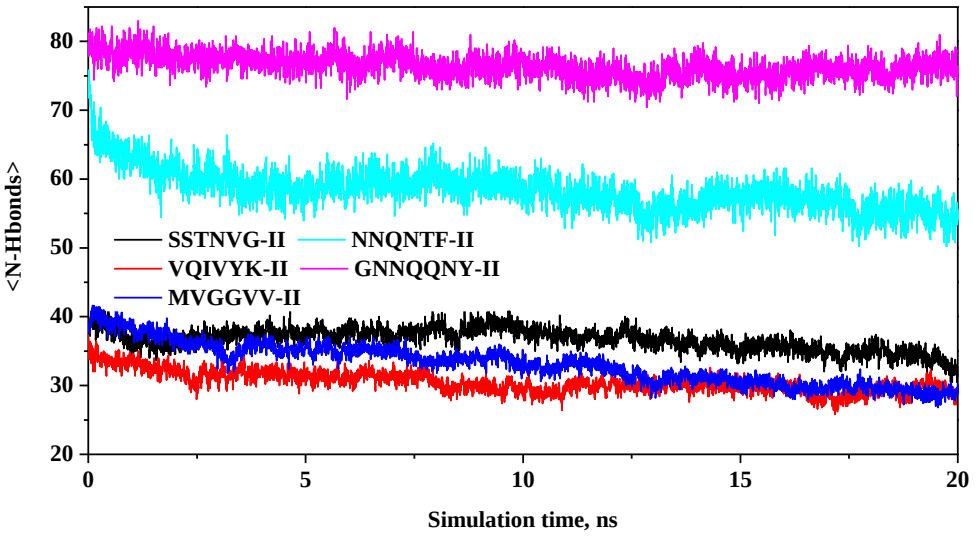


C

Figure 4-7 MM-PBSA per residue decomposition of backbone contribution to binding free energy. (A) SSNTVG polymorph form I and II , (B) NNQNTF polymorph form I and II and (C) GNNQQNY polymorph form I and II.



A



B

Figure 4-8 Comparison of the number of H-bonds as a function of MD simulation, for five pairs of polymorphic packing of amyloid oligomers: SSTNVG, VQIVYK, MVGGVV, NNQNTF and GNNQQNY. The hydrogen bonds were determined with respect to the energy minimized structures.

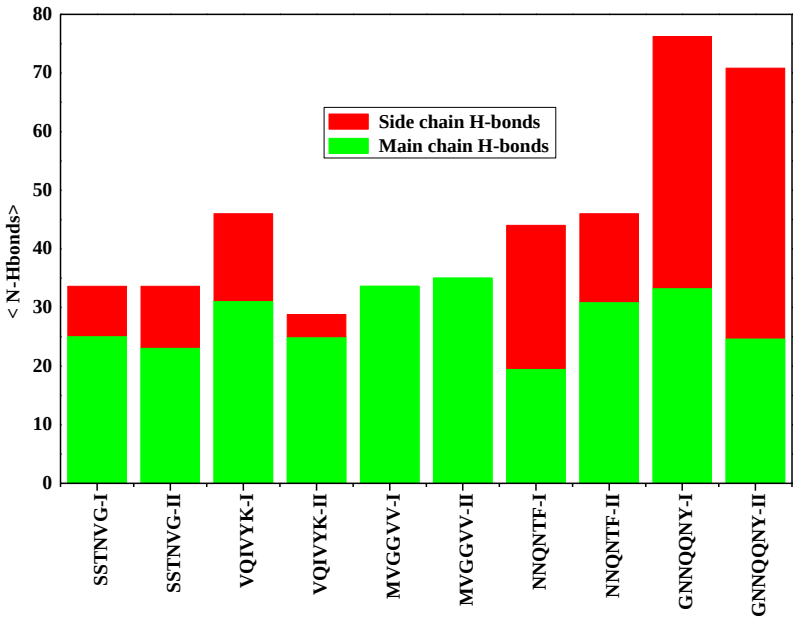
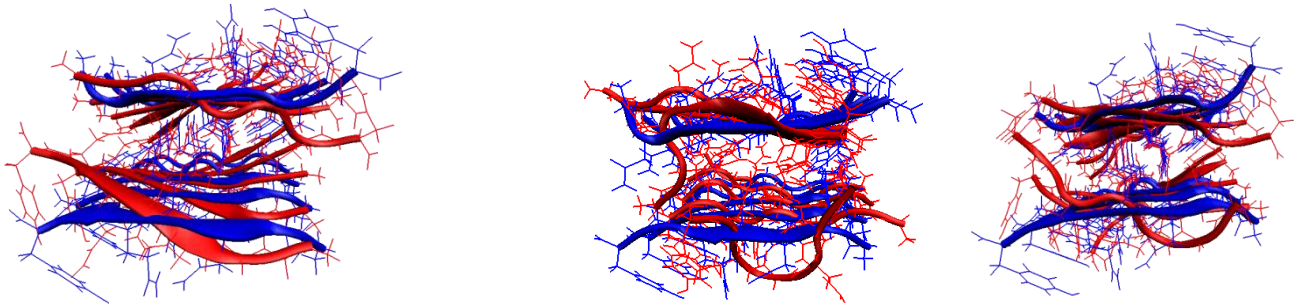


Figure 4-9 The hydrogen bond occupancy for the side-chain–side-chain and main-chain–side-chain atoms throughout the simulations. The cut off used for H-bond distance and angle for each system was 3.5 Å and 1200. The average occupancy of the main chain and side chain hydrogen bonds were calculated by using five trajectories. The average is over the entire simulation trajectories. Hydrogen bonds with occupancy $\geq$ 50% are considered here

GNNQQNY mutants: ***N2G*** ***Q4G*** ***N6G***



NNQNTF mutants: ***N1G*** ***Q3N***

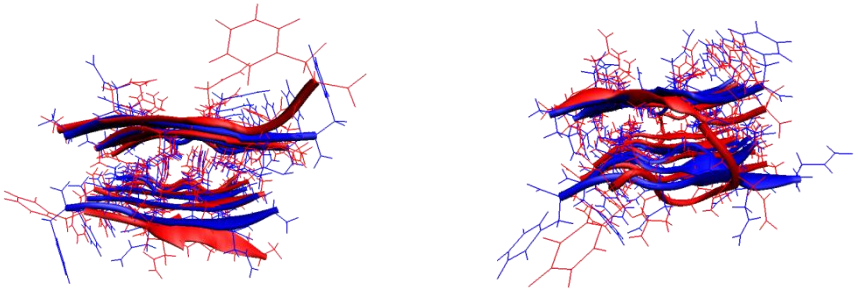


Figure 4-10 Superposition of the backbone atoms of the staring conformation of the Q/N rich mutants with the conformation at 20 ns. The initial structures are colored in blue while the structure after 20 ns simulation is shown in red. The structure for the conformation after 20 ns was taken from the trajectory with the smallest RMSD and RMSF values out of the five trajectories for each system.

4.5 References

132. Fandrich, M.; Meinhardt, J.; Grigorieff, N., Structural polymorphism of Alzheimer A beta and other amyloid fibrils. *Prion* **2009,** 3, (2), 89-93.

133. Miller, Y.; Ma, B.; Nussinov, R., Polymorphism in Alzheimer A beta Amyloid Organization Reflects Conformational Selection in a Rugged Energy Landscape. *Chemical Reviews* **2010,** 110, (8), 4820-4838.

134. Apostol, M. I.; Sawaya, M. R.; Cascio, D.; Eisenberg, D., Crystallographic Studies of Prion Protein (PrP) Segments Suggest How Structural Changes Encoded by Polymorphism at Residue 129 Modulate Susceptibility to Human Prion Disease. *Journal of Biological Chemistry* **2010,** 285, (39), 29671-29675.

135. Andersen, C. B.; Hicks, M. R.; Vetri, V.; Vandahl, B.; Rahbek-Nielsen, H.; Thogersen, H.; Thogersen, I. B.; Enghild, J. J.; Serpell, L. C.; Rischel, C.; Otzen, D. E., Glucagon Fibril Polymorphism Reflects Differences in Protofilament Backbone Structure. *Journal of Molecular Biology* **2010,** 397, (4), 932-946.

136. Calamai, M.; Chiti, F.; Dobson, C. M., Amyloid fibril formation can proceed from different conformations of a partially unfolded protein. *Biophysical Journal* **2005,** 89, (6), 4201-4210.

137. Chiti, F.; Dobson, C. M., Protein misfolding, functional amyloid, and human disease. *Annual Review of Biochemistry* **2006,** 75, 333-366.

138. Jones, E. M.; Surewicz, W. K., Fibril conformation as the basis of species- and strain-dependent seeding specificity of mammalian prion amyloids. *Cell* **2005,** 121, (1), 63-72.

139. Petkova, A. T.; Leapman, R. D.; Guo, Z. H.; Yau, W. M.; Mattson, M. P.; Tycko, R., Self-propagating, molecular-level polymorphism in Alzheimer's beta-amyloid fibrils. *Science* **2005,** 307, (5707), 262-265.

140. Chafekar, S. M.; Baas, F.; Scheper, W., Oligomer-specific A beta toxicity in cell models is mediated by selective uptake. *Biochimica Et Biophysica Acta-Molecular Basis of Disease* **2008,** 1782, (9), 523-531.

141. Deshpande, A.; Mina, E.; Glabe, C.; Busciglio, J., Different conformations of amyloid beta induce neurotoxicity by distinct mechanisms in human cortical neurons. *Journal of Neuroscience* **2006,** 26, (22), 6011-6018.

142. Iijima, K.; Chiang, H. C.; Hearn, S. A.; Hakker, I.; Gatt, A.; Shenton, C.; Granger, L.; Leung, A.; Iijima-Ando, K.; Zhong, Y., A beta 42 Mutants with Different Aggregation Profiles Induce Distinct Pathologies in Drosophila. *Plos One* **2008,** 3, (2), 8.

143. Poirier, M. A.; Li, H. L.; Macosko, J.; Cai, S. W.; Amzel, M.; Ross, C. A., Huntingtin spheroids and protofibrils as precursors in polyglutamine fibrilization. *Journal of Biological Chemistry* **2002,** 277, (43), 41032-41037.

144. Gosal, W. S.; Morten, I. J.; Hewitt, E. W.; Smith, D. A.; Thomson, N. H.; Radford, S. E., Competing pathways determine fibril morphology in the self-assembly of beta(2)-microglobulin into amyloid. *Journal of Molecular Biology* **2005,** 351, (4), 850-864.

145. Goldsbury, C.; Frey, P.; Olivieri, V.; Aebi, U.; Muller, S. A., Multiple assembly pathways underlie amyloid-beta fibril polymorphisms. *Journal of Molecular Biology* **2005,** 352, (2), 282-298.

146. Deechongkit, S.; Nguyen, H.; Powers, E. T.; Dawson, P. E.; Gruebele, M.; Kelly, J. W., Context-dependent contributions of backbone hydrogen bonding to beta-sheet folding energetics. *Nature* **2004,** 430, (6995), 101-105.

147. Tsemekhman, K.; Goldschmidt, L.; Eisenberg, D.; Baker, D., Cooperative hydrogen bonding in amyloid formation. *Protein Science* **2007,** 16, (4), 761-764.

148. Papaleo, E.; Invernizzi, G., Conformational Diseases: Structural Studies of Aggregation of Polyglutamine Proteins. *Current Computer-Aided Drug Design* **2011,** 7, (1), 23-43.

149. Berhanu, W. M.; Masunov, A. E., Molecular dynamic simulation of wildtype and mutants of the polymorphic amyloid NNQNTF segments of elk prion: structural stability and thermodynamic of association. *Biopolymers* **2011,** 95, 573-589.

150. De Simone, A.; Pedone, C.; Vitagliano, L., Structure, dynamics, and stability of assemblies of the human prion fragment SNQNNF. *Biochemical and Biophysical Research Communications* **2008,** 366, (3), 800-806.

151. Wu, C.; Bowers, M. T.; Shea, J. E., Molecular Structures of Quiescently Grown and Brain-Derived Polymorphic Fibrils of the Alzheimer Amyloid A beta(9-40) Peptide: A Comparison to Agitated Fibrils. *Plos Computational Biology* **2010,** 6, (3), 12.

152. Paparcone, R.; Sanchez, J.; Buehler, M. J., Comparative Study of Polymorphous Alzheimer's A beta(1-40) Amyloid Nanofibrils and Microfibers. *Journal of Computational and Theoretical Nanoscience* **2010,** 7, (7), 1279-1286.

153. Ma, B. Y.; Nussinov, R., Polymorphic C-terminal beta-Sheet Interactions Determine the Formation of Fibril or Amyloid beta-derived Diffusible Ligand-like Globulomer for the Alzheimer A beta 42 Dodecamer. *Journal of Biological Chemistry* 285, (47), 37102-37110.

154. Berryman, J. T.; Radford, S. E.; Harris, S. A., Systematic Examination of Polymorphism in Amyloid Fibrils by Molecular-Dynamics Simulation. *Biophysical Journal* **2011,** 100, (9), 2234-2242.

155. Humphrey, W.; Dalke, A.; Schulten, K., VMD: Visual molecular dynamics. *Journal of Molecular Graphics* **1996,** 14, (1), 33-&.

156. Massova, I.; Kollman, P. A., Computational alanine scanning to probe protein-protein interactions: A novel approach to evaluate binding free energies. *Journal of the American Chemical Society* **1999,** 121, (36), 8133-8143.

157. Buchete, N. V.; Hummer, G., Structure and dynamics of parallel beta-sheets, hydrophobic core, and loops in Alzheimer's A beta fibrils. *Biophysical Journal* **2007,** 92, (9), 3032-3039.

158. Huet, A.; Derreumaux, P., Impact of the mutation A21G (Flemish variant) on Alzheimer's beta-amyloid dimers by molecular dynamics simulations. *Biophysical Journal* **2006,** 91, (10), 3829-3840.

159. Miller, Y.; Ma, B. Y.; Nussinov, R., The Unique Alzheimer's beta-Amyloid Triangular Fibril Has a Cavity along the Fibril Axis under Physiological Conditions. *Journal of the American Chemical Society* 133, (8), 2742-2748.

160. Sievers SA, K. J., Chang HW, Zhao A, Jiang L, Zirafi O, Stevens JT, MÃ¼nch J, Baker D, Eisenberg D., Structure-based design of non-natural amino-acid inhibitors of amyloid fibril formation. *Nature Structural & Molecular Biology* **2011**, 1-7.

161. Paparcone, R.; Pires, M. A.; Buehler, M. J., Mutations Alter the Geometry and Mechanical Properties of Alzheimer's A beta(1-40) Amyloid Fibrils. *Biochemistry* **2010,** 49, (41), 8967-8977.

162. Paparcone, R.; Sanchez, J.; Buehler, M. J., Comparative Study of Polymorphous Alzheimer's A beta(1-40) Amyloid Nanofibrils and Microfibers. *Journal of Computational and Theoretical Nanoscience* 7, (7), 1279-1286.

163. Jager, M.; Dendle, M.; Kelly, J. W., Sequence determinants of thermodynamic stability in a WW domain-An all-beta-sheet protein. *Protein Science* **2009,** 18, (8), 1806-1813.

164. Sievers, S. A. K., J.; Chang, H. W.; Zhao, A.; Jiang, L.; Zirafi, O.; Stevens, J. T.; Munch, J.; Baker, D.; Eisenberg, D. , Structure-based design of non-natural amino-acid inhibitors of amyloid fibril formation. *Nature* **2011,** 475, 96-100.

CHAPTER 5 CONTROLLING THE AGGREGATION AND RATE OF RELEASE IN ORDER TO IMPROVE THE INSULIN FORMULATION

Chapter 5, in part, is a reprint of the material as it appears in Journal of Molecular Modeling, 2011, Workalemahu M. Berhanu & Artem E. Masunov, Published online: 15 June 2011; DOI 10.1007/s00894-011-1123-3

5.1 Background

Insulin is a 51-residue protein hormone consisting of two polypeptide chains: the chain A (comprising 21 residues) and the chain B (comprising 30 residues), linked together by three disulfide bonds (Figure 5-1A). Insulin is stored in the body in the secretory vesicles of the pancreas as a zinc-containing hexamer. When in the blood stream, insulin is present in its biologically active monomeric form.[165,166] The underproduction of insulin or lack of receptor sensitivity to insulin is known to cause diabetes that is affecting 171 million people worldwide.[167] Insulin is the mainstay of drug therapy for patients with type I diabetes, and can reduce the morbidity in the long term. The disease is caused by autoimmune destruction of insulin secreting β cells of the pancreas. Without sufficient levels of insulin, these patients cannot properly utilize glucose and typically have markedly elevated blood glucose (hyperglycemia) while intracellular glucose levels are generally low. The chronic complications of a consistently high blood sugar level are serious and include retinopathy (diabetes is the most common cause of blindness), neuropathy, nephropathy (diabetes is a leading cause of chronic renal failure), cardiovascular disease, peripheral vascular disease (diabetes is the leading cause of limb amputation) and makes the patient more susceptible to infection.[168]

Similar to many other proteins, insulin can misfold and form highly ordered fibrillar amyloid aggregates. Insulin fibrils have been observed *in vivo* following continuous subcutaneous insulin infusion [169] and repeated insulin injections [170]; they are the main factor in the pathogenesis called injection amyloidosis.[2,17] These insulin fibrils that form *in vivo* display the defining characteristics of amyloid aggregates associated with neurodegenerative diseases [111] including binding to the dye Congo red with "apple-green" birefringence, they show an elongated, unbranched fibrillar morphology[17], they exhibit nucleation-dependent polymerization, and they present a cross-β X-ray diffraction pattern.[2] Recently, serum samples from patients with Parkinson's disease have been found to display an autoimmune response to insulin oligomers and fibrils [110], possibly indicating the presence of insulin aggregates in this disease too. Insulin also forms amyloid-like fibrils *in vitro*, which are promoted by elevated temperatures, low pH, and increased ionic strength.[112,111] In addition, insulin fibril formation has been a limiting factor in long-term storage of insulin for treatment of diabetes. Amyloid fibrillation may cause problems during both production, storage and drug delivery of protein based pharmaceuticals.[94,6] In the case of commercial insulin, fibril formation is a problem in some of the isolation and purification steps, when pH is lowered 1-3.[111] The agitation of insulin solutions during transportation and in portable delivery systems may also induce fibrillar aggregation.[171,94,6] Moreover, in therapeutic use of protein drugs, it is essential to avoid the fibril formation, since amyloid fibrillated protein is biologically inactive [111,172] and may cause immunological responses in patients.[172,173] Future drug development may be aim to either stabilize native structure, inhibit the formation of crucial intermediates on the pathway of fibril formation, or to prevent interaction between fibrillation intermediates such as the partially unfolded monomer and oligomers.[6,111] Modifications of the amino acid sequence of insulin, such as single point

mutations, influence both insulin activity and protein aggregation.[174] The newer insulin analogues have several improvements due to their modified action profile.[175] Main advantages of short-acting preparations include faster onset of action and shorter duration time. Long-acting analogues afford structural changes that delay the onset of action, allow slow and continuous absorption into the systemic circulation, and prolong the duration, thus producing a time-concentration profile that imitates the normal insulin basal level and leads to physiological basal glycaemic control with less nocturnal hypoglycaemias.[176]

Upon aggregation, the molecule of insulin undergoes structural changes from a predominantly α-helical state to a β-sheet rich conformation, and many models of insulin fibrils have been suggested[6,111,177] repeatedly. The segment B11-B17 with sequence LVEALYL is the smallest segment that can both nucleate and inhibit the fibrillation of full-length insulin, depending on the molar ratio. This fact is suggesting that this segment is central to the cross-beta spine of the insulin fibril.[94] In addition, the point mutations H10D and L17Q in the chain B of insulin prolong the lag phase of insulin fibrillation, further supporting the importance of this segment in fibril formation.[178] Also, exposing this fibril-prone segment by truncating the five residues of the chain B C-terminal increases the propensity of insulin for fibril formation.[179]

Recent studies have shown that the chain A also contributes to insulin fibrillation. Both chain A and chain B can form fibrils on their own [108,109] and seeds of these chains can nucleate the fibrillation of full length insulin.[108] In addition, it was reported that segments as short as six residues from either chain A (residues A13–A18) or chain B (residues B12–B17) can form fibrils by themselves.[180]

The first atomic-resolution view of the fibrillar spine came from single crystal structures of the segments LYQLEN (residues A13–A18) and VEALYL (residues B12–B17).[22] The

combination of several complementary techniques (including X-ray diffraction of insulin fibrils and scanning-transmission electron microscopy analysis of the morphology of insulin fibrils) allowed a highly reliable structure of full-length insulin amyloid fibrils to be constructed.[94,181,182,183] This model has a β-solenoid structure consisting of repeated structural units of similar but not identical peptides, covalently connected by 2 disulfide bonds. [94,183] The solenoids are linked by a dry steric zipper formed by the mating of the central two LVEALYL (residues B11-B17) strands. Because LYQLENY contains a Tyr residue in the second position, this side chain superimposes on a Tyr from LVEALYL preserving the ‘‘kissing tyrosine’’ interaction observed across the wet interface of the crystal of LVEALYL (Figure 5.1).

Computational studies have complemented experiment to provide insights into insulin aggregation. All-atom molecular dynamics (MD) simulations have been applied to study the amyloid oligomer stability by testing different candidate β-sheet arrangements of preformed oligomers mimicking possible nucleus seeds at the very early stage of fibril formation.[21,184,185] Mark *et al.*[186] performed series of shorter molecular dynamics simulations to investigate the structure of monomeric and dimeric insulin in aqueous solution. Their simulation showed that in the absence of crystal contacts both monomeric and dimeric insulin have a high degrees of intrinsic flexibility in the absence of crystal contacts. Monomer MD simulations[187,188] established that the proposed binding site for glucose is stable, both statically and dynamically.[189] Other MD simulations of the insulin dimer (but not monomer) have been published.[190,191] They reveal details concerning the dynamics of the dimer during the simulation, including the hydrogen bond pattern and correlated motions.

In this study, we report on an MD study of the single layer insulin aggregates based on the high resolution models of insulin fibril with the purpose of understanding the nature of

insulin self-assembly. We present the information on energetics of the insulin association at an atomistic level that could be used to design new short- and long- acting insulin analogues. Mutant forms of insulin with altered aggregation properties that could potentially be used to in slow- or fast- acting therapeutic formulations are suggested on the basis of the observed contacts at the aggregates interface. There has been no pervious systemic study on how mutation affects the stability of the insulin oligomer aggregates. Our MD simulations of the different size of the insulin oligomer may contribute to a better understanding of the nucleation process and conformation change at the very early stage of fibril formation. This study aims to answer the following questions:

1. Which regions of the wild-type insulin oligiomer aggregate are flexible?
2. How do the single point mutations influence the structure and flexibility of these regions?
3. What are the effects of single glycine mutations of the side chains involved in the steric zipper?
4. What are the conformational differences among the aggregates of various sizes?

5.2 Methods

We conducted a total of ~0.35μs of explicit-solvent molecular dynamics (MD) simulation on the insulin single layer insulin aggregate oligomer and mutated sequences with intact disulfide bridges, using temperatures 330 K to emulate the experimental conditions of *in vitro* insulin fibrillization.[66,67]

In this study we rely on insulin fibrillar model constructed by Ivanova *et al* [94] using crystal structure of the LVEALYLV, SLYQLENY and fiber diffraction patterns. The C-terminal region of chain B (residues 20–30) is not involved in amyloid fibrillization [183], and was omitted.

Comparison of amino acid sequence of the insulin sequences from five different mammalian species (porcine, bovine, sheep, mouse and rate) in the residue 20-30 shows nine of the amino acids residues are conserved and B30 Tyr in human is replaced with Ala in the other species.[176] These residues are missing in the insulin model used in this study. Therefore, only the 40 amino acids are taken into account in the fibrils model.[94] The starting coordinates (Figure 5-1) for the MD simulations was taken from the amyloid fibril home page of David Eisenberg group available at: (http://people.mbi.ucla.edu/sawaya/jmol/fibrilmodels).

Ten different single point glycine mutant simulations were conducted to examine the effects of steric zipper. In chain A the following three single point glycine mutations were done: a) tyrosine (Y) in position 14 replaced with glycine (G), b) leucine (L) in position 16 replaced with glycine (G), c) asparagines (N) in position 18 replaced with glycine (G). While in chain B a total of seven mutations were done: d) leucine (L) in position 11 replaced with glycine (G), e) valine (V) in position 12 replaced with glycine (G), f) glutamic acid (E) on position 13, replaced with glycine(G), g) alanine (A) in position 14 replaced with glycine (G), h) leucine (L) in position 15 replaced with glycine (G), i) tyrosine (Y) in position 16 replaced with glycine (G), j) leucine (L) in position 16 replaced with glycine (G). The three mutants in chain A will be termed: $Y14G_A$, $L16G_A$ and $N18G_A$. The other seven mutants in chain B will be termed: $L11G_B$, $V12G_B$, $E13G_B$, $A14G_B$, $L15G_B$, $Y16G_B$ and $L17G_B$, respectively. All the starting structures of the mutants were built from the wild type [93] by replacing the side chains of the targeted residues with glycine, using VMD.[155] Such analogues may possibly allow increasing the potency of insulin-based medicines, extending the time of action and controlling it using prodrugs, as well as enhancing the bioavailability. Insulin analogues were developed to try and achieve more physiological insulin replacement from injection in the subcutaneous site.

The insulin single layer oligomer aggregates studied contains multiple protein-protein interfaces; the calculation of the free energy of the associations of monomers in single layer oligomer aggregates requires a suitable interface. Because the present study aimed to assess the stability of the insulin oligomer with respect to the increase in the number of strand (the longitudinal growth) and the effect of mutation of amino acids involved in intra-chain, we measured the interaction energy between the edge (B) chains and the central double layer (A) as shown in Figure 5-2. A molecular mechanics–Generalized Born surface area (MM-GBSA) method was used to calculate the binding free energies in the insulin single layer complex. The free energy analysis was done using a single trajectory approach.

5.3 Results

5.3.1 Relative structure stability of insulin oligomers

The conformational change and the conservation of the oligomers were monitored by the time evolution of the backbone root mean square (RMSD) and root mean square fluctuation (RMSF). The RMSDs provide useful information on relative stability of the oligomers, and were previously used in stability analyses of amyloid oligomers with β-sheet structure.[21,78,86,157,158] In Figure 5-3 we plot RMSD of the wild type and mutants oligomer aggregates relative to the corresponding initial structure as a function of simulation time.

The conformation change and conservation of oligomers stability of the different size wild type insulin oligomer was monitored by the time evaluation of the RMSD. In Figure 5-3, we plot the RMSD of the insulin oligomers of the main-chain heavy atoms relative to the corresponding initial structure as a function of simulation time. The profiles of RMSD deviation appear to reach reasonable plateaus during the 10 ns production run, indicating that statistical

convergence has been attained in these simulations. The average main-chain RMSD between the MD simulation and the initial structure were found in the range of 4.3 to 4.9 Å for WT and 3.75 to 4.75 Å for the mutants. Along the trajectories the systems tend to keep the original conformation. In the case of the wildtype monomer and dimer a large conformational flexibility is observed as indicated by the RMSD, RMSF (Figure 5-3 and 5-4), average secondary structure content (Table 5-1) and cluster analysis (Figure 5-6). The RMSF and cluster analysis presented in Figure 5-3 and Figure 5-6 for the monomer indicates it undergoes larger conformational changes forming a globular structure instead of the initial solenoid form. The C-terminus of the monomer bends to the central region, and forms an anti-parallel β-sheet between residues 12-16 and residues 35-40. The dimer largely preserves the solenoid conformation, but exhibits larger per residue fluctuation values in the beta sheet region in chains A ($_{11}$SLYQLENY$_{19}$) and B ($_{12}$VEALYL$_{17}$), that are twice as large as the RMSF in other cases (Figure 5-4).

The residue-based root mean square fluctuation (RMSF) of the backbones was used to assess the local dynamics and flexibility of each residue using ptraj tool in AMBER11. Detailed analysis of RMSF of the C_α, C, N atoms versus the residue number for insulin wild type and mutants oligomer aggregates is shown in Figure 5-4. The large size oligomer such as SH1-ST8 and SH1-ST10 are more flexible at their N- and C-termini as compared to its smaller size oligomers (except SH1-ST2). The relatively larger RMSF per residue of the SH1-ST1 and SH1-ST2 is found the β-sheet region indicates their instability and the loss of the initial fibril conformation (Figure 5-4A). The other oligomers (SH1-ST4 to SH1-ST10) the β-sheet region exhibits significantly smaller structural fluctuation from the fibril conformation. Figure 5-4B and C shows the RMSF values of atomic positions by each residue, computed throughout the simulation for wildtype of insulin (SH1-ST10) and corresponding single point glycine mutants.

The RMSF of single point mutants were found to be larger than the wildtype. The smallest fluctuation of the average RMSF in for chain A and chain B was found in the segments LYQLENY and LVEALYL respectively. The RMSF results for the wildtype and the mutants indicates that all chains have common characteristics of small variation in the residue located within the β-sheet region whereas large variations for residues in the termini regions. The larger flexibility of the two termini residues were due to the reduction of hydrogen bonds between the peptides. The side chains of the termini residues are more exposed to the water and tend to form hydrogen bonds with water molecules.[88]

5.3.2 Secondary structure content

We carried out secondary structure analysis using DSSP tool in AMBER11. Table 5-1 reports the average number of residues in a given secondary structure as a function of simulation time and the corresponding initial structure. When the average secondary structure content over time is considered differences between smaller oligomers and larger size oligomer are evident from the simulations. The single and double stranded aggregates exhibits lower β-sheets and more residues in helices and in coil-like conformation. The larger oligomers (such as SH1-ST4 to SH1-ST10) exhibit exhibits more β-sheets contents and fewer residues in helices and in coil-like conformation. The larger aggregates retain the fibril conformation mainly due to an increased number of backbone hydrogen bonds.[88]

5.3.3 Cluster Analysis

Cluster analysis or in short clustering puts similar samples of data into groups called clusters, such that an ensemble of data, for example, the different structures obtained from a MD

trajectory, is partitioned into groups of similar objects. Structural clustering is useful to understand the molecular motion within conformational space.[192] Conventional clustering algorithms are reducing any large MD trajectory to a set of conformational basins. To identify the most populated conformations sampled, clustering of all snapshots from the trajectories was performed using the ptraj program of AMBER11. The standard approach, which has been used with considerable success, is to cluster the configurations in terms of an RMSD. For clustering, we utilized the average linkage algorithm implemented in ptraj.[77] The uniqueness or equivalence of different clusters was assessed based on visual comparison of representative structures. The clustering was performed on a 5000-frame reference set (4 ps sampling rate). Figure 5-6 shows the superposition of the initial structure and the most populated cluster structure for single-layer insulin aggregates of different sizes. The analysis of the structures indicated the most populated clusters detected from the smaller size oligomers (single and double strand) indicating larger structural rearrangements compared to the initial conformation taken from the fibril model. The conformation was preserved for larger aggregate (SH1-ST8 and SH1-ST10).

5.3.4 Free energy calculation

Detailed characterization of individual energy terms of the calculated binding free energy of the studied insulin oligomer aggregates are shown in Table 5-2. An inspection of the free energy components for the wild types and mutants reveals that the electrostatic component of the free energy of binding (ΔE_{ele}) contributes unfavorably to binding ($\Delta G > 0$). The nonpolar component contributes favorably ($\Delta G < 0$) as expected, since formation of complexes reduces solvent-accessible surface area. In most cases the electrostatic component of the solvation free energy ΔG_{GS} is consistently favorable. The interaction energy due to electrostatic interaction

(ΔE_{vdw}) between strands led to unfavorable binding. These observations are consistent with previous calculations of the electrostatic component of the free energy of solvation. However, the less favorable electrostatics in each case is compensated by highly favorable nonpolar component of the free energy. In each case, favorable nature of the nonpolar interaction mostly originates from nonpolar component of solvation (ΔG_{GB}) and the van der Waals interaction energy (ΔE_{vdw}).

The result of the binding free energy calculation (Table 5-2) indicates the structurally stable models have the lowest binding free energy, while the models which are structurally unstable were found to have the largest binding free energy. The difference in binding free energy between the wild-type and mutated complex is defined as:

$$\Delta\Delta G_{mut} = \Delta G_{mut} - \Delta G_{wild} \quad (1)$$

A positive and negative $\Delta\Delta G_{mut}$ values indicate the unfavorable and favorable contributions. The positive $\Delta\Delta G_{mut}$ values of 37.3 to 1.4 kcal/mol of the mutants in the β-sheet region (except $Y14G_A$ and $L15G_B$) indicate their lower tendency to aggregation compared to the wildtype. This result from our simulation could be used in rational design new insulin analogues with decreased propensity for self-association avoiding injection amyloidosis of insulin. The relatively larger positive value of $\Delta\Delta G_{mut}$ for mutants ($Y14G_A$, $L11G_B$, $V12G_B$ and $E12G_B$) indicates the less favorable association compared to the wildtype. In general substitution of β-sheet region of chain A and B by a small, short Gly disrupts the steric zipper shape complementary and weakens hydrophobic interactions (see Table 5-2).The single point glycine mutation reduces the unfavorable electrostatic interaction. The mutation of the negatively charged glutamate (E) to G in the mutant $E12G_B$ reduces the electrostatic repulsion in the wildtype as is evidence in the significant reduction of the unfavorable electrostatic interaction (Table 5-2). Mutation of Tyr14

in chain A with glycine eliminates the hydrogen bond between the residues Tyr14 of chain and the Tyr16 of chain B as the result the calculated binding free energy was high. The negative value of value of $\Delta\Delta G_{mut}$ for a mutant is due to the increased hydrophobic interaction in the steric zipper between the chains. Complete hydrophobic side chain by Gly substation impedes fibril growth.[193] The trend in the calculated binding free energy is in agreement with the observed instability based on RMSD, RMSF. Those aggregate oligomer models which show structural instability were found to have unfavorable binding energy compared to the stable once.

5.3.5 Decomposition free energy on a per-residue basis

The free energy decomposition not only identifies the binding energy hot spots, but also gives insight into the nature of the key interactions.[189] To provide the basic information on the intermolecular interactions contributed from the individual residues in the insulin single layer aggregate interaction decomposition of free energy (the per residue total, side chain and backbone binding free energy) was evaluated using the decomposition energy module in AMBER11. The calculation was performed over the 2500 MD snapshots taken from the 20 ns simulation. The summations of per residue interaction free energies were separated into the residue backbone (ΔG backbone bind) and the side chain (ΔG side chain bind). The energy contributions from the selected residues are summarized in Figure 5-5.

The result form the energy decomposition shows the major contribution to the binding energy of insulin oligomer aggregate is gained from the key amino acid residues (those with a $G_{binding} \leq$-0.50 kcal/mol) occurring mainly in the β-sheet region. These residues are in chain A (Q5, L13, Y14, Q15, L16, N18 and Y19) and in chain B (S9, L11, V12, L15, L17 and V18).The result of the per residue decomposition indicates the important of the particular residues in the β-sheet

region with regards to the formation and stabilization of insulin and this is in agreement with experimental observation.[94] To establish the interaction associated with these residues, their electrostatic, van der Waals and solvation energy terms are shown in Table 5-3 . Table 5-3 shows that favorable contributions to the binding free energy arising from these residues relates to E_{ele}, E_{vdw}, ΔG_{GS}, while unfavorable comes from ΔG_{GB}. The favorable E_{ele} terms from the residue in the β-sheet region are compensated by highly unfavorable repulsion from three glutamate between adjacent insulin layers.

5.3.6 Fibril nucleation and the structure of insulin oligomers

Understanding the process of amyloid fibril formation is an important goal of protein aggregation studies.[47] Amyloids grow in a nucleation-dependent manner.[35,47] Fibrillation kinetics is typically characterized by an initial apparent lag phase related to the formation of oligomers, protofibrils, and aggregation nuclei.[194] The typical fibril formation process is characterized by a lag phase in which no detectable fibers are formed. This is then followed by an elongation phase in which fiber is formed over a time period often shorter than the lag phase. Eventually, the process reaches equilibrium when most soluble proteins are converted into fibrils.[195] On the other hand, if fibers (oligomers, protofibrils, and fibrils) are already formed they grow extremely fast with very short lag-times. A recent experimental work on oligomers capability to stabilize fibril nucleation activity by Ono *et al*.[196] on Aβ amyloid, on amylin [195], and on insulin in various labs [197, 198, 199,200] have indicated the oligomers and the fibril showed different capability to act as seeds. Anselm *et al* [184] used the degree of structural similarity to the fibril conformation detected for the oligomers in their simulation as an possible reason for difference among various size of oligomer with respect to the effective as nucleation seeds. The degree of structural similarity

between fibril conformation and the conformation of the oligomers after MD simulation could be used in explaining the shorting of the lag phase in the present of different size of oligomers. The trend in retaining the initial fibril conformation will help us in getting an atomic level explanation to the observed difference in the seed effects of various sizes of oligomers. The results from our simulation show the single layer insulin oligomer as small as the trimer is capable of preserving the conformation present in the fibril (See Figure 5-6). The dimer shares some properties of the mature fibril. The monomer adopts a structure that differs significantly from that of the fibril.

The secondary structure content and clustering analysis on the trajectories from the various size single layer insulin oligomer shows the larger aggregates retain the fibril conformation and the smaller ones (SH1-ST1 and SH1-ST2) lose this conformation. The observation could be used to explain the shortening of the nucleation lag phase of insulin aggregation with oligomer seeds. Insulin like other amyloid peptides follow an apparent nucleation-dependent polymerization kinetics[47,201] whereby a small number of monomers associate through a free energy barrier corresponding to a critical nucleus size, beyond which initiates gradient of favorable free energy or "down-hill" polymerization. Based on the result of the secondary structure and cluster analysis we proposed the SH1-ST4 to be a critical nucleus for the single insulin fibril oligomer growth. To characterize the critical nucleation we computed the association energy different between our proposed minimum nuclei and the larger size oligomers (SH1-ST6, SH1-ST8 and SH1-ST10) we using the equation.

$$\Delta\Delta G(n) = G(n) - (G4); \quad n=4,6,8,10 \qquad (2)$$

The results are shown in Table 5-2. Our calculation shows that, for a high number of strands, the oligomer is stable and its free energy is favorable for the addition of the new chains. The result

of our semi-quantative approaches for insulin single layers of limited size of insulin single layer are in agreement with those obtained from pervious extensive simulations done on Aβ amyloid critical nucleus and mechanism of fibril elongation.[202,203]

5.4 Conclusions

The results from this work provided valuable insight into the forces that drive the stability of the peptide-peptide complexes of the single layer aggregate oligomer models of insulin and those that lead to unstable complexes. The study of the wild type and mutants in an explicit solvent will provide valuable to future efforts aimed at the design of short –and long-acting insulin analogs. The major findings of this study can be summarized as follows:

1. The stability of the insulin single layer peptides oligomers increases as the number the number of strands increases (dynamic cooperative effect) .

2. The binding energy calculated by MM-GBSA method shows the hydrophobic interactions play an important role in stabilizing the structural organizations of the single layer insulin. Per residue decomposition shows the key amino acid residues (those with a $G_{binding} \leq$-1.00 kcal/mol) occur mainly in the β-sheet region of both chain A and chain B. Due to the electrostatic repulsion between the three negatively charged glutamates in adjacent insulin strands, electrostatic repulsion to the binding energy is unfavorable.

3. A single glycine substitution at the steric zipper interface disrupts the hydrophobic contacts and reduces the van der Waals interactions in the mutants thus reducing the binding free energy. The result of the binding free energy calculation indicated that the wild type is more structurally stable than most of the mutants. A comparison of the binding free energy between the wildtype and the chain A mutants ($Y14G_A$, $L16G_A$ and $N18G_A$) indicates shape

complementarity between neighboring strands plays a key role to stabilize the entire oligomeric structure.

4. The secondary structure contents and clustering analysis of the trajectories of the single layer insulin oligomers shows the larger aggregates retain the fibril conformation but the smaller once (SH1-ST1 and SH1-ST2) lose this conformation. This observation could explain the observed shortening of the nucleation lag phase of insulin aggregation with oligomer seeds. Based on the result of the secondary structure and cluster analysis we proposed the SH1-ST4 is a critical nucleus for the single-layer insulin fibril oligomer growth.

Our simulations provide detailed insights into understanding the structural stability and aggregation behavior of the insulin single layer aggregates (obtained from a high resolution insulin fibril model) at atomic level. In the search for clinically advantageous fast acting insulin analogs several approaches were found to be useful for altering the monomer/monomer interface. One of them is disruption of β-sheet interactions in the β-chain through charge repulsion, or changes in hydrophobic interactions in the C-terminus of chain B [176]. Our simulations on the wildtype and single glycine mutants at the steric zipper region show other parts of insulin molecule can be targeted in the design of both short and long action insulin analogs as well. Aside from the design of such insulin analogs, the present study may prove useful for rational design of insulin aggregation inhibitors which could be used in stabilizing insulin formulations, leading to their safer handling and more cost-effective storage especially in developing countries.

Table 5-1 Average secondary structure contents of different size insulin wildtype and its (SH1-ST10) single point glycine mutants

	Starting	Average	Starting	Average	Starting	Average	Starting	Average	Starting	Average
	WT (SH1-ST2)		WT (SH1-ST4)		WT (SH1-ST6)		WT (SH1-ST8)		WT (SH1-ST10)	
β-Sheet	0(0)	0(0)	0(0)	0(0),[0]	0(0)	0(0),[0]	0(0)	0(0),[0]	0(0)	0(0),[0]
β-Bridge	23(63.9)	15.3(56.1),[3.6]	47(63.5)	42.3(62.3),[3.1]	83(68)	69.9(79.6), [5.2]	115(68.1)	114(69.2), [5.7]	158(73.8)	135.6,(70.1),[7.6]
coil	0(0)	0(0)	0(0)	0(0),[0]	0(0)	0(0),[0]	0(0)	0(0),[0]	0(0)	0(0),[0]
turn	2(5.5)	6.4(23.5),[2.4]	9(12.2)	10.3(15.2),[3.4]	25(20.5)	16.4,(18.7).[4]	35(20.7)	28.1,(17.1).[6.1]	45(21)	28.1,(17.1).[6.1]
α-Helix	11(30.6)	3.9(14.3), [2.9]	14(18.9)	12.1(17.8),[4.6]	8(6.6)	14.4,(8.7),[5.9]	12(7.1)	14.4,(8.7),[5.9]	8(3.7)	11.2(5.8),[4.3]
3_{10}-Helix	0(0)	1.6(5.9),[2.0]	2(2.7)	2.7(4.0). [2.8]	6(4.9)	8.4,(9.6), [3.9]	7(4.1)	5.7(3.5), [4.1]	3(1.4)	11.8,(6.1),[4.6]
π-Helix	0(0)	0.05(0.2),[0.1]	2(2.7)	0.5(0.7), [1.2]	0(0)	1.3,(1.5), [1.9]	0(0)	2.4(1.5), [2.7]	0(0)	0.5(0.3),[1.0]
	$L11G_B$ (SH1-ST10)		$V12G_B$(SH1-ST10)		$E13G_B$ (SH1-ST10)		$A14G_B$ (SH1-ST10)		$L15G_B$ (SH1-ST10)	
β-Sheet	0(0)	0(0)	0(0)	0(0),[0]	0(0)	0(0),[0]	0(0)	0(0),[0]	0(0)	0(0),[0]
β-Bridge	135(71)	127.7,(68.3),[7.9]	146(72.3)	128,(68.8),[6.4]	131(68.6)	130.5(69.6), [7.9]	134(69.1)	120.9(69.2), [7.7]	157(70.7)	142.5,(70.2),[7.6]
coil	0(0)	0(0)	0(0)	0(0),[0]	0(0)	0(0),[0]	0(0)	0(0),[0]	0(0)	0(0),[0]
turn	28(14.7)	31.5(16.8),[4.9]	36(17.8)	37.6,(20.2),[5.0]	45(23.6)	34.9,(18.6).[5.7]	46(23.7)	38.8,(21.6).[5.2]	40(18)	32.6,(16.1).[6.7]
α-Helix	16(8.4)	16.5(8.8), [4.9]	9(4.5)	11.5,(6.2),[4.6]	4(2.1)	12.6,(6.7),[5.3]	8(4.1)	7.0,(3.9),[3.6]	18(8.1)	17.5,(8.6),[6.2]
3_{10}-Helix	9(4.7)	8.8,(4.7),[4.7]	6(3.0)	8.4, (4.5). [4.3]	9(4.7)	9.5,(5.1), [4.2]	6(3.1)	9.8,(5.4), [4.1]	7(3.2)	8.2,(4.0),[4.5]
π-Helix	2(1)	2.5,(1.4),[2.2]	5(2.5)	0.5(0.3), [1.3]	4(2.1)	0.3,(0.2), [1.0]	0(0)	1.4,(0.8), [2.4]	0(0)	2.3,(1.1),[3.0]
	Y16G (SH1-ST10)		$L17G_A$ (SH1-ST10)		$Y14G_A$ (SH1-ST10)		$L16G_A$ (SH1-ST10)		$N18G_A$ (SH1-ST10)	
β-Sheet	0(0)	0(0)	0(0)	0(0),[0]	0(0)	0(0),[0]	0(0)	0(0)	0(0)	0(0),[0]
β-Bridge	133(72.3)	127.7,(68.3),[6.0]	146(72.6)	135.1,(71.5),[6.7]	143(69.1)	121.1,(66.0),[7.9]	139(67.8)	127.8,(65.8),[6.8]	151(70.9)	122.9,(68.7),[8.2]
coil	0(0)	0(0)	0(0)	0(0),[0]	0(0)	0(0),[0]	0(0)	0(0)	0(0)	0(0),[0]
turn	39(21.2)	31.5(16.8),[4.9]	23(11.4)	30.6,(16.2),[5.4]	29(14)	30.0,(16.4).[5.9]	38(18.5)	39.0,(20.1),[5.8]	36(16.9)	36.4,(20.4),[4.4]
α-Helix	10(5.4)	16.5(8.8), [4.9]	24(12)	17.9,(9.5),[5.2]	24(11.6)	21.0(11.4),[7.6]	13(6.3)	16.3,(8.4), [5.1]	10(4.7)	7.3,(4.1),[5.6]
3_{10}-Helix	0(0)	8.8,(4.7),[4.7]	3(1.5)	3.1, (1.6). [2.8]	6(2.9)	10.2,(5.6),[5.6]	3(1.5)	9.8,(5.0),[4.5]	9(4.2)	9.0, (5.0). [4.1]
π-Helix	2(1.1)	2.5,(1.4),[2.2]	5(2.5)	2.5(1.2), [2.2]	5(2.4)	1.2,(0.6),[2.0]	12(5.8)	1.2,(0.6),[2.0]	7(3.3)	3.2, (3.2), [2.4]

*Percentages and standard deviations are given in parenthesis and square brackets, respectively

Table 5-2 Individual energy components of the binding free energy of insulin amyloid aggregate peptide

MM-GBSA binding energy components of the different size of the single layer insulin amyloid aggregates							
System	ΔE_{vdw}	ΔE_{ele}	ΔG_{GB}	ΔG_{GS}	ΔG_{solv}	ΔG_{total}	$\Delta\Delta G_{(6\text{-}n)}$
WT(*SH1-ST4*)	-163.96±0.18	576.65±1.02	-502.40±0.95	-21.35±0.02	-523.75±0.95	-111.06±0.17	11.43
WT(*SH1-ST6*)	-177.21±0.28	1149.66±0.91	-1071.50±0.86	-23.43±0.03	-1094.94±0.86	-122.49±0.24	0.0
WT(*SH1-ST8*)	-149.64±0.19	1514.62±2.27	-1433.42±2.22	-25.81±0.02	-1459.27±2.22	-142.29±0.22	-19.8
WT(*SH1-ST10*)	-196.31±0.14	1827.75±	-1742.52±1.00	-25.20±0.02	-1767.28±0.13	-136.28±0.13	-13.79
MM-GBSA binding energy components of single layer insulin amyloid aggregates mutant of chain A (SH1-ST10)							
System	ΔE_{vdw}	ΔE_{ele}	ΔG_{GB}	ΔG_{GS}	ΔG_{solv}	ΔG_{total}	$\Delta\Delta G_{mut}$
$Y14G_A$ (*chain A*)	-165.16±0.12	1974.31±0.84	-1887.76±0.84	-20.36±0.02	-1908.11±0.84	-98.96±016	37.3
$L16G_A$ (*chain A)*	-208.14±0.14	1710.42±1.74	-1628.03±1.73	-27.10±0.03	-1655.15±1.73	-152.87±0.16	-16.59
$N18G_A$ (*chain A*)	-201.272±0.12	1623.99±0.93	-1527.81±0.91	-26.02±0.01	-1553.82±0.91	-131.11±0.15	5.17
MM-GBSA binding energy components of single layer insulin amyloid aggregates mutant of chain B (SH1-ST10)							
System	ΔE_{vdw}	ΔE_{ele}	ΔG_{GB}	ΔG_{GS}	ΔG_{solv}	ΔG_{total}	$\Delta\Delta G_{mut}$
$L11G_B$ (*chain B*)	-171.7±0.23	1598.8±3.4	-1522.2±3.4	-22.0±0.02	-1544.2±3.4	-117.1±0.2	19.2
$V12G_B$ (*chain B*)	-167.1±0.2	1789.9±1.1	-1715.79±1.0	-21.5±0.01	-1737.2±1.0	-114.4±0.2	21.9
$E13G_B$ (*chain B*)	-186.7±0.3	981.4±1.0	-893. 7±1.0	-25.3±0.02	-918.9±1.06	-124.2±0.25	12.1
$A14G_B$ (*chain B*)	-191.2±0.2	1620.8±4.0	-1533.3±4.0	-25.4±0.03	-1558.7±4.0	-129.2±0.25	7.1
$L15G_B$ (*chain B*)	-209.3±0.2	1622.9±2.4	-1529±2.4	-26.6±0.02	-1555.6±2.4	-142.03±	-5.7
$Y16G_B$ (*chain B*)	-194.8±0.2	1752.6±1.9	-1663.1±1.9	-23.3±0.03	-1686.4±1.9	-128.622±0.2	7.6
$L17G_B$ (*chain B*)	-204.4±0.3	1823.9±2.3	-1728.42±2.2	-26.0±0.02	-1754.43±2.3	-134.904±0.2	1.4

$\mathbf{E_{vdw}}$ and $\mathbf{E_{elec}}$ are the van deer waals and electrostatic binding terms, $\Delta\mathbf{G_{GB}}$, $\Delta\mathbf{G_{GS}}$, $\Delta\mathbf{G_{solv}}$ are the polar, non polar and total solvation energies. Data are shown as mean±SD. Standard error of the mean $\Delta\mathbf{G_{total}} = \Delta\mathbf{E_{vdw}} + \Delta\mathbf{E_{ele}} + \Delta\mathbf{G_{sol}}$, $\Delta G_{sol} = \Delta G_{GB} + \Delta G_{GS,}$ $\Delta G_{mut} = \Delta G_{mut} - G_{wild}$, the change of mutant binding free energy as to wildtype. $\Delta\Delta\mathbf{G_{(n\text{-}6)}} =$ Oligomer free energies expressed relative to the hexamer state for β -sheet oligomers.

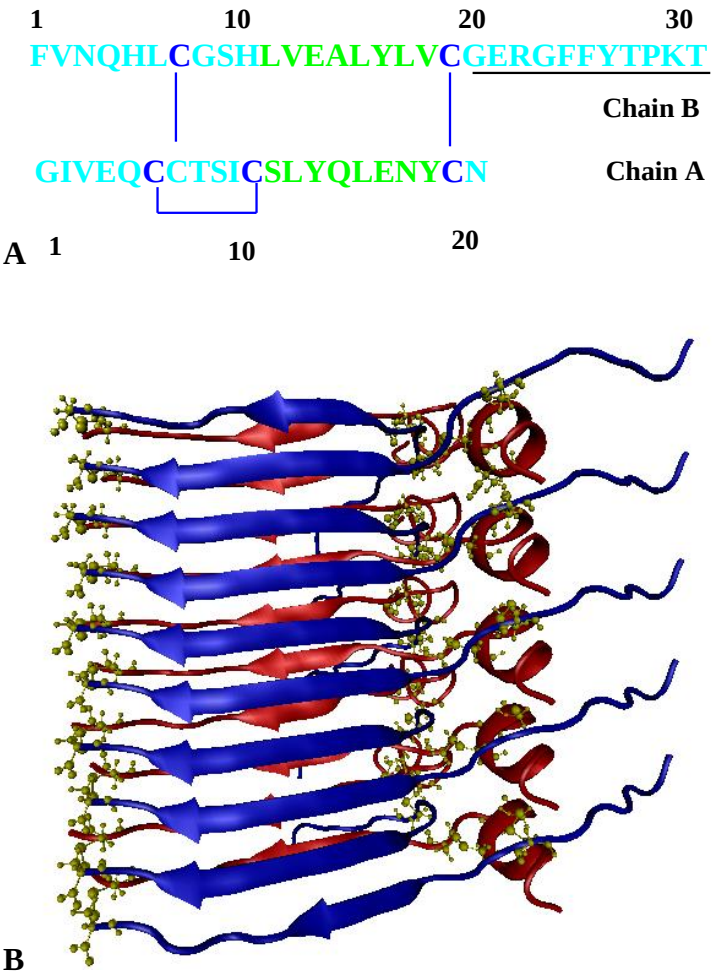


Figure 5-1 Amino acid sequence and structure of single-layer insulin oligomer (A) Amino acid sequence of insulin (chain B top and chain A bottom). Segments LVEAYLV of chain B and SLYQLENT of chain A are colored in green. Disulfide bonds are colored in blue. The C-terminal region of chain B underlined and in italic, is not involved in amyloid fibrillization. The residues underlined are missing in the insulin model used in this study. Therefore, only the 40 amino acids are taken into account in the fibrils model. (B) Single -layered structural models of insulin oligomers (10 stranded). Two chains are associated together via interdigitated pair of LYQLENY molecules of chain A and LVEALYL molecules of chain B which interlock tightly to form the dry steric zipper. The chain A is red and the chain B is blue. Disulfide bonds are indicated in the yellow.

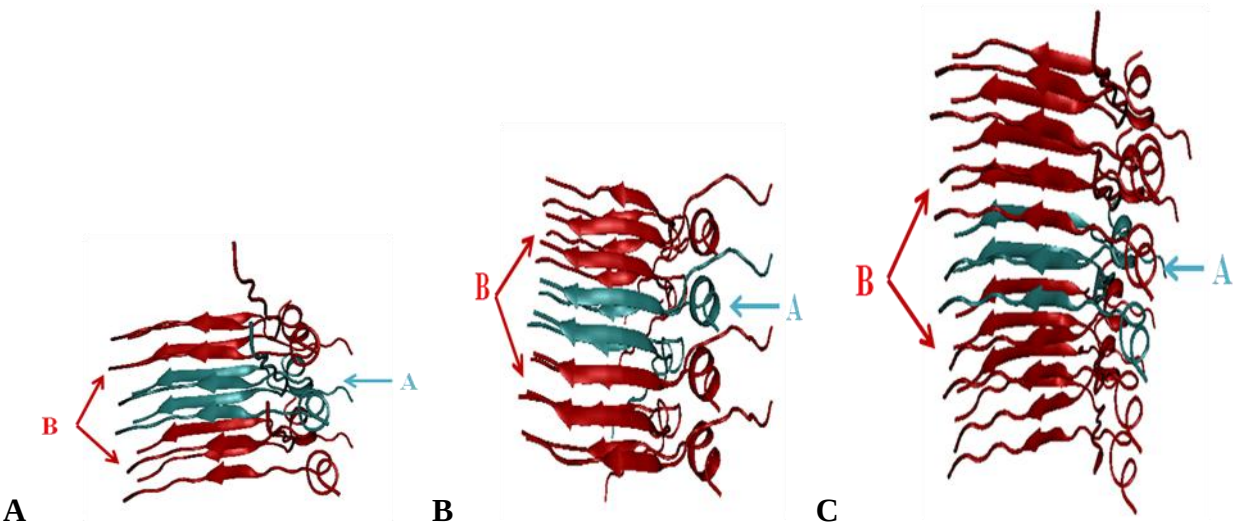
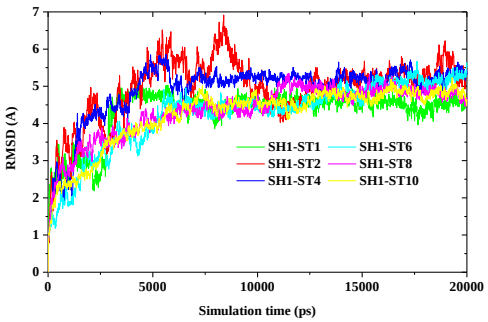
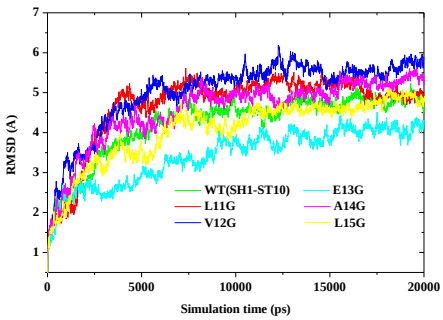


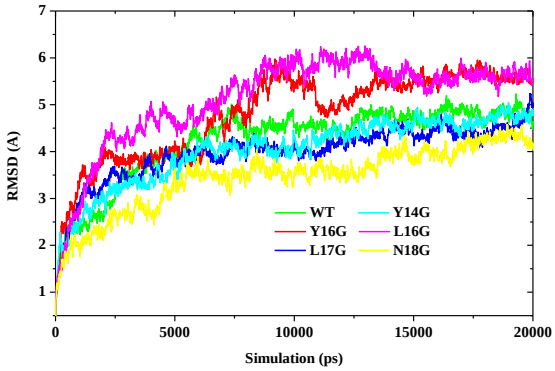
Figure 5-2 Schematic drawing of the setup used for estimating the internal stability of insulin single layer aggregates and mutants. Free energies of interaction were calculated between the middle chains A (cyan) and the remaining edge chains B(red) reflecting the strength by which chain A clamps the insulin stack in the β-solenoid structure.



A

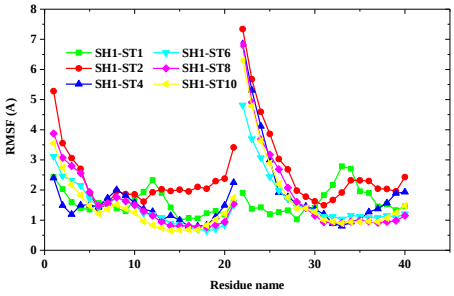


B

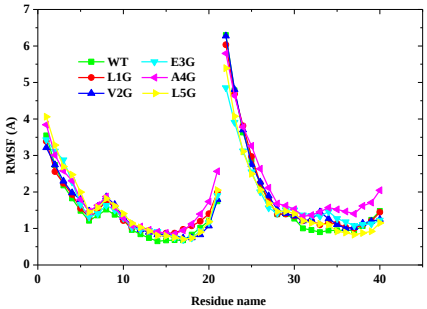


C

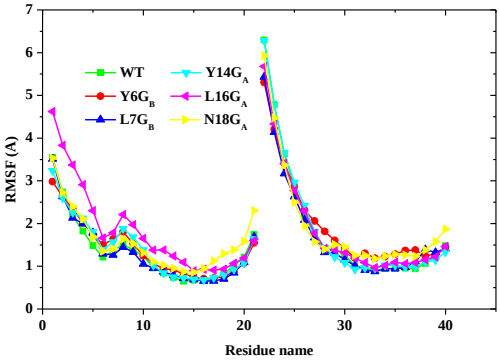
Figure 5-3 Backbone RMSDs of single layered insulin models and single glycine mutants (SH1-ST10)



A

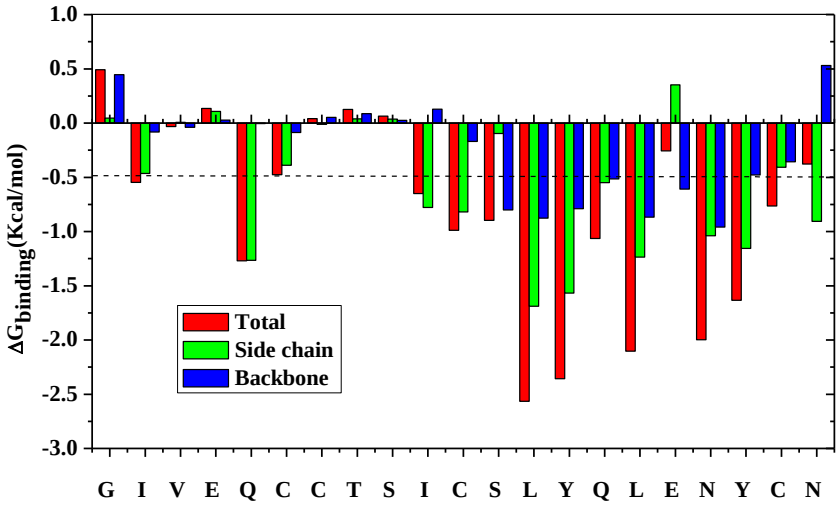


B

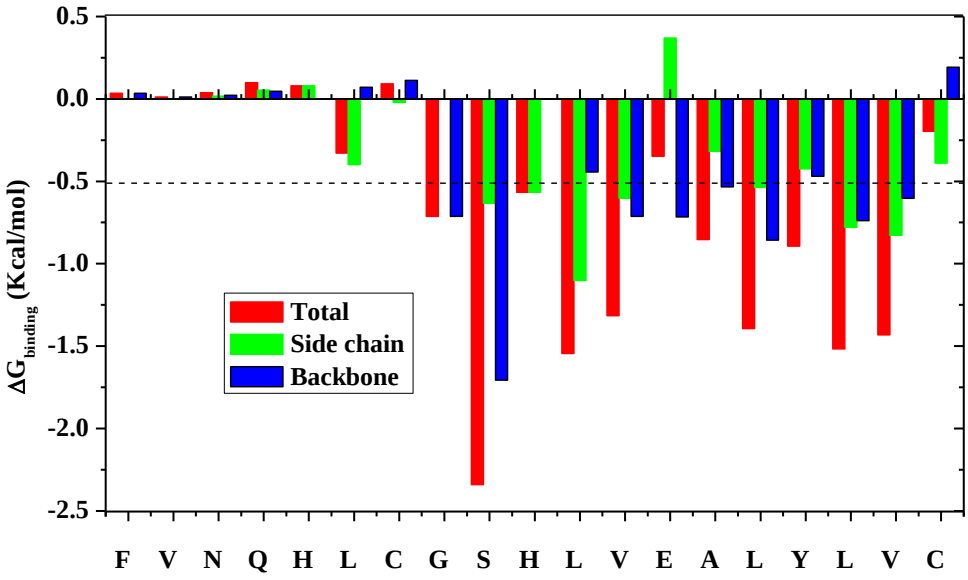


C

Figure 5-4 Plot of the root mean square fluctuation (RMSF) of single layered insulin models and single point glycine mutants (SH1-ST10)



A



B

Figure 5-5 Decomposition of the free energy on a per residue basis for chain A (A) and chain B (B) of the single layer insulin aggregate (SH1-ST10)

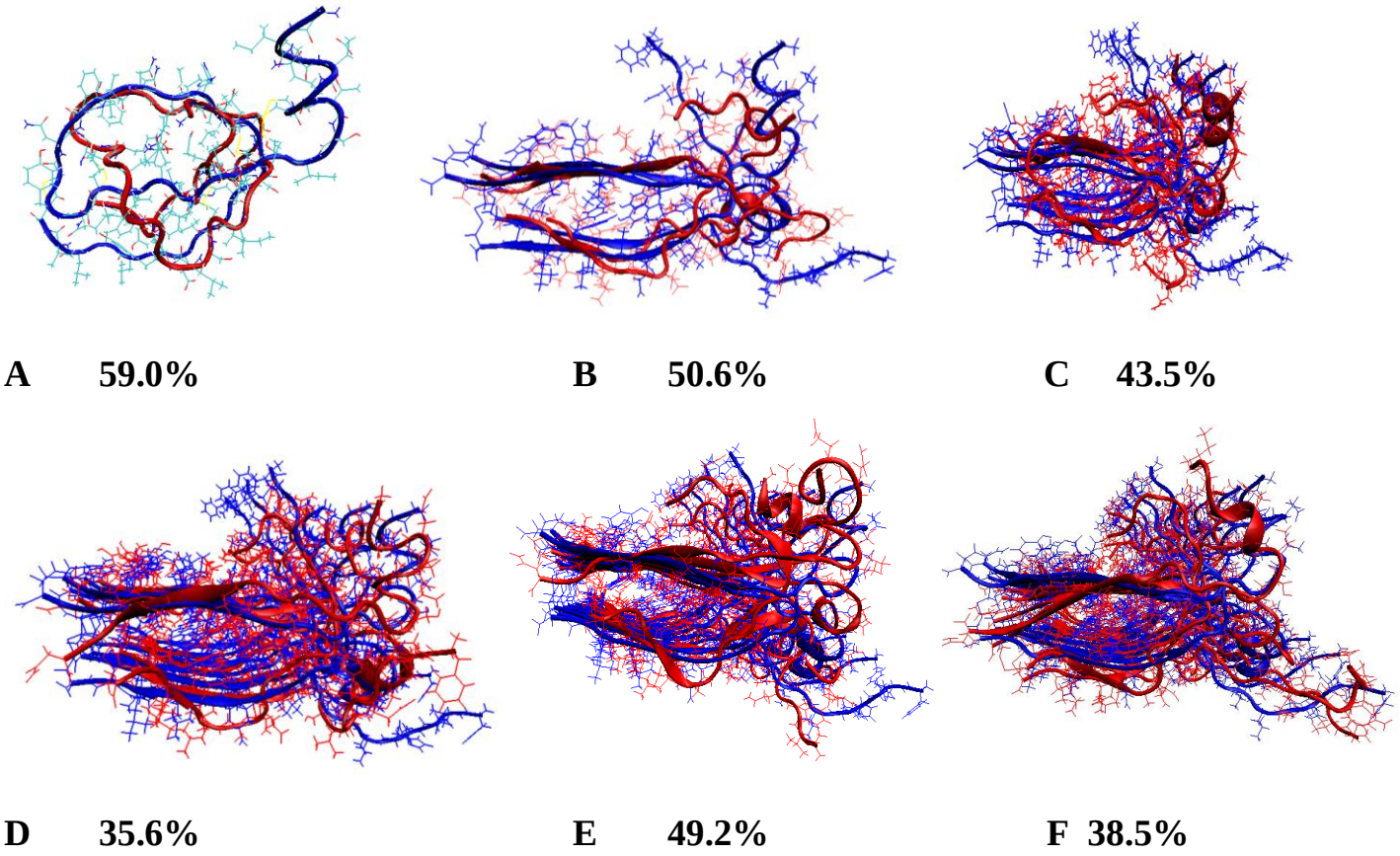


Figure 5-6 Superposition of insulin single layer oligomer aggregates of the initial structure and representative structures of the most populated clusters.(A: SH1-ST1, B: SH1-ST2, C: SH1-ST4, D: SH1-ST6, E: SH1-ST8, and F: SH1-ST10). The initial structures are shown in blue and the most populated cluster with the corresponding percentages is shown in red.

5.5 References

165. Nystrom, F. H.; Quon, M. J., Insulin signalling: Metabolic pathways and mechanisms for specificity. *Cellular Signalling* **1999,** 11, (8), 563-574.

166. Ottensmeyer, F. P.; Beniac, D. R.; Luo, R. Z. T.; Yip, C. C., Mechanism of transmembrane signaling: Insulin binding and the insulin receptor. *Biochemistry* **2000,** 39, (40), 12103-12112.

167. Wild, S.; Roglic, G.; Green, A.; Sicree, R.; King, H., Global prevalence of diabetes - Estimates for the year 2000 and projections for 2030. *Diabetes Care* **2004,** 27, (5), 1047-1053.

168. Rewers, M., Why do people with diabetes die too soon? *Diabetes Care* **2008,** 31, (4), 830-832.

169. Storkel, S.; Schneider, H. M.; Muntefering, H.; Kashiwagi, S., Iatrogenic, insulin-dependent, local amyloidosis. *Laboratory Investigation* **1983,** 48, (1), 108-111.

170. 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, (3), 158-161.

171. Sluzky, V.; Klibanov, A. M.; Langer, R., Mechanism of insulin aggregation and stabilization in agitated aqueous-solutions. *Biotechnology and Bioengineering* **1992,** 40, (8), 895-903.

172. Grillo, A. O.; Edwards, K. L. T.; Kashi, R. S.; Shipley, K. M.; Hu, L.; Besman, M. J.; Middaugh, C. R., Conformational origin of the aggregation of recombinant human factor VIII. *Biochemistry* **2001,** 40, (2), 586-595.

173. Valla, V., Therapeutics of diabetes mellitus: focus on insulin analogues and insulin pumps. *Experimental Diabetes Research* **2010,** 2010, 1-14.

174. Valla, V., Therapeutics of Diabetes Mellitus: Focus on Insulin Analogues and Insulin Pumps. *Experimental Diabetes Research* **2010**, 14.

175. Bell, D. S. H., Insulin therapy in diabetes mellitus - How can the currently available injectable insulins be most prudently and efficaciously utilised? *Drugs* **2007,** 67, (13), 1813-1827.

176. Mayer, J. P.; Zhang, F.; DiMarchi, R. D., Insulin structure and function. *Biopolymers* **2007,** 88, (5), 687-713.

177. Geddes, A. J.; Parker, K. D.; Atkins, E. D. T.; Beighton, E., Cross-beta conformation in proteins. *Journal of Molecular Biology* **1968,** 32, (2), 343-358

178. 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, (28), 8397-8409.

179. Brange, J.; Dodson, G. G.; Edwards, D. J.; Holden, P. H.; Whittingham, J. L., A model of insulin fibrils derived from the x-ray crystal structure of a monomeric insulin (despentapeptide insulin). *Proteins-Structure Function and Genetics* **1997,** 27, (4), 507-516.

180. Ivanova, M. I.; Thompson, M. J.; Eisenberg, D., A systematic screen of beta(2)-microglobulin and insulin for amyloid-like segments. *Proceedings of the National Academy of Sciences of the United States of America* **2006,** 103, (11), 4079-4082.

181. Jimenez, J. L.; Nettleton, E. J.; Bouchard, M.; Robinson, C. V.; Dobson, C. M.; Saibil, H. R., The protofilament structure of insulin amyloid fibrils. *Proceedings of the National Academy of Sciences of the United States of America* **2002,** 99, (14), 9196-9201.

182. Vestergaard, B.; Groenning, M.; Roessle, M.; Kastrup, J. S.; van de Weert, M.; Flink, J. M.; Frokjaer, S.; Gajhede, M.; Svergun, D. I., A helical structural nucleus is the primary elongating unit of insulin amyloid fibrils. *Plos Biology* **2007,** 5, (5), 1089-1097.

183. Choi, J. H.; May, B. C. H.; Wille, H.; Cohen, F. E., Molecular Modeling of the Misfolded Insulin Subunit and Amyloid Fibril. *Biophysical Journal* **2009,** 97, (12), 3187-3195.

184. Horn, A. H. C.; Sticht, H., Amyloid-beta 42 Oligomer Structures from Fibrils: A Systematic Molecular Dynamics Study. *Journal of Physical Chemistry B* **2010,** 114, (6), 2219-2226.

185. Tsai, H. H.; Reches, M.; Tsai, C. J.; Gunasekaran, K.; Gazit, E.; Nussinov, R., Energy landscape of amyloidogenic peptide oligomerization by parallel-tempering molecular dynamics simulation: Significant role of Asn ladder. *Proceedings of the National Academy of Sciences of the United States of America* **2005,** 102, (23), 8174-8179.

186. Mark, A. E.; Berendsen, H. J. C.; Vangunsteren, W. F., Conformational flexibility of aqueous monomeric and dimeric insulin - a molecular-dynamics study. *Biochemistry* **1991,** 30, (45), 10866-10872.

187. Zoete, V.; Meuwly, M.; Karplus, M., Investigation of glucose binding sites on insulin. *Proteins-Structure Function and Bioinformatics* **2004,** 55, (3), 568-581.

188. Zoete, V.; Meuwly, M., Importance of individual side chains for the stability of a protein fold: Computational alanine scanning of the insulin monomer. *Journal of Computational Chemistry* **2006,** 27, (15), 1843-1857.

189. Zoete, V.; Meuwly, M.; Karplus, M., Study of the insulin dimerization: Binding free energy calculations and per-residue free energy decomposition. *Proteins-Structure Function and Bioinformatics* **2005,** 61, (1), 79-93.

190. Falconi, M.; Cambria, M. T.; Cambria, A.; Desideri, A., Structure and stability of the insulin dimer investigated by molecular dynamics simulation. *Journal of Biomolecular Structure & Dynamics* **2001,** 18, (5), 761-772.

191. Lu, B. Z.; Chen, W. Z.; Wang, C. X.; Xu, X. J., Protein molecular dynamics with electrostatic force entirely determined by a single Poisson-Boltzmann calculation. *Proteins-Structure Function and Genetics* **2002,** 48, (3), 497-504.

192. Keller, B.; Daura, X.; van Gunsteren, W. F., Comparing geometric and kinetic cluster algorithms for molecular simulation data. *Journal of Chemical Physics* 132, (7), 16.

193. Takeda, T.; Klimov, D. K., Probing Energetics of A beta Fibril Elongation by Molecular Dynamics Simulations. *Biophysical Journal* **2009,** 96, (11), 4428-4437.

194. Soto, C.; Estrada, L.; Castilla, J., Amyloids, prions and the inherent infectious nature of misfolded protein aggregates. *Trends in Biochemical Sciences* **2006,** 31, (3), 150-155.

195. Padrick, S. B.; Miranker, A. D., Islet amyloid: Phase partitioning and secondary nucleation are central to the mechanism of fibrillogenesis. *Biochemistry* **2002,** 41, (14), 4694-4703.

196. Ono, K.; Condron, M. M.; Teplow, D. B., Structure-neurotoxicity relationships of amyloid beta-protein oligomers. *Proceedings of the National Academy of Sciences of the United States of America* **2009,** 106, (35), 14745-14750.

197. Sorci, M.; Grassucci, R. A.; Hahn, I.; Frank, J.; Belfort, G., Time-dependent insulin oligomer reaction pathway prior to fibril formation: Cooling and seeding. *Proteins-Structure Function and Bioinformatics* **2009,** 77, (1), 62-73.

198. Heldt, C. L.; Sorci, M.; Posada, D.; Hirsa, A.; Belfort, G., Detection and Reduction of Microaggregates in Insulin Preparations. *Biotechnology and Bioengineering* **2010,** 108, (1), 237-241.

199. Manno, M.; Giacomazza, D.; Newman, J.; Martorana, V.; San Biagio, P. L., Amyloid Gels: Precocious Appearance of Elastic Properties during the Formation of an Insulin Fibrillar Network. *Langmuir* **2009,** 26, (3), 1424-1426.

200. Fodera, V.; Cataldo, S.; Librizzi, F.; Pignataro, B.; Spiccia, P.; Leone, M., Self-Organization Pathways and Spatial Heterogeneity in Insulin Amyloid Fibril Formation. *Journal of Physical Chemistry B* **2009,** 113, (31), 10830-10837.

201. Xue, W. F.; Homans, S. W.; Radford, S. E., Systematic analysis of nucleation-dependent polymerization reveals new insights into the mechanism of amyloid self-assembly. *Proceedings of the National Academy of Sciences of the United States of America* **2008,** 105, (26), 8926-8931.

202. Fawzi, N. L.; Okabe, Y.; Yap, E. H.; Head-Gordon, T., Determining the critical nucleus and mechanism of fibril elongation of the Alzheimer's A beta(1-40) peptide. *Journal of Molecular Biology* **2007,** 365, (2), 535-550.

203. Fawzi, N. L.; Kohlstedt, K. L.; Okabe, Y.; Head-Gordon, T., Protofibril assemblies of the arctic, dutch, and flemish mutants of the Alzheimer's A beta(1-40) peptide. *Biophysical Journal* **2008,** 94, (6), 2007-2016.

CHAPTER 6 THE ATOMIC LEVEL INTERACTION OF POLYPHENOLS WITH THE OLIGOMER AGGREGATE OF VQIVYK SEGMENT FROM TAU PEPTIDE

6.1 Background

The presence in tissues of amyloid plaques consisting mainly of amyloid fibrils arising from the polymerization of specific peptides/proteins is a key hallmark of several degenerative conditions including Alzheimer's, Parkinson's and type II diabetes.[137] In Alzheimer's disease, the tau peptide forms intracellular amyloid in the form of paired helical filaments (PHF).[204,205] The tau protein consisting of 441 residues binds and stabilizes microtubules.[204] The tau peptides aggregate via cross-β hydrogen bonding, where two monomer β-sheet structures adhere together.[205] The most important amino acid sequence in the tau peptide is ^{306}VQIVYK,[311] as this sequence has been shown to be necessary for amyloid formation through the cross-β interactions.[205] The hexapeptide VQIVYK by itself forms insoluble β-sheet aggregates spontaneously in aqueous solution. [22] The VQIVYK segment of tau was suggested as the minimal interaction motif for fiber formation.[205] Landau *et al* [206] recently determined the atomic structures of VQIVYK segment from the tau in complex with small molecule binders, determined by X-ray micro-crystallography. The fiber-like complexes consist of pairs of β-sheets, with small molecules binding between the sheets, roughly parallel to the fiber axis.[206] Landau *et al* [206] proposed that the tube-like cavity along the β-sheets provides an adequate site for the binding various aromatic compounds, such as polyphenols.[207]

Amyloid fibrils exhibit a common molecular architecture in which arrays of β-strands are connected by hydrogen bonds oriented parallel to the fiber long axis, into an array known as a cross-β structure. The substructure of mature fibrils consists of one or more protofilament units, which can assemble laterally or intertwine in various ways as rope-like or ribbon-like modifications to the common fibrillar framework. [2] The finding that amyloid fibrils are stabilized primarily by hydrogen bonds involving the polypeptide main chain explains why fibrils formed from polypeptides of different sequence are morphologically and structurally similar.[208] Presently, prefibrillar aggregates of different proteins and peptides are considered the most toxic amyloid species, whereas mature fibrils are substantially devoid of cytotoxicity.[137] Accordingly, intra-cellular or extracellular prefibrillar aggregates are considered the main factors for cell impairment and tissue degeneration in amyloid diseases.[209,210] Therefore, agents that interfere with early oligomerization are expected to be especially valuable for use in the therapy or prevention of amyloid diseases. The toxic effects of amyloid aggregates to exposed cells, includes nonspecific membrane permeabilization, oxidative stress, mitochondria impairment and eventually apoptosis.[211,212] A lot of many efforts are presently spent to find out naturally occurring molecules, including polyphenols, or to design synthetic ones, that are capable to protect cells against oxidative stress or the inhibition of the amyloid formation at its earliest stages and disruption of the fibrillar structures.[42,212] At present, there are no approved therapies that target amyloid formation directly, but many organic molecules have been shown to inhibit fibrillation *in vitro*, and thus represent an increasing list of proposed antiamyloid lead compounds. Natural polyphenolic compounds from foods and traditional herbal medicines, having broad pharmacological activities and exhibiting inhibition of amyloid formation, have

been extensively investigated in the disruption of mature amyloid fibrils and reduction of the toxicity of fibrils to living cells.[213, 214, 215,216] Recent publications [217, 218, 219], have studied the antiamyloid effects of natural polyphenols on three consecutive processes: formation of nascent fibrils, elongation or extension of the fibrils, and destabilization of the formed mature assemblies. The destabilizing effects include disaggregation/fragmentation of the fibrils and conversion of the fibrils into amorphous deposition.[220] Although many attempts have been made to elucidate the molecular mechanism of natural polyphenols against amyloidogenesis, the structure-activity relationship is still obscure and remains to be further explored.

Recent *in vitro* evidence has suggested that polyphenolic compounds (flavonoids) from food products such as red wine and green tea have been reported to show antiamyloid activity. [221 222, 223, 219] Despite progress in experimental observations, there are still many questions about amylin-resveratrol or amylin polyphenol interactions on a molecular level. For example, 1) What are the physicochemical factors controlling polyphenol binding? 2) What are the molecular interactions between polyphenols and the VQIVYK oligomers? 3) Does the polyphenols binding induce changes the VQIVYK oligomers structure? 4) Are there difference in binding affinity among the different polyphenols? Answering these questions will be important to our understanding of the mechanism of VQIVYK oligomer fibril dissociation induced by polyphenol and may aid in designing new antiaggregation agents. Many compounds have been reported to show antiamyloid activity in various *in vitro* and *in vivo* experiments. Detailed structural studies of the mechanism of action of already available antiamyloids can help in future development and characterization of druggable modalities. All-atom computer simulations, such as molecular

dynamics (MD), are well suited to provide molecular-level details of VQIVYK oligomer – polyphenol interactions.

There have been several theoretical attempts to study the interactions between current inhibitors and oligomers at the atomic level. Using all-atom molecular dynamics simulations with explicit solvation model, Wu *et a*l. [100] have identified and characterized two specific binding modes of Congo red molecules to protofibrils formed by GNNQQNY. Binding of the fluorescence dye thioflavin T (ThT) to the fibrils formed by Aβ fragments, Aβ16–22, has been probed using all-atom MD.[224] Two ThT binding sites were identified, one in the hydrophobic groove on the fibril side and another on the fibril edge. From MD simulations, the binding energetics for ThT was also computed. More recently, binding of tricyclic planar ligands (9, 10-anthraquinone and anthracene) to fibril forming Aβ fragments (Aβ14–20) was investigated using MD.[225] The results showed that 9, 10-anthraquinone interferes with the formation of inter-strand hydrogen bonds and reduces the accumulation of ordered aggregates. Dmitri and coworkers [226, 87] using replica exchange molecular dynamics and atomistic implicit solvent model, studied the mechanisms of binding of naproxen and ibuprofen to the A beta fibril. Liu *et al*. [227] investigated the molecular mechanism of the inhibition effect of trehalose on Aβ16–22 and Aβ40 peptides with MD in explicit solvent. Neil *et al* [228] using molecular dynamics simulations compared the mode of interaction of an active (LPFFD) and inactive (LHFFD) β-sheet breaker peptide with an Aβ fibril structure. They found that LHFFD had a weaker interaction with the fibril than the active peptide, LPFFD, from geometric and energetic considerations. [228] Recently we performed an implicit solvent molecular simulations of amyloidogenic peptides (GNNQNNY) co-incubated with polyphenols to probe the interaction between the ligand the amyloid aggregate models [86].

Lemkul *et al* [229] using multiple dynamic simulations found the flavoniod morin can bind to the ends of the fibrils blocking the attachment of an incoming peptide, penetrate into the hydrophobic core to disrupt the Asp23−Lys28 salt bridges. They found combination of hydrophobicity, aromaticity, and hydrogen bonding capacity of morin as a main factor destabilizing the $A\beta_{42}$ protofibril. Various labs reported that polyphenols physically disrupt tau aggregates.[230,231,232,233] It is this information that motivates this study. Clearly, there is great interest in understanding how small molecules might interact with, and ultimately destabilize, amyloid assemblies.

Polyphenolic compounds are attractive therapeutic candidates, as they are found in natural food products, are capable of crossing the blood−brain barrier, and are nontoxic in clinically relevant doses.[234,235] Studies of the interaction of polyphenols such as epigallocatechin [236] and resveratrol [237] with α-synuclein and Aβ lead to the proposal, based in part on seeding studies, that polyphenols functions as amyloid aggregation inhibitor by diverting polypeptides from their normal amyloid formation pathway into nonproductive off-pathway states. Polyphenols are characterized by the presence of several phenolic hydroxyl groups with acidic property and with their planar structures their planar structure forms hydrogen bonds with peptides.[238]

The VQIVYK segment of tau was suggested as the critical for tau polymerization. Therefore, the tau peptide segment (VQIVYK) has been used as a simplified model system to facilitate the discovery of key factors underlying amyloid fibril formation and the development of antiamyloid agents. The structure of the hexapeptide with an amino acid sequence VQIVYK (Residues 306-311) from tau protein in complex with small molecules has been recently been

determined by Landau *et al*.[206] The atomic structures of small molecules bound to amyloid reveal the molecular framework of small-molecule binding, within cylindrical cavities running along the β-spines of the fibers. These complexes reveal a molecular framework which partially defines the amyloid pharmacophore, the structural features responsible for the binding of small molecules to amyloid aggregates.

MD studies of polyphenols binding to VQIVYK oligomer of tau peptide to the best of our knowledge have not been performed. We used MD simulations with explicit-solvent to study the interaction of curcumin, myricetin and exifone with a preformed oligomers aggregate of VQIVYK from tau peptide. Detailed binding free energies between curcumin, exifone and myricetin and individual protein residues of the oligomers of VQIVYK were computed by using a per-residue basis decomposition method, which provides insights into the inhibitor-protein binding mechanism and also explains the mechanisms of the aggregation inhibitor effect of polyphenols.

6.2 Methods

The structure of the oligomer aggregate of the hexapeptide with an amino acid sequence VQIVYK (Residues 306-311) from tau protein in complex with small inhibitor molecules has been recently been determined by Landau *et al* [206] The X-ray structure of VQIVYK oligomer bound the polyphenol curcumin, taken from the web page http://people.mbi.ucla.edu/meytal/CoCrystalPaper/#V6K-CUR, served as the starting point for modeling VQIVYK complex with myricetin and exifone. The myricetin and exifone were docked to the peptide fibrillar structure using Sirius graphics program

(http://www.ngbw.org/sirius/) The schematic representation for the strcutureof the polyphenols is shown in Figure 6-1. The molecular dynamic (MD) simulation was performed using AMBER11 packcage [63] with an all atom amber99SB force field and explicit TIP3P water models. Each of the VQIVYK oligomer models with and without the polyphenols were solvated by explicit water molecules that extends 10 Å from any edge of the octahedral box to the protein atoms. Table 6-1 shows a summary of the simulations. Each system was simulated for 20 ns and the trajectories were saved at 8.0 ps intervals for further analysis. A hydrogen bond was assigned if the distance between donor D and acceptor A is ≤ 3.5 Å and the angle D-H …A ≥ 120^{0} [239]. Structural analysis was performed using the PTRAJ module of the AMBER 11[63] software package. VMD (visual molecular dynamics) [155] program was used for the visualization of trajectories. The MM-PBSA single trajectory approach implemented as script (MMPBSA.py) in Amber11, was used to calculate the binding energy. Solute entropic contributions were not calculated in this study since they are only crudely estimated by normal mode analysis. Although the MM-GBSA(MM-PBSA) calculations may overestimate the absolute binding free energy due to the missing terms (e.g., conformational entropy change of the solute upon binding) and underestimate the desolvation free energy, they usually give a reasonable qualitative estimate on the relative binding free energy when two similar ligands are compared.[80,240]

6.3 Results

To examine the structural stability of the VQIVYK oligomers with and without curcumin, exifone and myricetin we analyzed, the C_α root-mean-square deviation (RMSD) versus time, the C_α root-mean-square fluctuation (RMSF) as a function of the residue number,

twist angle and hydrogen-bonding pattern. All the reported quantities have been computed over the 20 ns of the production simulations of each system. The conformational stabilities of the VQIVYK oligomers with and without the ligands were monitored by the time evolution of the backbone root mean square (RMSD) and root mean square fluctuation (RMSF) relative to their initial energy minimized structure as shown in Table 6-2 and Figures 6-2. The RMSDs and RMSF provide useful information on relative stability of the oligomers, and were previously used in stability analyses of short amyloid oligomers with β-sheet structure.[86, 241, 149] To compare the effect of polyphenols binding to the VQIVYK oligomer we calculated the RMSD of Cα of the apo form and in complex with the ligand of the oligomer as shown in Figure 6-2A. The overall structure of the aggregates of the VQIVYK segment of tau peptide in complex with the polyphenol ligands is changing as is evidence by larger (≥2.0 Å) deviation in Cα of the complex compared to the negative control aggregate model. The lowest deviation from the starting structure is detected for VQIVYK when bound to exifone. The highest rms deviations for VQIVYK oligiomer was observed when the oligomer is bound to curcumin. Root mean square fluctuation(in Å) from the initial structure of the VQIVYK-backbone atoms over the time course of the molecular dynamics simulation when bound to different polyphenol ligands is shown in Figure 6-2B. The complexation with the ligand also affects the RMSF compared to the apo form of the aggregate (see Figure 6-3). The RMSD of the ligands along the simulation time is shown in Figure 6-2B curcumin shows largest RMSD (1-3.5 Å), the RMSD of myricetin was within 0.5 to 1.0 Å while the RMSD of exifone remains about 0.5 Å. Armstrong *et al.*[242] have suggested the planar structure of phenolic compounds could contribute to their effectiveness as inhibiting aggregation by allowing them to intercalate between monomer layers. The hydrogen bond

analysis of curcumin (the most flexible ligand with few hydroxyl groups) indicates ligand flexibility and number of strong hydrogen bond acceptor (the ketone carbonyl: two in curcumin versus one in the other ligand) plays an important role in interaction of the ligand with the peptides. A proper balance between molecular flexibility and number of strong hydrogen bond donor/acceptor group could play a role in the designing of more potent polyphenols as aggregation inhibitor.[229]

6.3.1 Average twist angle

Amyloid fibrils typically exhibit twisted β-sheets, as observed by electron microscopy and solid state NMR. Since twisted β-sheets optimize the hydrogen bonds, side chain stacking, and electrostatic interactions, it is commonly accepted that twisted sheets are more stable than flat ones. While twisting, the β-sheets pairs remain to be complimentary via the steric zippers.[88]The twisting in the SH4-ST7 aggregate of VQIVTK hexapeptide was evaluated by considering pairs of dihedral angles, one per each sheet of the pair. Each dihedral angle is calculated from the coordinates of the Cα(Gln2) and the Cα(Tyr6) atom of the second and the six strand of the sheet. Twisting angles have been computed by using the five inner strands.[79] As shown in Table 6-2, for the VQIVYK oligomer with and without polyphenolic ligand, the average twist angle of the oligomer with the ligand is larger than without the ligands. A large twist angle between two adjacent strands may lead them to tear away from one another at a relatively early stage of the simulation by disrupting the main chain and side-chain interactions necessary for maintaining the bulk structure. The presence of the ligand leads is structural

disruption exposing hydrophobic segments towards the polar solvent and the consequent destabilization of the bulk organization.

6.3.2 Hydrogen bonding analysis

The ordered oligomer of VQIVYK is stabilized by an extensive network of inter-peptide H-bonds. Hydrogen bonds contents of β-sheets were used previously to judge the structural integrity and stability of the various β-sheet aggregates. We did an inter-peptide hydrogen analysis and found the inter-peptide hydrogen bond content of VQIVYK oligomer with and without polyphenols was stable across all simulations (see Figure 6-3A and Table 6-2). Experimental and theoretical studies have suggested that the protein–polyphenol strong association is driven by hydrophobic effects and stabilized by hydrogen-bonding interactions. [88] The polyphenol with hydroxyl groups, competitively interact with peptides through hydrogen bonds. The analysis of the hydrogen bonds present between the oligomer aggregates and polyphenols with time was obtained from a trajectory using ptraj module in AMBER11 and is shown in Table 6-2 and Figure 6-3B. Polyphenols binds reversibly and relatively weakly with peptide molecules. The pattern of H-bond formation between the polyphenols and VQIVYK oligomer have been analyzed in detail by considering percentage of occupancy (percentage of structures exhibiting the particular type of H-bond) and are shown in Table 6-3. We investigated the hydrogen bonds between polyphenols and the adjacent residues. The hydrogen bond occupancy (Table 6-4) shows the hydrogen bond with significantly high occupancy between the ligand and the peptide occurs with tyrosine and lysine residues that are close to the ligand as

shown in Table 6-3. In the case of curcumin the hydrogen bond acceptor with highest percentage occupancy is from the carbonyl ketones (see Table 6-3).

6.3.3 Energetic Analysis of the binding

By the MM-PB(GB)SA analysis, the total free energy of binding could be separated into electrostatic, van der Waals, and solute-solvent interactions, gaining, additional insights into the physics of the VQIVYK oligomer-polyphenol association process. A single trajectory method was employed such that the snapshot coordinates for both the bound and unbound states were obtained from a single molecular dynamics simulation. For this analysis, 2500 equally spaced snapshots were taken at intervals of 80 ps from the 20.0 ns production simulation of each MD trajectory. The binding free energy and the energy components of the polyphenols and VQIVYK oligomer complexes are summarized in Table 6-4. Both van der Waals and electrostatic contributions are relevant to the interaction. According to Table 6-4, electrostatic (ΔE_{ele}) and van der Waals (ΔE_{vdw}) terms in the gas phase provide the major favorable contributions to the polyphenols binding, whereas polar solvation energies ($\Delta G_{PB(GB)}$) impair the binding. The nonpolar solvation energies (ΔG_{SA}), which correspond to the burial of SASA upon binding, barely contribute to the polyphenols binding. Further insight into the forces involved in polyphenols and VQIVYK oligomer complex formation can be obtained by analyzing the electrostatic and non-electrostatic contributions in Table 6-4. As demonstrated by numerous studies, the electrostatic contribution generally disfavors the docking of ligand and receptor molecules because the unfavorable change in the electrostatics of solvation is mostly, but not fully, compensated by the favorable electrostatics within the resulting ligand-receptor complex[243]

Indeed, from Table 6-4, despite the favorable electrostatic energies in the gas phase (ΔE_{ele}), the contributions of polar solvation energies to binding ($\Delta G_{PB(GB)}$) are unfavorable for the 3 complexes, and the sum of ΔE_{ele} and $\Delta G_{PB(GB)}$, does not favor the binding. Table 6-4 also suggests that the net result of non-electrostatic interaction which is the sum of ΔE_{vdw} and ΔG_{SA}, is favorable for the formation of the complexes, and it this behavior has been proposed previously as a general trend for noncovalent ligand-receptor associations.[244] From the above results, we can conclude that the binding free energies obtained for these complexes are driven by more favorable nonpolar interactions rather than by electrostatic interactions. To provide basic information on the most important residues in the binding of polyphenols to the VQIVYK oligomer an inhibitor-residue free energy decomposition analysis was performed. The calculation was done over the 2500 MD snapshots taken from the 20 ns simulation. According to the free energy decomposition analysis (Figure 6-4), the binding between the VQIVYK and the polyphenol is driven by selected "hot spots" that play a major role in VQIVYK −polyphenol recognition. The most important residues are Ile9, Ile33, Ile57, Ile81, Ile129 and Tyr149. Myricetin and exifone have been reported as tau aggregation inhibitors with a 1.2 μM and 3.2 μM IC50 respectively. The binding energy analysis is in qualitative agreement with the experimental observation.[231] The result form the simulation indicates that using the VQIVYK oligomer structure as a pharmacophore for tau amyloid in combination with docking and MD simulation could be effective in the virtual screening for lead discovery of small molecule aggregation inhibitors. Recently Okimoto *et al* [83]used MD simulation of a protein-ligand conformation obtained from molecular docking to estimate the binding free energies using MM-PBSA method and for ligand ranking. The combined docking and MD simulation was found to

improve by 1.6 to 4.0 time the enrichment performance compared to the used of docking method alone. The binding of the polyphenols with the peptide disrupts the surface pattern thus increases the solubility of small protofibrils. Such interaction remodels the tau oligomer into a confirmation that is different from the peptide without the ligand (see Figure 6-5). The remodeling of the peptide by the polyphenols will prevent the growth of the aggregate and will lead to disaggregation of the tau.

6.4 Conclusions

The results from this work provide a valuable insight into the mechanism of the interaction of polyphenols with the short segment of tau amyloid peptide. The study of the VQIVYK oligomer pharmacophore of tau amyloid with polyphenols in an explicit solvent may prove valuable in the future design and search of tau amyloid aggregation inhibitor.

1. Polyphenol planarity with certain flexibility and presence of a strong hydrogen bond acceptor (the ketone carbonyl) for formation of hydrogen bond with the residues of the peptide closer to the ligand.

2. The overall structure of the aggregates of the VQIVYK segment of tau peptide in complex with the polyphenol ligands compared to the negative control aggregate model is changing as is evidence by larger RMSD, RMSF and twist angles indicating the remodeling of the aggregate by the polyphenol molecules.

3. The binding free energy calculation showed electrostatic (ΔE_{ele}) and van der Waals (ΔE_{vdw}) terms in the gas phase provide the major favorable contributions to the polyphenols

binding, whereas polar solvation energies ($\Delta G_{PB(GB)}$) impair the binding. The nonpolar solvation energies (ΔG_{SA}), which correspond to the burial of SASA upon binding, barely contribute to the polyphenols binding. The free energy decomposition analysis of the binding between the VQIVYK and the polyphenol is driven by selected "hot spots" that play a major role in VQIVYK −polyphenol recognition. The most important residues are Ile9, Ile33, Ile57, Ile81, Ile129, and Tyr149.

4. The MM-PBSA (MM-GBSA) ranking of the polyphenols is in qualitative agreement with their experimental binding ranking. Thus use of VQIVYK oligomer as pharmacophore for tau amyloid in combination with docking and MD simulation could be an effective strategy in the virtual screening for lead discovery of small molecule tau aggregation inhibitors.

Table 6-1 Summary of simulated system of polyphenols with tau peptide segment VQIVYK and control oligomer

ID[†]	Content	Number of water molecules	Systems	Box size	Simulation length (ns)
1	VQIVYK oligomer with a single of curcumin	5774	Four sheets, seven strands (SH4-ST7)	77.21×77.21×77.21	20
2	VQIVYK oligomer with a single of exofine	5434	Four sheets, seven strands (SH4-ST7)	75.99×75.99 ×75.99	20
3	VQIVYK oligomer with a single of myricetin	5767	Four sheets, seven strands (SH4-ST7)	77.24×77.24×77.24	20
4	Control	5956	Four sheets, seven strands (SH4-ST7)	77.96×77.96×77.96	20

[†] The VQIVYK oligomer consists of a total four sheets with seven strands per sheets. The oligomer is organization with parallel β-sheet between strand per sheet and antiparallel between two sheets).

Table 6-2 Summary of structural analysis VQIVYK oligomers with and without polyphenols

Geometric parameters		Models name			
		VQIVYK oligomer, control	VQIVYK oligomer with curcumin	VQIVYK oligomer with exofine	VQIVYK oligomer with myricetin
	<RMSD>	5.45(1.69)	10.58(2.61)	7.21(1.14)	8.82(1.87)
	<RMSF>	2.59(1.95)	4.70(1.93)	2.66(0.98)	3.51(1.22)
Average inter-strand twist angle	Sheet 1	18.16(14.69)	-8.70(3.85)	-17.45(3.04)	-4.50(2.75)
	Sheet 2	-13.60(1.18)	-14.22 (1.46)	-14.41(1.32)	-15.37(3.83)
	Sheet 3	-21.47(2.16)	5.16(9.96)	-11.33(3.13)	-12.83(3.03)
	Sheet 4	-12.06(1.34)	-13.67(1.22)	-19.60(3.19)	-13.84(1.41)
Oligomer inter-peptide H-bonds		87.31(3.43)	85.22(3.32)	84.41(4.30)	81.42(5.06)
H-bond of ligand with oligomer		-	0.92(1.41)	0.51(0.86)	1.02(0.69)

Each number is averaged over 20 ns trajectory and the numbers in parenthesis are standard deviations.

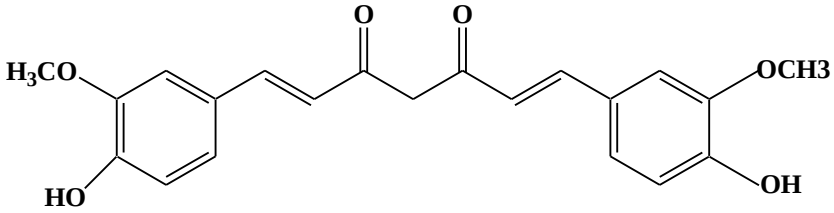
Table 6-3 Hydrogen bond occupancy between the polyphenols and the oligomer of VQIVYK segment of tau peptide

VQIVYK with curcumin Hydrogen bond type		(%)	VQIVYK with exifone Hydrogen bond type		(%)	VQIVYK with myricetin Hydrogen bond type		(%)
donor	Acceptor		donor	Acceptor		donor	Acceptor	
curcumin@O34	TYR131@HH	71.8	exifone @O20	TYR 35@HH	16.8	myricetin @O12	TYR35@HH	60.5
curcumin @O33	TYR131@HH	34.8	exifone @H30	TYR35@HH	14.6	TYR11@OH	myricetin @H30	46.4
curcumin @O33	TYR11@HH	26.2	exifone @O18	TYR35@HH	8.9	TYR131@OH	myricetin @H31	21.0
curcumin @OXT	curcumin @H36	15.8	exifone @O18	TYR11@HH	5.7	myricetin 169@O22	TYR11@HH	12.2
curcumin @O34	TYR155@HH	13.4	LYS54@OXT	LYS103@H1	2.9	LYS54@OXT	myricetin @H29	12.0

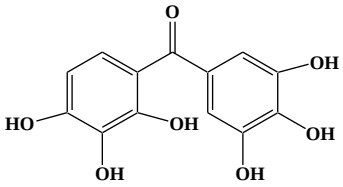
Table 6-4 Summary of the MM-GB(PB)SA energy component analysis of the polyphenols with the VQIVYK oligomers

MMGBSA	Model name		
	VQIVYK oligomer with curcumin	VQIVYK oligomer with exifone	VQIVYK oligomer with myricetin
$<\Delta E_{vdw}>$	-28.29± 4.53	-22.06±2.86	-32.39±5.35
$<\Delta E_{ele}>$	-18.58±9.09	-8.71±8.38	-15.62±8.31
$<\Delta G_{PB}>$	36.71±9.06	29.14±7.96	38.84±8.66
$<\Delta G_{SA}>$	-5.26±0.59	-3.89±0.33	-4.83±0.36
$<\Delta G_{solv}>$	31.45±9.08	25.25±7.97	34.01±8.67
$<\Delta G_{binding}>$	-15.42±3.26	-5.5147±2.64	-14.02±4.33
MMPBSA			
$<\Delta E_{vdw}>$	-28.29±4.53	-22.06±2.86	-32.39±5.35
$<\Delta E_{ele}>$	-18.58±9.09	-8.71±8.38	-15.62±8.31
$<\Delta G_{PB}>$	35.67± 9.67	28.35± 8.46	41.03± 10.85
$<\Delta G_{SA}>$	-3.02± 0.45	-1.84±0.25	-2.53±0.25
$<\Delta G_{solv}>$	32.66± 9.68	26.51± 8.46	38.50± 10.85
$<\Delta G_{binding}>$	-14.21±3.88	-4.25±4.59	-9.51±4.25

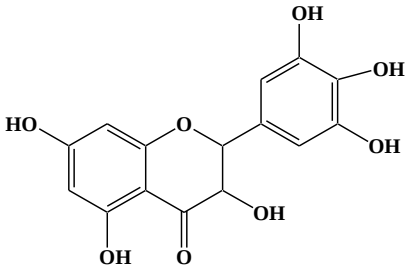
* Binding free energy components (kcal mol^{-1}) and standard deviations calculated with MM-PBSA/ MM-GBSA for VQIVYK oligomer (SH4-ST7): Average over 2500 snapshots of the trajectory. [b] Δ $\mathbf{E_{vdw}}$, non-bonded van der Waals energy; ΔE_{ele}, non-solvent electrostatic potential energy; ΔG_{PB}, electrostatic contributions to the solvation free energy calculated with Poisson-Boltzmann equation; ΔG_{SA} , ΔG_{Solv} are nonpolar and total solvation energies; $\Delta G_{\mathbf{binding}}$ binding energy of the system. All energies are in kcal/mol: $\Delta G_{binding} = \Delta E_{vdw} + \Delta E_{ele} + \Delta G_{sol}$; $\Delta G_{sol} = \Delta G_{PB} + \Delta G_{SA}$;



A

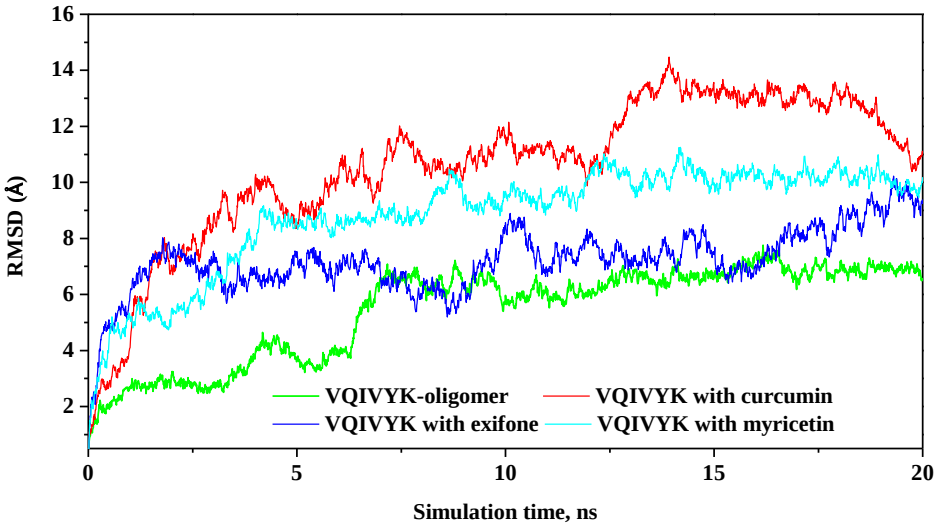


B



C

Figure 6-1 Chemical structures of curcumin (A), exifone (B) and myricetin (C)



A

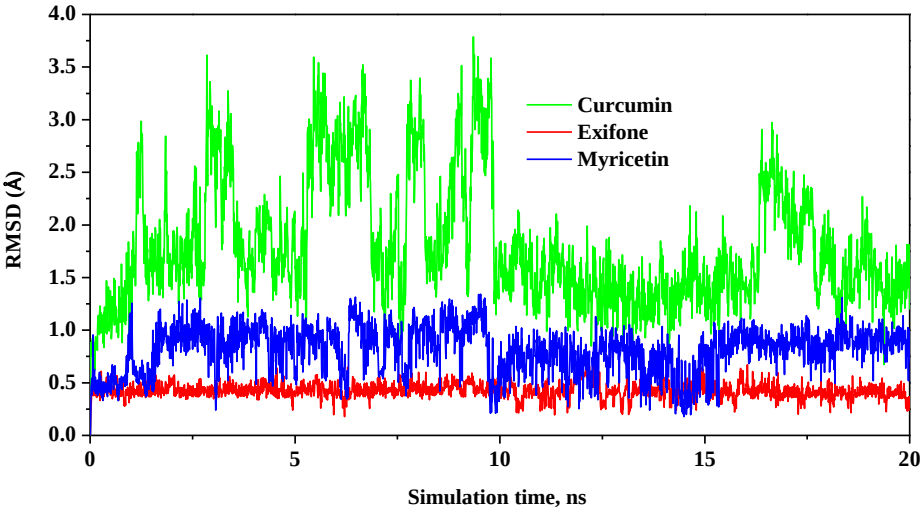


Figure 6-2 RMDS as a function of the simulation time (A) Comparative RMSD analyses of VQIVYK oligomer control and in complex with curcumin, exifone and myricetin and (B) comparative RMSD of the inhibitor molecules with respect to the first snapshot

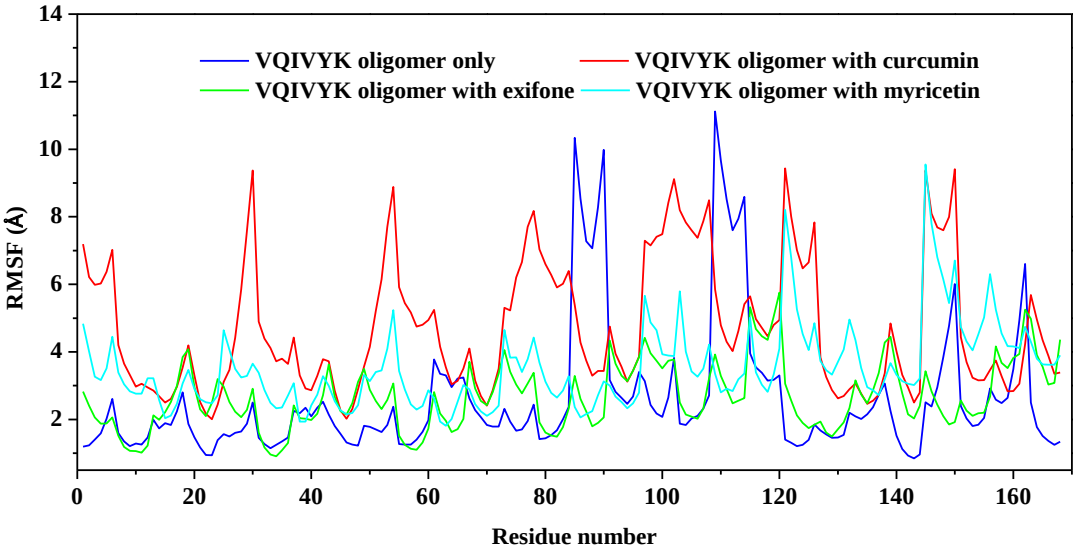
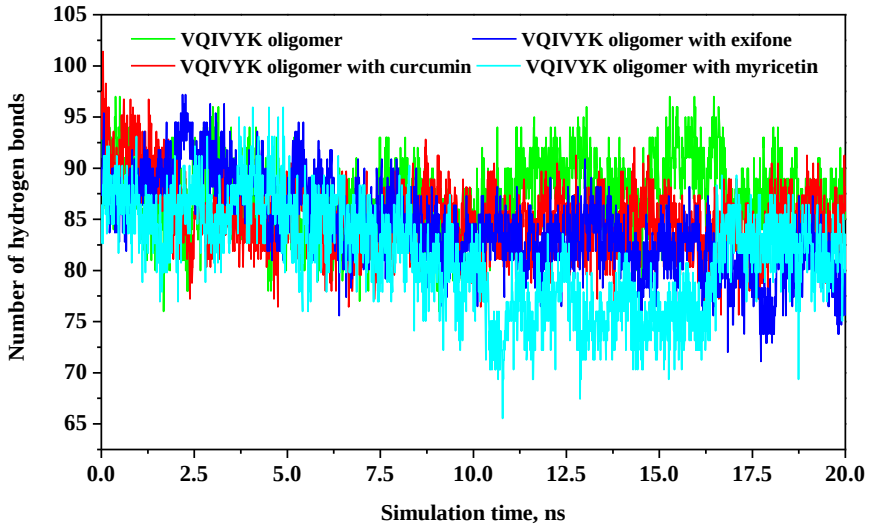
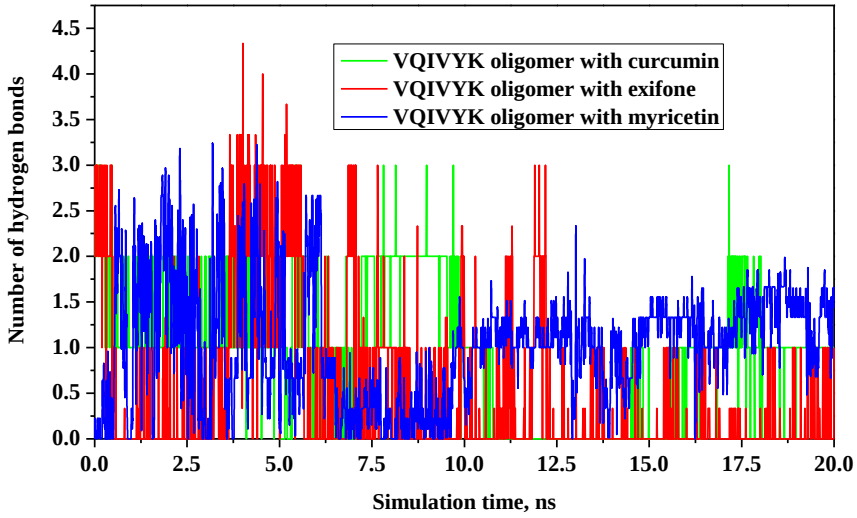


Figure 6-3 Comparative RMSF analysis of the VQIVYK oligomer control and in complex with curcumin, exifoneand myricetin



A



B

Figure 6-4 Time evolution of the number of hydrogen bonds (A) Total inter peptide hydrogen bonds (B) hydrogen bonds between the polyphenols and VQIVYK oligomer

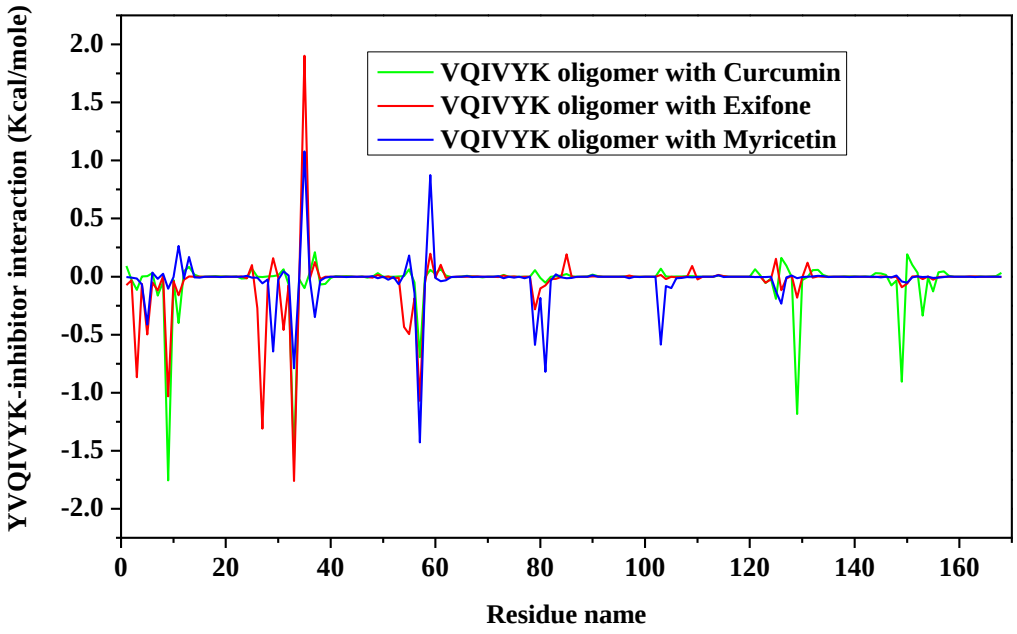


Figure 6-5 Decomposition of the free energy on a per residue basis for VQIVYK oligomer and the polyphenol interaction

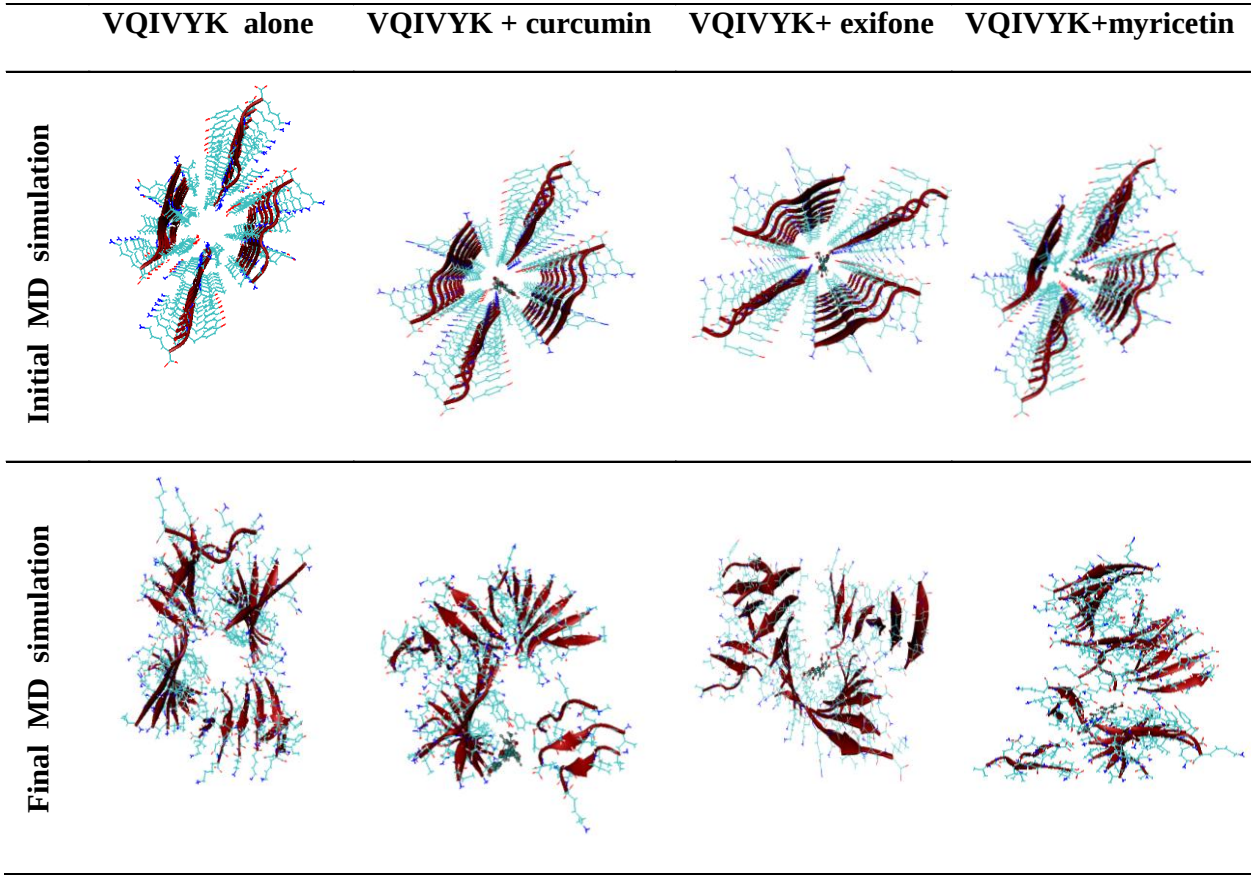


Figure 6-6 The structure of the initial and final structure after 20 ns for the VQIVYK oligomer with and without polyphenols

6.5 Refereneces

204. Chisato Nishiuraa, K. T., Katsuhiko Minouraa, Miho Sumidab, Taizo Taniguchib, Koji Tomooa, and Toshimasa Ishidaa, Importance of Tyr310 residue in the third repeat of microtubule binding domain for filament formation of tau protein. *J Biochem* **2010,** 147, 405-414.

205. M. von Bergen, P. F., J. Biernat, J. Heberle, E.-M. Mandelkow, and E. Mandelkow, Assembly of τ protein into Alzheimer paired helical filaments depends on a local sequence motif (306VQIVYK311) forming β structure. *Proc Natl Acad sci* **2000,** 97, 5129-5134.

206. Landau M, S. M., Faull KF, Laganowsky A, Jiang L, Sievers SA, Liu J, Barrio JR, Eisenberg D., Towards a Pharmacophore for Amyloid. *PLoS Biol.* **2011,** 9(6):e1001080.

207. Porat Y, A. A., Gazit E Inhibition of amyloid fibril formation by polyphenols: structural similarity and aromatic interactions as a common inhibition mechanism. *Chem Biol Drug Des* **2006,** 67, 27–37.

208. Dobson, C. M., Protein folding and misfolding. *Nature* **2003,** 426, (6968), 884-890.

209. Stefani, M.; Dobson, C. M., Protein aggregation and aggregate toxicity: new insights into protein folding, misfolding diseases and biological evolution. *Journal of Molecular Medicine-Jmm* **2003,** 81, (11), 678-699.

210. Haataja, L.; Gurlo, T.; Huang, C. J.; Butler, P. C., Islet amyloid in type 2 diabetes, and the toxic oligomer hypothesis. *Endocrine Reviews* **2008,** 29, (3), 303-316.

211. Bucciantini, M.; Calloni, G.; Chiti, F.; Formigli, L.; Nosi, D.; Dobson, C. M.; Stefani, M., Prefibrillar amyloid protein aggregates share common features of cytotoxicity. *Journal of Biological Chemistry* **2004,** 279, (30), 31374-31382.

212. Butterfield, D. A., Amyloid beta-peptide (1-42)-induced oxidative stress and neurotoxicity: Implications for neurodegeneration in Alzheimer's disease brain. A review. *Free Radical Research* **2002,** 36, (12), 1307-1313.

213. Zhu, M.; Rajamani, S.; Kaylor, J.; Han, S.; Zhou, F. M.; Fink, A. L., The flavonoid baicalein inhibits fibrillation of alpha-synuclein and disaggregates existing fibrils. *Journal of Biological Chemistry* **2004,** 279, (26), 26846-26857.

214. Zhu, J. T. T.; Choi, R. C. Y.; Chu, G. K. Y.; Cheung, A. W. H.; Gao, Q. T.; Li, J.; Jiang, Z. Y.; Dong, T. T. X.; Tsim, K. W. K., Flavonoids possess neuroprotective effects on cultured pheochromocytoma PC12 cells: A comparison of different flavonoids in activating estrogenic effect and in preventing beta-amyloid-induced cell death. *Journal of Agricultural and Food Chemistry* **2007,** 55, (6), 2438-2445.

215. Akaishi, T.; Morimoto, T.; Shibao, M.; Watanabe, S.; Sakai-Kato, K.; Utsunomiya-Tate, N.; Abe, K., Structural requirements for the flavonoid fisetin in inhibiting fibril formation of amyloid beta protein. *Neuroscience Letters* **2008,** 444, (3), 280-285.

216. Kim, H.; Park, B. S.; Lee, K. G.; Choi, C. Y.; Jang, S. S.; Kim, Y. H.; Lee, S. E., Effects of naturally occurring compounds on fibril formation and oxidative stress of beta-amyloid. *Journal of Agricultural and Food Chemistry* **2005,** 53, (22), 8537-8541.

217. Masuda, M.; Suzuki, N.; Taniguchi, S.; Oikawa, T.; Nonaka, T.; Iwatsubo, T.; Hisanaga, S.; Goedert, M.; Hasegawa, M., Small molecule inhibitors of alpha-synuclein filament assembly. *Biochemistry* **2006,** 45, (19), 6085-6094.

218. Shoval, H.; Lichtenberg, D.; Gazit, E., The molecular mechanisms of the anti-amyloid effects of phenols. *Amyloid-Journal of Protein Folding Disorders* **2007,** 14, (1), 73-87.

219. Porat, Y.; Abramowitz, A.; Gazit, E., Inhibition of amyloid fibril formation by polyphenols: Structural similarity and aromatic interactions as a common inhibition mechanism. *Chemical Biology & Drug Design* **2006,** 67, (1), 27-37.

220. He, J.; Xing, Y. F.; Huang, B.; Zhang, Y. Z.; Zeng, C. M., Tea Catechins Induce the Conversion of Preformed Lysozyme Amyloid Fibrils to Amorphous Aggregates. *Journal of Agricultural and Food Chemistry* **2009,** 57, (23), 11391-11396.

221. Ho, L.; Pasinetti, G. M., Polyphenolic compounds for treating neurodegenerative disorders involving protein misfolding. *Expert Review of Proteomics* **2010,** 7, (4), 579-589.

222. Ono, K.; Yamada, M., Antioxidant compounds have potent anti-fibrillogenic and fibril-destabilizing effects for alpha-synuclein fibrils in vitro. *Journal of Neurochemistry* **2006,** 97, (1), 105-115.

223. Bastianetto, S.; Krantic, S.; Quirion, R., Polyphenols as potential inhibitors of amyloid aggregation and toxicity: Possible significance to Alzheimer's disease. *Mini-Reviews in Medicinal Chemistry* **2008,** 8, (5), 429-435.

224. Wu, C.; Wang, Z. X.; Lei, H. X.; Duan, Y.; Bowers, M. T.; Shea, J. E., The Binding of Thioflavin T and Its Neutral Analog BTA-1 to Protofibrils of the Alzheimer's Disease A beta(16-22) Peptide Probed by Molecular Dynamics Simulations. *Journal of Molecular Biology* **2008,** 384, (3), 718-729.

225. Convertino, M.; Pellarin, R.; Catto, M.; Carotti, A.; Caflisch, A., 9,10-Anthraquinone hinders beta-aggregation: How does a small molecule interfere with A beta-peptide amyloid fibrillation? *Protein Science* **2009,** 18, (4), 792-800.

226. Takeda, T.; Chang, W. L. E.; Raman, E. P.; Klimov, D. K., Binding of nonsteroidal anti-inflammatory drugs to A beta fibril. *Proteins-Structure Function and Bioinformatics* **2010,** 78, (13), 2849-2860.

227. Liu, F. F.; Ji, L.; Dong, X. Y.; Sun, Y., Molecular Insight into the Inhibition Effect of Trehalose on the Nucleation and Elongation of Amyloid beta-Peptide Oligomers. *Journal of Physical Chemistry B* **2009,** 113, (32), 11320-11329.

228. Bruce, N. J.; Chen, D. L.; Dastidar, S. G.; Marks, G. E.; Schein, C. H.; Bryce, R. A., Molecular dynamics simulations of A beta fibril interactions with beta-sheet breaker peptides. *Peptides* **2010,** 31, (11), 2100-2108.

229. Lemkul, J. A.; Bevan, D. R., Destabilizing Alzheimer's A beta(42) Protofibrils with Morin: Mechanistic Insights from Molecular Dynamics Simulations. *Biochemistry* **2010,** 49, (18), 3935-3946.

230. Ho, L.; Yemul, S.; Wang, J.; Pasinettia, G. M., Grape Seed Polyphenolic Extract as a Potential Novel Therapeutic Agent in Tauopathies. *Journal of Alzheimers Disease* **2009,** 16, (2), 433-439.

231. Taniguchi, S.; Suzuki, N.; Masuda, M.; Hisanaga, S.; Iwatsubo, T.; Goedert, M.; Hasegawa, M., Inhibition of heparin-induced tau filament formation by phenothiazines, polyphenols, and porphyrins. *Journal of Biological Chemistry* **2005,** 280, (9), 7614-7623.

232. Wang, J.; Santa-Maria, I.; Ho, L.; Ksiezak-Reding, H.; Ono, K.; Teplow, D. B.; Pasinetti, G. M., Grape Derived Polyphenols Attenuate Tau Neuropathology in a Mouse Model of Alzheimer's Disease. *Journal of Alzheimers Disease* **2010,** 22, (2), 653-661.

233. Yang, F. S.; Lim, G. P.; Begum, A. N.; Ubeda, O. J.; Simmons, M. R.; Ambegaokar, S. S.; Chen, P. P.; Kayed, R.; Glabe, C. G.; Frautschy, S. A.; Cole, G. M., Curcumin inhibits

formation of amyloid beta oligomers and fibrils, binds plaques, and reduces amyloid in vivo. *Journal of Biological Chemistry* **2005,** 280, (7), 5892-5901.

234. Amijee, H.; Scopes, D. I. C., The Quest for Small Molecules as Amyloid Inhibiting Therapies for Alzheimer's Disease. *Journal of Alzheimers Disease* **2009,** 17, (1), 33-47.

235. Ho, L.; Pasinetti, G. M., Polyphenolic compounds for treating neurodegenerative disorders involving protein misfolding. *Expert Review of Proteomics* 7, (4), 579-589.

236. Meng, F. L.; Abedini, A.; Plesner, A.; Verchere, C. B.; Raleigh, D. P., The Flavanol (-)-Epigallocatechin 3-Gallate Inhibits Amyloid Formation by Islet Amyloid Polypeptide, Disaggregates Amyloid Fibrils, and Protects Cultured Cells against IAPP-Induced Toxicity. *Biochemistry* 49, (37), 8127-8133.

237. Ladiwala, A. R. A. L., J. C.; Bale, S. S.; Marcelino-Cruz, A. M.; Bhattacharya, M.; Dordick, J. S.; Tessier, P. M., , Resveratrol Selectively Remodels Soluble Oligomers and Fibrils of Amyloid A beta into Off-pathway Conformers. *Journal of Biological Chemistry* **2010,** 285, (31), 24228-24237.

238. Bulic B, P. M., Schmidt B, Mandelkow EM, Waldmann H, Mandelkow E., Development of tau aggregation inhibitors for Alzheimer's disease. *Angew Chem Int Ed Engl.* **2009,** 48, 1740-1752.

239. Fabiola, F.; Bertram, R.; Korostelev, A.; Chapman, M. S., An improved hydrogen bond potential: Impact on medium resolution protein structures. *Protein Science* **2002,** 11, (6), 1415-1423.

240. Gilson, M. K.; Zhou, H. X., Calculation of protein-ligand binding affinities. *Annual Review of Biophysics and Biomolecular Structure* **2007,** 36, 21-42.

241. Berhanu, W. M. M., A. E., Controlling the aggregation and rate of release in order to improve insulin formulation: molecular dynamics study of full-length insulin amyloid oligomer models. *Journal of Molecular Modeling* published online: 15 June 2011; DOI 10.1007/s00894-011-1123-3.

242. Armstrong, A. H.; Chen, J.; McKoy, A. F.; Hecht, M. H., Mutations That Replace Aromatic Side Chains Promote Aggregation of the Alzheimer's A beta Peptide. *Biochemistry* **2011,** 50, (19), 4058-4067.

243. Gutierrez, L. J.; Enriz, R. D.; Baldoni, H. A., Structural and Thermodynamic Characteristics of the Exosite Binding Pocket on the Human BACE1: A Molecular Modeling Approach. *Journal of Physical Chemistry A* **2010,** 114, (37), 10261-10269.

244. Miyamoto, S. K., P. A.,, What determines the strength of noncovalent association of ligands to proteins in aqueous-solution. *Proceedings of the National Academy of Sciences of the United States of America* **1993,** 90, (18), 8402–8406.

SUMMARY

This Thesis describes Molecular dynamic study aimed at understanding the effect steric zipper mutation, polymorphic packing and polyphenols on the aggregation of amyloid peptides.

First we investigated the effect of various sizes and arrangements of oligomer seeds of the wild-type and mutants of the three hexa-peptides fragments of Tau, Insulin and Aβ peptide (VQIVYK, MVGGVV and LYQLEN) on their structural stability and dynamics. We found the stability of the VQIVYK, MVGGVV and LYQLEN peptide oligomers increases with increasing the number of β-strand. The Sh2-St4 model was found to stable enough that could possible act as a stable seed in prompting amyloid fibril formation for all the three peptides. The binding energy calculated by MM-PBSA method and the analysis of individual contributions to the binding energy shows the hydrophobic interactions play an important role in stabilizing the structural organizations between β-sheet layers in the oligomers. The result of the binding free energy calculation also indicated that the wild type is the most stable structure compared to the mutants. The single glycine substitution at the steric zipper interface disrupts the hydrophobic steric zipper remarkably, indicating that the hydrophobic attraction is a major driving force for stabilizing and aggregation of oligomers. Consequently, the substantial reduction in the van der Waals intersheet interactions leads to destabilization of the oligomers. Overall, aggregation of both wild type and mutant peptides is driven by nonpolar interaction. Thus, designing new peptidomimetic inhibitors able to prevent the fibril formation based on the modification of steric zipper motif of the oligomers, similar to the ones examined in this study may become a viable therapeutic strategy.

Next we investigated effects of sequence and packing arrangements on five pairs of short segments of amyloid peptides with steric zipper polymorphism. The simulation revealed the amyloid peptide rich in Q/N amino acid (GNNQQNY and NNQNTF) have a greater structural stability than the short segments amyloid peptide lacking the Q/N amino acids (SSTNVG, VQIVYK and MVGGVV). The Q/N residue rich short amyloid segments have larger hydrogen bond contents and hydrogen bond occupancy. The overall increase of hydrogen bond in the Q/N residue rich peptides with smaller RMSD, RMSF and greater stability suggests their stability is mainly associated with an increase in side chain interaction and hydrogen bond contents. The simulations of Q/N→G mutants disrupted the steric zipper, leading to unstable oligomers. The The MM-PBSA binding free energy method was applied to the study of the β-sheet association. The nonpolar component of free energy is more favorable, while the electrostatic solvation is unfavorable for sheet to sheet interaction. This explains the acceleration of aggregation by adding nonpolar co-solvents (methanol, trifluoroethanol, and hexafluoroisopropanol).The decomposition of the binding energy per residue showed the contribution of the N/Q side-chains to the association of the 5 stranded double layer oligomers is larger than the other nonpolar and small size amino acids at the interface, underlining the importance of Q/N amino acid in stabilizing the short segment amyloid peptides in crystal free context.

Next we investigated the structural stability of the wild type and mutants of a single layer models of insulin aimed at the design of short- and long-acting insulin analogs. We found the stabilities of the single-layer insulin peptide oligomers increase as the number of strands increases (dynamic cooperative effect). The binding energy calculated by the MM-GBSA method shows that hydrophobic interactions play an important role in stabilizing the structural

organization of the single-layer insulin. Per-residue decomposition shows that the key amino acid residues for single layer insulin stability occur mainly in the β-sheet regions of chains A and B. The binding energy decomposition also reveals due to the electrostatic repulsion between the three negatively charged glutamates in adjacent insulin strands, electrostatic contribution to the binding energy is unfavorable. The single glycine substitution at the steric zipper interface was found to disrupts the hydrophobic contacts and reduces the van der Waals interactions in the mutants, thus reducing the binding free energy. The binding free-energy calculation indicated that the wild type is more structurally stable than most of the mutants. A comparison of the binding free energy between the wild type and the chain-A mutants (Y14GA, L16GA and N18GA) indicated that shape complementarity between neighboring strands plays a key role in stabilizing the entire oligomeric structure. The secondary structure contents and the clustering analysis of the trajectories of the single-layer insulin oligomers of various sizes showed that the larger aggregates retain the fibril conformation but the smaller ones (SH1-ST1 and SH1-ST2) lose this conformation. This observation could explain the observed shortening of the nucleation lag phase of insulin aggregation with oligomer seeds. Based on the secondary structure contents and the cluster analysis, we propose that SH1-ST4 is a critical nucleus for single-layer insulin fibril oligomer growth. Our simulations of wildtype and single glycine mutants at the steric zipper region can be targeted in the design of both short- and long-acting insulin analogs as well. Aside from the design of such insulin analogs, the present study may prove useful in the rational design of insulin aggregation inhibitors that can be used to stabilize insulin formulations, leading to safer handling and more cost-effective storage of such formulations, especially in developing countries.

Finally we investigated the interaction of polyphenols with short amyloid aggregates. The results from this work provide a valuable insight into the mechanism of the interaction of polyphenols with the short segment of tau amyloid peptide. The study of the VQIVYK oligomer pharmacophore of tau amyloid with polyphenols in an explicit solvent may prove valuable in the future design and search of tau amyloid aggregation inhibitor. Polyphenol planarity with certain flexibility and presence of a strong hydrogen bond acceptor (the ketone carbonyl) for formation of hydrogen bond with the residues of the peptide closer to the ligand. The overall structure of the aggregates of the VQIVYK segment of tau peptide in complex with the polyphenol ligands compared to the negative control aggregate model is changing as is evidence by larger RMSD, RMSF and twist angles indicating the remodeling of the aggregate by the polyphenol molecules. The binding free energy calculation showed electrostatic (ΔEele) and van der Waals (ΔEvdw) terms in the gas phase provide the major favorable contributions to the polyphenols binding, whereas polar solvation energies (ΔGPB(GB)) impair the binding. The nonpolar solvation energies (ΔGSA), which correspond to the burial of SASA upon binding, barely contribute to the polyphenols binding. The free energy decomposition analysis of the binding between the VQIVYK and the polyphenol is driven by selected "hot spots" that play a major role in VQIVYK −polyphenol recognition. The most important residues are Ile9, Ile33, Ile57, Ile81, Ile129, and Tyr149. The MM-PBSA (MM-GBSA) ranking of the polyphenols is in qualitative agreement with their experimental binding ranking. Thus use of VQIVYK oligomer as pharmacophore for tau amyloid in combination with docking and MD simulation could be an effective strategy in the virtual screening for lead discovery of small molecule tau aggregation inhibitors.

In summary, we conclude MD simulation could be used in the atomic level understanding of amyloid aggregation formation, could contributes to elucidating the driving force for the thermodynamics of the aggregation, could contribute in the structure based designing of aggregation inhibitors and in combination with docking and MM-PBSA binding free energy calculation could be useful in the virtual screening of inhibitors.

REFERENCES

1. Devlin, T. M., *Text Book of Biochemistry with clinical correlations*. Seventh ed.; John Wily and Sons, Inc New Jersey, 2010.

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

3. Whitford, D., *Protein structure and functions*. John Wiley & sons, ltd: Chichester, 2005.

4. Scheibel, J. P. Z. a. T., *Protein folding-misfolding: some current concepts of protein chemistry*. Nova Science Publishers: New York, 2007.

5. Dobson, C. M., The structural basis of protein folding and its links with human disease. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **2001,** 356, (1406), 133-145.

6. Groenning, M.; Frokjaer, S.; Vestergaard, B., Formation Mechanism of Insulin Fibrils and Structural Aspects of the Insulin Fibrillation Process. *Current Protein & Peptide Science* **2009,** 10, (5), 509-528.

7. Stine, W. B.; Dahlgren, K. N.; Krafft, G. A.; LaDu, M. J., In vitro characterization of conditions for amyloid-beta peptide oligomerization and fibrillogenesis. *Journal of Biological Chemistry* **2003,** 278, (13), 11612-11622.

8. Senguen, F. T.; Doran, T. M.; Anderson, E. A.; Nilsson, B. L., Clarifying the influence of core amino acid hydrophobicity, secondary structure propensity, and molecular volume on amyloid-beta 16-22 self-assembly. *Molecular Biosystems* **2011,** 7, (2), 497-510.

9. Tyedmers, J.; Mogk, A.; Bukau, B., Cellular strategies for controlling protein aggregation. *Nature Reviews Molecular Cell Biology* **2010,** 11, (11), 777-788.

10. Dahlgren, K. N.; Manelli, A. M.; Stine, W. B.; Baker, L. K.; Krafft, G. A.; LaDu, M. J., Oligomeric and fibrillar species of amyloid-beta peptides differentially affect neuronal viability. *Journal of Biological Chemistry* **2002,** 277, (35), 32046-32053.

11. Walsh, D. M.; Selkoe, D. J., A beta Oligomers - a decade of discovery. *Journal of Neurochemistry* **2007,** 101, (5), 1172-1184.

12. Quist, A.; Doudevski, L.; Lin, H.; Azimova, R.; Ng, D.; Frangione, B.; Kagan, B.; Ghiso, J.; Lal, R., Amyloid ion channels: A common structural link for protein-misfolding disease. *Proceedings of the National Academy of Sciences of the United States of America* **2005,** 102, (30), 10427-10432.

13. Naqvi, S. H.; Wang, W. S.; Miao, J. Y.; He, R. Q., Pore-like Aggregates of Tau Protein Induced by Formaldehyde. *Progress in Biochemistry and Biophysics* **2010,** 37, (11), 1195-1203.

14. Fowler, D. M.; Koulov, A. V.; Alory-Jost, C.; Marks, M. S.; Balch, W. E.; Kelly, J. W., Functional amyloid formation within mammalian tissue. *Plos Biology* **2006,** 4, (1), 100-107.

15. Maji, S. K.; Perrin, M. H.; Sawaya, M. R.; Jessberger, S.; Vadodaria, K.; Rissman, R. A.; Singru, P. S.; Nilsson, K. P. R.; Simon, R.; Schubert, D.; Eisenberg, D.; Rivier, J.; Sawchenko, P.; Vale, W.; Riek, R., Functional Amyloids As Natural Storage of Peptide Hormones in Pituitary Secretory Granules. *Science* **2009,** 325, (5938), 328-332.

16. Wiltzius, J. J. W.; Landau, M.; Nelson, R.; Sawaya, M. R.; Apostol, M. I.; Goldschmidt, L.; Soriaga, A. B.; Cascio, D.; Rajashankar, K.; Eisenberg, D., Molecular mechanisms for protein-encoded inheritance. *Nature Structural & Molecular Biology* **2009,** 16, (9), 973-U98.

17. Sipe, J. D.; Benson, M. D.; Buxbaum, J. N.; Ikeda, S.; Merlini, G.; Saraiva, M. J. M.; Westermark, P., Amyloid fibril protein nomenclature: 2010 recommendations from the nomenclature committee of the International Society of Amyloidosis. *Amyloid-Journal of Protein Folding Disorders* **2010,** 17, (3-4), 101-104.

18. Teng, P. K.; Eisenberg, D., Short protein segments can drive a non-fibrillizing protein into the amyloid state. *Protein Engineering Design & Selection* **2009,** 22, (8), 531-536.

19. Glenner, G. G.; Eanes, E. D.; Page, D. L., Relation of properties of congo red-stained amyloid fibrils to beta-conformation. *Journal of Histochemistry & Cytochemistry* **1972,** 20, (10), 821-&.

20. Serpell, L. C.; Berriman, J.; Jakes, R.; Goedert, M.; Crowther, R. A., Fiber diffraction of synthetic alpha-synuclein filaments shows amyloid-like cross-beta conformation. *Proceedings of the National Academy of Sciences of the United States of America* **2000,** 97, (9), 4897-4902.

21. Zheng, J.; Jang, H.; Ma, B.; Tsai, C. J.; Nussinov, R., Modeling the Alzheimer A beta(17-42) fibril architecture: Tight intermolecular sheet-sheet association and intramolecular hydrated cavities. *Biophysical Journal* **2007,** 93, (9), 3046-3057.

22. Sawaya, M. R.; Sambashivan, S.; Nelson, R.; Ivanova, M. I.; Sievers, S. A.; Apostol, M. I.; Thompson, M. J.; Balbirnie, M.; Wiltzius, J. J. W.; McFarlane, H. T.; Madsen, A. O.; Riekel, C.; Eisenberg, D., Atomic structures of amyloid cross-beta spines reveal varied steric zippers. *Nature* **2007,** 447, (7143), 453-457.

23. Luca, S.; Yau, W. M.; Leapman, R.; Tycko, R., Peptide conformation and supramolecular organization in amylin fibrils: Constraints from solid-state NMR. *Biochemistry* **2007,** 46, (47), 13505-13522.

24. Marshall, K. E.; Serpell, L. C., Fibres, crystals and polymorphism: the structural promiscuity of amyloidogenic peptides. *Soft Matter* **2010,** 6, (10), 2110-2114.

25. Bockmann, A., 3D protein structures by solid-state NMR spectroscopy: Ready for high resolution. *Angewandte Chemie-International Edition* **2008,** 47, (33), 6110-6113.

26. Goldschmidt, L.; Teng, P. K.; Riek, R.; Eisenberg, D., Identifying the amylome, proteins capable of forming amyloid-like fibrils. *Proceedings of the National Academy of Sciences of the United States of America* **2010,** 107, (8), 3487-3492.

27. Sunde, M.; Blake, C., The structure of amyloid fibrils by electron microscopy and X-ray diffraction. In *Advances in Protein Chemistry, Vol 50*, Academic Press Inc: San Diego, 1997; Vol. 50, pp 123-159.

28. Astbury, W. T., Dickinson, S, and Bailey, K., The X-ray interpretation of denaturation and the structure of the seed globulins. *Biochem. J.* **1935,** 29, 2351–0.

29. Sunde, M.; Serpell, L. C.; Bartlam, M.; Fraser, P. E.; Pepys, M. B.; Blake, C. C. F., Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *Journal of Molecular Biology* **1997,** 273, (3), 729-739.

30. Nelson, R.; Sawaya, M. R.; Balbirnie, M.; Madsen, A. O.; Riekel, C.; Grothe, R.; Eisenberg, D., Structure of the cross-beta spine of amyloid-like fibrils. *Nature* **2005,** 435, (7043), 773-778.

31. Xu, Z. P.; Paparcone, R.; Buehler, M. J., Alzheimer's A beta(1-40) Amyloid Fibrils Feature Size-Dependent Mechanical Properties. *Biophysical Journal* **2009,** 98, (10), 2053-2062.

32. Kajava, A. V.; Baxa, U.; Steven, A. C., Beta arcades: recurring motifs in naturally occurring and disease-related amyloid fibrils. *Faseb Journal* **2010,** 24, (5), 1311-1319.

33. Yu, X.; Wang, J. D.; Yang, J. C.; Wang, Q. M.; Cheng, S. Z. D.; Nussinov, R.; Zheng, J., Atomic-Scale Simulations Confirm that Soluble beta-Sheet-Rich Peptide Self-Assemblies Provide Amyloid Mimics Presenting Similar Conformational Properties. *Biophysical Journal* **2009,** 98, (1), 27-36.

34. Fowler, D. M.; Koulov, A. V.; Balch, W. E.; Kelly, J. W., Functional amyloid - from bacteria to humans. *Trends in Biochemical Sciences* **2007,** 32, (5), 217-224.

35. Harper, J. D.; Lansbury, P. T., Models of amyloid seeding in Alzheimier's disease and scrapie: Mechanistic truths and physiological consequences of the time-dependent solubility of amyloid proteins. *Annual Review of Biochemistry* **1997,** 66, 385-407.

36. Lasagna-Reeves, C. A.; Castillo-Carranza, D. L.; Guerrero-Munoz, M. J.; Jackson, G. R.; Kayed, R., Preparation and Characterization of Neurotoxic Tau Oligomers. *Biochemistry* **2010,** 49, (47), 10039-10041.

37. Andreetto, E.; Yan, L. M.; Tatarek-Nossol, M.; Velkova, A.; Frank, R.; Kapurniotu, A., Identification of Hot Regions of the A beta-IAPP Interaction Interface as High-Affinity Binding Sites in both Cross- and Self-Association. *Angewandte Chemie-International Edition* **2010,** 49, (17), 3081-3085.

38. Guo, J. P.; Arai, T.; Miklossy, J.; McGeer, P. L., A beta and tau form soluble complexes that may promote self aggregation of both into the insoluble forms observed in Alzheimer's disease. *Proceedings of the National Academy of Sciences of the United States of America* **2006,** 103, (6), 1953-1958.

39. Nicolls, M. R., The Clinical and Biological Relationship between Type II Diabetes Mellitus and Alzheimer's Disease. *Current Alzheimer Research* **2004,** 1, (1), 47-54.

40. Wills, J.; Jones, J.; Haggerty, T.; Duka, V.; Joyce, J. N.; Sidhu, A., Elevated tauopathy and alpha-synuclein pathology in postmortem Parkinson's disease brains with and without dementia. *Experimental Neurology* **2010,** 225, (1), 210-218.

41. Du, J. L.; Murphy, R. M., Characterization of the Interaction of beta-Amyloid with Transthyretin Monomers and Tetramers. *Biochemistry* **2010,** 49, (38), 8276-8289.

42. Aguzzi, A.; O'Connor, T., Protein aggregation diseases: pathogenicity and therapeutic perspectives. *Nature Reviews Drug Discovery* **2010,** 9, (3), 237-248.

43. Bartolini, M.; Andrisano, V., Strategies for the Inhibition of Protein Aggregation in Human Diseases. *Chembiochem* **2010,** 11, (8), 1018-1035.

44. Harrison, R. S.; Sharpe, P. C.; Singh, Y.; Fairlie, D. P., Amyloid peptides and proteins in review. In *Reviews of Physiology, Biochemistry and Pharmacology, Vol 159*, Springer-Verlag Berlin: Berlin, 2007; Vol. 159, pp 1-77.

45. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E., UCSF chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **2004,** 25, (13), 1605-1612.

46. Han, H. Y.; Weinreb, P. H.; Lansbury, P. T., The core Alzheimers peptide NAC forms amyloid fibrils which seed and are seeded by beta-amyloid - is NAC a common trigger or target in neurodegenerative disease. *Chemistry & Biology* **1995,** 2, (3), 163-169.

47. Bhak, G.; Choe, Y. J.; Paik, S. R., Mechanism of amyloidogenesis: nucleation-dependent fibrillation versus double-concerted fibrillation. *Bmb Reports* **2009,** 42, (9), 541-551.

48. Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y. B.; Wriggers, W., Atomic-Level Characterization of the Structural Dynamics of Proteins. *Science* **2010,** 330, (6002), 341-346.

49. Becker, O. M., *Computational biochemistry and biophysics* Marcel Dekker: New York 2001.

50. Adcock, S. A.; McCammon, J. A., Molecular dynamics: Survey of methods for simulating the activity of proteins. *Chemical Reviews* **2006,** 106, (5), 1589-1615.

51. Ciccotti, G. a. H., W. G. , *Molecular dynamics simulation of statistical-mechanical systems*. North-Holland: 1986.

52. Heermann, D. W., *Computer simulation methods*. Springer: 1986.

53. Ponder, J. W.; Case, D. A., Force fields for protein simulations. In *Protein Simulations*, Academic Press Inc: San Diego, 2003; Vol. 66, pp 27-87.

54. Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J., The Amber biomolecular simulation programs. *Journal of Computational Chemistry* **2005,** 26, (16), 1668-1688.

55. Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M., CHARMM: The Biomolecular Simulation Program. *Journal of Computational Chemistry* **2009,** 30, (10), 1545-1614.

56. Van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C., GROMACS: Fast, flexible, and free. *Journal of Computational Chemistry* **2005,** 26, (16), 1701-1718.

57. Verlet, L., Computer experiments on classical fluids I Thermodynamical properties of Lennard Jones molecules. *Physical Review* **1967,** 159, (1), 98-103.

58. Hockney, R. W., The potential calculation and some applications. *Method in Computational Physics* **1970,** 9, 136-211.

59. Swope, W. C.; Andersen, H. C.; Berens, P. H.; Wilson, K. R., A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules application to small water clusters. *Journal of Chemical Physics* **1982,** 76, (1), 637-649.

60. Senn, H. M.; Thiel, W., QM/MM methods for biological systems. In *Atomistic Approaches in Modern Biology: From Quantum Chemistry to Molecular Simulations*, Springer-Verlag Berlin: Berlin, 2007; Vol. 268, pp 173-290.

61. Swope, W. C.; Andersen, H. C.; Berens, P. H.; Wilson, K. R., A computer simulation method for the calculation of equilibrium-constants for the formation of physical clusters of molecules application to small water clusters. *Journal of Chemical Physics* **1982,** 76, (1), 637-649.

62. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of simple potential functions for simulating liquid water. *Journal of Chemical Physics* **1983,** 79, (2), 926-935.

63. Case DA, D. T., Cheatham TE, Simmerling CL, Wang J, Duke RE, Luo R, Walker RC, Zhang W, Merz KM, Roberts B, Wang B, Hayik S, Roitberg A, Seabra G, Kolossváry I, Wong KF, Paesani F, Vanicek J, Liu J, Wu X, Brozell SR, Steinbrecher T, Gohlke H, Cai Q, Ye X, Wang J, Hsieh MJ, Cui G, Roe DR, Mathews DH, Seetin MG, Sagui C, Babin V, Luchko T, Gusarov S, Kovalenko A, Kollman PA (2010) AMBER 11 University of California *AMBER 11*, University of California, San Francisco: San Francisco, 2010.

64. Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C., Numerical-integration of cartesian equations of motion of a system with constraints - molecular-dynamics of n-alkanes. *Journal of Computational Physics* **1977,** 23, (3), 327-341.

65. Humphrey, W.; Dalke, A.; Schulten, K., VMD: Visual molecular dynamics. *Journal of Molecular Graphics* **1996,** 14, (1), 33-38.

66. Sasahara, K.; Naiki, H.; Goto, Y., Kinetically controlled thermal response of beta(2)-microglobulin amyloid fibrils. *Journal of Molecular Biology* **2005,** 352, (3), 700-711.

67. Meersman, F.; Dobson, C. M., Probing the pressure-temperature stability of amyloid fibrils provides new insights into their molecular properties. *Biochimica Et Biophysica Acta-Proteins and Proteomics* **2006,** 1764, (3), 452-460.

68. Mauro, M.; Craparo, E. F.; Podesta, A.; Bulone, D.; Carrotta, R.; Martorana, V.; Tiana, G.; San Biagio, P. L., Kinetics of different processes in human insulin amyloid formation. *Journal of Molecular Biology* **2007,** 366, (1), 258-274.

69. Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A., Development and testing of a general amber force field. *Journal of Computational Chemistry* **2004,** 25, (9), 1157-1174.

70. Frisch, M. J. T., G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb,; M. A.; Cheeseman, J. R. M., J. A., Jr.; Vreven, T.; Kudin,; K. N.; Burant, J. C. M., J. M.; Iyengar, S. S.; Tomasi, J.;; Barone, V. M., B.; Cossi, M.; Scalmani, G.; Rega, N.;; Petersson, G. A. N., H.; Hada, M.; Ehara, M.; Toyota, K.;; Fukuda, R. H., J.; Ishida, M.; Nakajima, T.; Honda, Y.;; Kitao, O. N., H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H.; P.; Cross, J. B. A., C.; Jaramillo, J.; Gomperts, R.; Stratmann,; R. E.; Yazyev, O. A., A. J.; Cammi, R.; Pomelli, C.; Ochterski,; J. W.; Ayala, P. Y. M., K.; Voth, G. A.; Salvador, P.; Dannenberg,; J. J.; Zakrzewski, V. G. D., S.; Daniels, A. D.;; Strain, M. C. F., O.; Malick, D. K.; Rabuck, A. D.;; Raghavachari, K. F., J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A.; G.; Clifford, S. C., J.; Stefanov, B. B.; Liu, G.; Liashenko,; A.; Piskorz, P. K., I.; Martin, R. L.; Fox, D. J.; Keith, T.;; Al-Laham, M. A. P., C. Y.; Nanayakkara, A.; Challacombe, M.;; Gill, P. M. W. J., B.; Chen, W.; Wong, M. W.; Gonzalez, C.;; Pople, J. A. Gaussian.

71. Tatsis, V. A.; Tsoulos, I. G.; Stavrakoudis, A., Molecular Dynamics Simulations of the TSSPSAD Peptide Antigen in Free and Bound with CAMPATH-1H Fab Antibody States: The Importance of the beta-Turn Conformation. *International Journal of Peptide Research and Therapeutics* **2009,** 15, (1), 1-9.

72. Kuzmanic, A.; Zagrovic, B., Determination of Ensemble-Average Pairwise Root Mean-Square Deviation from Experimental B-Factors. *Biophysical Journal* **2010,** 98, (5), 861-871.

73. Sadiq, S. K.; De Fabritiis, G., Explicit solvent dynamics and energetics of HIV-1 protease flap opening and closing. *Proteins-Structure Function and Bioinformatics* **2010,** 78, (14), 2873-2885.

74. Kabsch, W.; Sander, C., Dictionary of protein secondary structure - pattern-recognition of hydrogen-bonded and geometrical features. *Biopolymers* **1983,** 22, (12), 2577-2637.

75. Zheng, J.; Ma, B. Y.; Tsai, C. J.; Nussinov, R., Structural stability and dynamics of an amyloid-forming peptide GNNQQNY from the yeast prion sup-35. *Biophysical Journal* **2006,** 91, (3), 824-833.

76. Schlick, T., Molecular Modeling. An Interdisciplinary Guide. In [Online] Edition, S., Ed. Springer Verlag: New York, 2010.

77. Shao, J. Y.; Tanner, S. W.; Thompson, N.; Cheatham, T. E., Clustering molecular dynamics trajectories: 1. Characterizing the performance of different clustering algorithms. *Journal of Chemical Theory and Computation* **2007,** 3, (6), 2312-2334.

78. Berhanu, W. M.; Masunov, A. E., Can molecular dynamics simulations assist in design of specific inhibitors and imaging agents of amyloid aggregation? Structure, stability and free energy predictions for amyloid oligomers of VQIVYK, MVGGVV and LYQLEN. *J Mol Model, Published online : 21 Dec.2010; DOI 10.1007/s00894-010-0912-4*.

79. De Simone, A.; Esposito, L.; Pedone, C.; Vitagliano, L., Insights into stability and toxicity of amyloid-like oligomers by replica exchange molecular dynamics analyses. *Biophysical Journal* **2008,** 95, (4), 1965-1973.

80. Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S. H.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E., Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. *Accounts of Chemical Research* **2000,** 33, (12), 889-897.

81. Gohlke, H.; Case, D. A., Converging free energy estimates: MM-PB(GB)SA studies on the protein-protein complex Ras-Raf. *Journal of Computational Chemistry* **2004,** 25, (2), 238-250.

82. Chong, L. T.; Pitera, J. W.; Swope, W. C.; Pande, V. S., Comparison of computational approaches for predicting the effects of missense mutations on p53 function. *Journal of Molecular Graphics & Modelling* **2009,** 27, (8), 978-982.

83. Okimoto, N.; Futatsugi, N.; Fuji, H.; Suenaga, A.; Morimoto, G.; Yanai, R.; Ohno, Y.; Narumi, T.; Taiji, M., High-Performance Drug Discovery: Computational Screening by Combining Docking and Molecular Dynamics Simulations. *Plos Computational Biology* **2009,** 5, (10), 13.

84. Ma, B. Y.; Nussinov, R., Simulations as analytical tools to understand protein aggregation and predict amyloid conformation. *Current Opinion in Chemical Biology* **2006,** 10, (5), 445-452.

85. Jiang, P.; Li, W. F.; Shea, J. E.; Mu, Y. G., Resveratrol Inhibits the Formation of Multiple-Layered beta-Sheet Oligomers of the Human Islet Amyloid Polypeptide Segment 22-27. *Biophysical Journal* **2011,** 100, (6), 1550-1558.

86. Berhanu, W. M.; Masunov, A. E., Natural polyphenols as inhibitors of amyloid aggregation. Molecular dynamics study of GNNQQNY heptapeptide decamer. *Biophysical Chemistry* **2010,** 149, (1-2), 12-21.

87. Raman, E. P.; Takeda, T.; Klimov, D. K., Molecular Dynamics Simulations of Ibuprofen Binding to A beta Peptides. *Biophysical Journal* **2009,** 97, (7), 2070-2079.

88. Berhanu, W. M.; Masunov, A. E., Unique example of amyloid aggregates stabilized by main chain H-bond instead of the steric zipper: molecular dynamics study of the amyloidogenic segment of amylin wild-type and mutants. *Journal of Molecular Modeling,Published online: 28; May 2011 DOI 10.1007/s00894-011-1030-7*

89. Selkoe, D. J., Folding proteins in fatal ways. *Nature* **2003,** 426, (6968), 900-904.

90. Hamley, I. W., Peptide fibrillization. *Angewandte Chemie-International Edition* **2007,** 46, (43), 8128-8147.

91. Makin, O. S.; Serpell, L. C., Structures for amyloid fibrils. *Febs Journal* **2005,** 272, (23), 5950-5961.

92. Nelson, R.; Eisenberg, D., Recent atomic models of amyloid fibril structure. *Current Opinion in Structural Biology* **2006,** 16, (2), 260-265.

93. Wiltzius, J. J. W.; Sievers, S. A.; Sawaya, M. R.; Cascio, D.; Popov, D.; Riekel, C.; Eisenberg, D., Atomic structure of the cross-beta spine of islet amyloid polypeptide (amylin). *Protein Science* **2008,** 17, (9), 1467-1474.

94. Ivanova, M. I.; Sievers, S. A.; Sawaya, M. R.; Wall, J. S.; Eisenberg, D., Molecular basis for insulin fibril assembly. *Proceedings of the National Academy of Sciences of the United States of America* **2009,** 106, (45), 18990-18995.

95. Park, J.; Kahng, B.; Hwang, W., Thermodynamic Selection of Steric Zipper Patterns in the Amyloid Cross-beta Spine. *Plos Computational Biology* **2009,** 5, (9), 17.

96. Gsponer, J.; Haberthur, U.; Caflisch, A., The role of side-chain interactions in the early steps of aggregation: Molecular dynamics simulations of an amyloid-forming peptide from the yeast prion Sup35. *Proceedings of the National Academy of Sciences of the United States of America* **2003,** 100, (9), 5154-5159.

97. Esposito, L.; Paladino, A.; Pedone, C.; Vitagliano, L., Insights into structure, stability, and toxicity of monomeric and aggregated polyglutamine models from molecular dynamics simulations. *Biophysical Journal* **2008,** 94, (10), 4031-4040.

98. Esposito, L.; Pedone, C.; Vitagliano, L., Molecular dynamics analyses of cross-beta-spine steric zipper models: beta-sheet twisting and aggregation. *Proceedings of the National Academy of Sciences of the United States of America* **2006,** 103, (31), 11533-11538.

99. Wei, G. H.; Song, W.; Derreumaux, P.; Mousseau, N., Self-assembly of amyloid-forming peptides by molecular dynamics simulations. *Frontiers in Bioscience* **2008,** 13, 5681-5692.

100. Wu, C.; Wang, Z. X.; Lei, H. X.; Zhang, W.; Duan, Y., Dual binding modes of Congo red to amyloid protofibril surface observed in molecular dynamics simulations. *Journal of the American Chemical Society* **2007,** 129, (5), 1225-1232.

101. Berryman, J. T.; Radford, S. E.; Harris, S. A., Thermodynamic Description of Polymorphism in Q- and N-Rich Peptide Aggregates Revealed by Atomistic Simulation. *Biophysical Journal* **2009,** 97, (1), 1-11.

102. Chang, L. K.; Zhao, J. H.; Liu, H. L.; Liu, K. T.; Chen, J. T.; Tsai, W. B.; Ho, Y., Molecular Dynamics Simulations to Investigate the Structural Stability and Aggregation Behavior of the GGVVIA Oligomers Derived from Amyloid beta Peptide. *Journal of Biomolecular Structure & Dynamics* **2009,** 26, (6), 731-740.

103. Nerelius, C.; Johansson, J.; Sandegren, A., Amyloid beta-peptide aggregation. What does it result in and how can it be prevented? *Frontiers in Bioscience* **2009,** 14, 1716-U3856.

104. Rauk, A., Why is the amyloid beta peptide of Alzheimer's disease neurotoxic? *Dalton Transactions* **2008**, (10), 1273-1282.

105. Rauk, A., The chemistry of Alzheimer's disease. *Chemical Society Reviews* **2009,** 38, (9), 2698-2715.

106. Teplow, D. B.; Lazo, N. D.; Bitan, G.; Bernstein, S.; Wyttenbach, T.; Bowers, M. T.; Baumketner, A.; Shea, J. E.; Urbanc, B.; Cruz, L.; Borreguero, J.; Stanley, H. E., Elucidating amyloid beta-protein folding and assembly: A multidisciplinary approach. *Accounts of Chemical Research* **2006,** 39, (9), 635-645.

107. Berriman, J.; Serpell, L. C.; Oberg, K. A.; Fink, A. L.; Goedert, M.; Crowther, R. A., Tau filaments from human brain and from in vitro assembly of recombinant protein show cross-beta structure. *Proceedings of the National Academy of Sciences of the United States of America* **2003,** 100, (15), 9034-9038.

108. Devlin, G. L.; Knowles, T. P. J.; Squires, A.; McCammon, M. G.; Gras, S. L.; Nilsson, M. R.; Robinson, C. V.; Dobson, C. M.; MacPhee, C. E., The component polypeptide chains of bovine insulin nucleate or inhibit aggregation of the parent protein in a conformation-dependent manner. *Journal of Molecular Biology* **2006,** 360, (2), 497-509.

109. Hong, D. P.; Fink, A. L., Independent heterologous fibrillation of insulin and its B-chain peptide. *Biochemistry* **2005,** 44, (50), 16701-16709.

110. Wilhelm, K. R.; Yanamandra, K.; Gruden, M. A.; Zamotin, V.; Malisauskas, M.; Casaite, V.; Darinskas, A.; Forsgren, L.; Morozova-Roche, L. A., Immune reactivity towards insulin, its amyloid and protein S100B in blood sera of Parkinson's disease patients. *European Journal of Neurology* **2007,** 14, (3), 327-334.

111. Brange, J.; Andersen, L.; Laursen, E. D.; Meyn, G.; Rasmussen, E., Toward understanding insulin fibrillation. *Journal of Pharmaceutical Sciences* **1997,** 86, (5), 517-525.

112. Ahmad, A.; Uversky, V. N.; Hong, D.; Fink, A. L., Early events in the fibrillation of monomeric insulin. *Journal of Biological Chemistry* **2005,** 280, (52), 42669-42675.

113. Zhang, Z. Q.; Chen, H.; Bai, H. J.; Lai, L. H., Molecular dynamics Simulations on the oligomer-formation process of the GNNQQNY peptide from yeast prion protein sup35. *Biophysical Journal* **2007,** 93, (5), 1484-1492.

114. Xu, W. X.; Ping, J.; Li, W. F.; Mu, Y. G., Assembly dynamics of two-beta sheets revealed by molecular dynamics simulations. *Journal of Chemical Physics* **2009,** 130, (16), 8.

115. Kayed, R.; Pensalfini, A.; Margol, L.; Sokolov, Y.; Sarsoza, F.; Head, E.; Hall, J.; Glabe, C., Annular Protofibrils Are a Structurally and Functionally Distinct Type of Amyloid Oligomer. *Journal of Biological Chemistry* **2009,** 284, (7), 4230-4237.

116. Vitagliano, L.; Stanzione, F.; De Simone, A.; Esposito, L., Dynamics and Stability of Amyloid-Like Steric Zipper Assemblies with Hydrophobic Dry Interfaces. *Biopolymers* **2009,** 91, (12), 1161-1171.

117. http://www.doe-mbi.ucla.edu/~sawaya/chime/xtalpept (September, 2009),

118. Case DA, D. T., Cheatham TE, Simmerling CL, Wang J, Duke RE, Luo R, Walker RC, Zhang W, Merz KM, Roberts B, Wang B, Hayik S, Roitberg A, Seabra G, Kolossváry I, Wong KF, Paesani F, Vanicek J, Liu J, Wu X, Brozell SR, Steinbrecher T, Gohlke H, Cai Q, Ye X, Wang J, Hsieh MJ, Cui G, Roe DR, Mathews DH, Seetin MG, Sagui C, Babin V, Luchko T, Gusarov S, Kovalenko A, Kollman PA *Amber 10*, University of California, : San Francisco,, 2008.

119. Chong, L. T.; Duan, Y.; Wang, L.; Massova, I.; Kollman, P. A., Molecular dynamics and free-energy calculations applied to affinity maturation in antibody 48G7. *Proceedings of the National Academy of Sciences of the United States of America* **1999,** 96, (25), 14330-14335.

120. Hardy, J.; Selkoe, D. J., Medicine - The amyloid hypothesis of Alzheimer's disease: Progress and problems on the road to therapeutics. *Science* **2002,** 297, (5580), 353-356.

121. Lesne, S.; Koh, M. T.; Kotilinek, L.; Kayed, R.; Glabe, C. G.; Yang, A.; Gallagher, M.; Ashe, K. H., A specific amyloid-beta protein assembly in the brain impairs memory. *Nature* **2006,** 440, (7082), 352-357.

122. Sciarretta, K. L.; Gordon, D. J.; Meredith, S. C., Peptide-based inhibitors of amyloid assembly. In *Amyloid, Prions, and Other Protein Aggregates, Part C*, Elsevier Academic Press Inc: San Diego, 2006; Vol. 413, pp 273-312.

123. Bogan, A. A.; Thorn, K. S., Anatomy of hot spots in protein interfaces. *Journal of Molecular Biology* **1998,** 280, (1), 1-9.

124. Haydar, S. N.; Yun, H. E. D.; Staal, R. G. W.; Hirst, W. D., Small-Molecule Protein-Protein Interaction Inhibitors as Therapeutic Agents for Neurodegenerative Diseases: Recent Progress and

Future Directions. In *Annual Reports in Medicinal Chemistry, Vol 44*, Elsevier Academic Press Inc: San Diego, 2009; Vol. 44, pp 51-69.

125. Blazer, L. L.; Neubig, R. R., Small Molecule Protein-Protein Interaction Inhibitors as CNS Therapeutic Agents: Current Progress and Future Hurdles. *Neuropsychopharmacology* **2009,** 34, (1), 126-141.

126. Otto, M.; Lewczuk, P.; Wiltfang, J., Neurochemical approaches of cerebrospinal fluid diagnostics in neurodegenerative diseases. *Methods* **2008,** 44, (4), 289-298.

127. Tenidis, K.; Waldner, M.; Bernhagen, J.; Fischle, W.; Bergmann, M.; Weber, M.; Merkle, M. L.; Voelter, W.; Brunner, H.; Kapurniotu, A., Identification of a penta- and hexapeptide of islet amyloid polypeptide (IAPP) with amyloidogenic and cytotoxic properties. *Journal of Molecular Biology* **2000,** 295, (4), 1055-1071.

128. Porat, Y.; Mazor, Y.; Efrat, S.; Gazit, E., Inhibition of islet amyloid polypeptide fibril formation: A potential role for heteroaromatic interactions. *Biochemistry* **2004,** 43, (45), 14454-14462.

129. Sato, T.; Kienlen-Campard, P.; Ahmed, M.; Liu, W.; Li, H. L.; Elliott, J. I.; Aimoto, S.; Constantinescu, S. N.; Octave, J. N.; Smith, S. O., Inhibitors of amyloid toxicity based on beta-sheet packing of A beta 40 and A beta 42. *Biochemistry* **2006,** 45, (17), 5503-5516.

130. Takahashi, T.; Ohta, K.; Mihara, H., Rational design of amyloid beta peptide-binding proteins: Pseudo-A beta beta-sheet surface presented in green fluorescent protein binds tightly and preferentially to structured A beta. *Proteins-Structure Function and Bioinformatics* **2010,** 78, (2), 336-347.

131. Kim, Y. S.; Lee, J. H.; Ryu, J.; Kim, D. J., Multivalent & Multifunctional Ligands to beta-Amyloid. *Current Pharmaceutical Design* **2009,** 15, (6), 637-658.

132. Fandrich, M.; Meinhardt, J.; Grigorieff, N., Structural polymorphism of Alzheimer A beta and other amyloid fibrils. *Prion* **2009,** 3, (2), 89-93.

133. Miller, Y.; Ma, B.; Nussinov, R., Polymorphism in Alzheimer A beta Amyloid Organization Reflects Conformational Selection in a Rugged Energy Landscape. *Chemical Reviews* **2010,** 110, (8), 4820-4838.

134. Apostol, M. I.; Sawaya, M. R.; Cascio, D.; Eisenberg, D., Crystallographic Studies of Prion Protein (PrP) Segments Suggest How Structural Changes Encoded by Polymorphism at Residue 129 Modulate Susceptibility to Human Prion Disease. *Journal of Biological Chemistry* **2010,** 285, (39), 29671-29675.

135. Andersen, C. B.; Hicks, M. R.; Vetri, V.; Vandahl, B.; Rahbek-Nielsen, H.; Thogersen, H.; Thogersen, I. B.; Enghild, J. J.; Serpell, L. C.; Rischel, C.; Otzen, D. E., Glucagon Fibril Polymorphism Reflects Differences in Protofilament Backbone Structure. *Journal of Molecular Biology* **2010,** 397, (4), 932-946.

136. Calamai, M.; Chiti, F.; Dobson, C. M., Amyloid fibril formation can proceed from different conformations of a partially unfolded protein. *Biophysical Journal* **2005,** 89, (6), 4201-4210.

137. Chiti, F.; Dobson, C. M., Protein misfolding, functional amyloid, and human disease. *Annual Review of Biochemistry* **2006,** 75, 333-366.

138. Jones, E. M.; Surewicz, W. K., Fibril conformation as the basis of species- and strain-dependent seeding specificity of mammalian prion amyloids. *Cell* **2005,** 121, (1), 63-72.

139. Petkova, A. T.; Leapman, R. D.; Guo, Z. H.; Yau, W. M.; Mattson, M. P.; Tycko, R., Self-propagating, molecular-level polymorphism in Alzheimer's beta-amyloid fibrils. *Science* **2005,** 307, (5707), 262-265.

140. Chafekar, S. M.; Baas, F.; Scheper, W., Oligomer-specific A beta toxicity in cell models is mediated by selective uptake. *Biochimica Et Biophysica Acta-Molecular Basis of Disease* **2008,** 1782, (9), 523-531.

141. Deshpande, A.; Mina, E.; Glabe, C.; Busciglio, J., Different conformations of amyloid beta induce neurotoxicity by distinct mechanisms in human cortical neurons. *Journal of Neuroscience* **2006,** 26, (22), 6011-6018.

142. Iijima, K.; Chiang, H. C.; Hearn, S. A.; Hakker, I.; Gatt, A.; Shenton, C.; Granger, L.; Leung, A.; Iijima-Ando, K.; Zhong, Y., A beta 42 Mutants with Different Aggregation Profiles Induce Distinct Pathologies in Drosophila. *Plos One* **2008,** 3, (2), 8.

143. Poirier, M. A.; Li, H. L.; Macosko, J.; Cai, S. W.; Amzel, M.; Ross, C. A., Huntingtin spheroids and protofibrils as precursors in polyglutamine fibrilization. *Journal of Biological Chemistry* **2002,** 277, (43), 41032-41037.

144. Gosal, W. S.; Morten, I. J.; Hewitt, E. W.; Smith, D. A.; Thomson, N. H.; Radford, S. E., Competing pathways determine fibril morphology in the self-assembly of beta(2)-microglobulin into amyloid. *Journal of Molecular Biology* **2005,** 351, (4), 850-864.

145. Goldsbury, C.; Frey, P.; Olivieri, V.; Aebi, U.; Muller, S. A., Multiple assembly pathways underlie amyloid-beta fibril polymorphisms. *Journal of Molecular Biology* **2005,** 352, (2), 282-298.

146. Deechongkit, S.; Nguyen, H.; Powers, E. T.; Dawson, P. E.; Gruebele, M.; Kelly, J. W., Context-dependent contributions of backbone hydrogen bonding to beta-sheet folding energetics. *Nature* **2004,** 430, (6995), 101-105.

147. Tsemekhman, K.; Goldschmidt, L.; Eisenberg, D.; Baker, D., Cooperative hydrogen bonding in amyloid formation. *Protein Science* **2007,** 16, (4), 761-764.

148. Papaleo, E.; Invernizzi, G., Conformational Diseases: Structural Studies of Aggregation of Polyglutamine Proteins. *Current Computer-Aided Drug Design* **2011,** 7, (1), 23-43.

149. Berhanu, W. M.; Masunov, A. E., Molecular dynamic simulation of wildtype and mutants of the polymorphic amyloid NNQNTF segments of elk prion: structural stability and thermodynamic of association. *Biopolymers* **2011,** 95, 573-589.

150. De Simone, A.; Pedone, C.; Vitagliano, L., Structure, dynamics, and stability of assemblies of the human prion fragment SNQNNF. *Biochemical and Biophysical Research Communications* **2008,** 366, (3), 800-806.

151. Wu, C.; Bowers, M. T.; Shea, J. E., Molecular Structures of Quiescently Grown and Brain-Derived Polymorphic Fibrils of the Alzheimer Amyloid A beta(9-40) Peptide: A Comparison to Agitated Fibrils. *Plos Computational Biology* **2010,** 6, (3), 12.

152. Paparcone, R.; Sanchez, J.; Buehler, M. J., Comparative Study of Polymorphous Alzheimer's A beta(1-40) Amyloid Nanofibrils and Microfibers. *Journal of Computational and Theoretical Nanoscience* **2010,** 7, (7), 1279-1286.

153. Ma, B. Y.; Nussinov, R., Polymorphic C-terminal beta-Sheet Interactions Determine the Formation of Fibril or Amyloid beta-derived Diffusible Ligand-like Globulomer for the Alzheimer A beta 42 Dodecamer. *Journal of Biological Chemistry* 285, (47), 37102-37110.

154. Berryman, J. T.; Radford, S. E.; Harris, S. A., Systematic Examination of Polymorphism in Amyloid Fibrils by Molecular-Dynamics Simulation. *Biophysical Journal* **2011,** 100, (9), 2234-2242.

155. Humphrey, W.; Dalke, A.; Schulten, K., VMD: Visual molecular dynamics. *Journal of Molecular Graphics* **1996,** 14, (1), 33-&.

156. Massova, I.; Kollman, P. A., Computational alanine scanning to probe protein-protein interactions: A novel approach to evaluate binding free energies. *Journal of the American Chemical Society* **1999,** 121, (36), 8133-8143.

157. Buchete, N. V.; Hummer, G., Structure and dynamics of parallel beta-sheets, hydrophobic core, and loops in Alzheimer's A beta fibrils. *Biophysical Journal* **2007,** 92, (9), 3032-3039.

158. Huet, A.; Derreumaux, P., Impact of the mutation A21G (Flemish variant) on Alzheimer's beta-amyloid dimers by molecular dynamics simulations. *Biophysical Journal* **2006,** 91, (10), 3829-3840.

159. Miller, Y.; Ma, B. Y.; Nussinov, R., The Unique Alzheimer's beta-Amyloid Triangular Fibril Has a Cavity along the Fibril Axis under Physiological Conditions. *Journal of the American Chemical Society* 133, (8), 2742-2748.

160. Sievers SA, K. J., Chang HW, Zhao A, Jiang L, Zirafi O, Stevens JT, MÃ¼nch J, Baker D, Eisenberg D., Structure-based design of non-natural amino-acid inhibitors of amyloid fibril formation. *Nature Structural & Molecular Biology* **2011**, 1-7.

161. Paparcone, R.; Pires, M. A.; Buehler, M. J., Mutations Alter the Geometry and Mechanical Properties of Alzheimer's A beta(1-40) Amyloid Fibrils. *Biochemistry* **2010,** 49, (41), 8967-8977.

162. Paparcone, R.; Sanchez, J.; Buehler, M. J., Comparative Study of Polymorphous Alzheimer's A beta(1-40) Amyloid Nanofibrils and Microfibers. *Journal of Computational and Theoretical Nanoscience* 7, (7), 1279-1286.

163. Jager, M.; Dendle, M.; Kelly, J. W., Sequence determinants of thermodynamic stability in a WW domain-An all-beta-sheet protein. *Protein Science* **2009,** 18, (8), 1806-1813.

164. Sievers, S. A. K., J.; Chang, H. W.; Zhao, A.; Jiang, L.; Zirafi, O.; Stevens, J. T.; Munch, J.; Baker, D.; Eisenberg, D. , Structure-based design of non-natural amino-acid inhibitors of amyloid fibril formation. *Nature* **2011,** 475, 96-100.

165. Nystrom, F. H.; Quon, M. J., Insulin signalling: Metabolic pathways and mechanisms for specificity. *Cellular Signalling* **1999,** 11, (8), 563-574.

166. Ottensmeyer, F. P.; Beniac, D. R.; Luo, R. Z. T.; Yip, C. C., Mechanism of transmembrane signaling: Insulin binding and the insulin receptor. *Biochemistry* **2000,** 39, (40), 12103-12112.

167. Wild, S.; Roglic, G.; Green, A.; Sicree, R.; King, H., Global prevalence of diabetes - Estimates for the year 2000 and projections for 2030. *Diabetes Care* **2004,** 27, (5), 1047-1053.

168. Rewers, M., Why do people with diabetes die too soon? *Diabetes Care* **2008,** 31, (4), 830-832.

169. Storkel, S.; Schneider, H. M.; Muntefering, H.; Kashiwagi, S., Iatrogenic, insulin-dependent, local amyloidosis. *Laboratory Investigation* **1983,** 48, (1), 108-111.

170. 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, (3), 158-161.

171. Sluzky, V.; Klibanov, A. M.; Langer, R., Mechanism of insulin aggregation and stabilization in agitated aqueous-solutions. *Biotechnology and Bioengineering* **1992,** 40, (8), 895-903.

172. Grillo, A. O.; Edwards, K. L. T.; Kashi, R. S.; Shipley, K. M.; Hu, L.; Besman, M. J.; Middaugh, C. R., Conformational origin of the aggregation of recombinant human factor VIII. *Biochemistry* **2001,** 40, (2), 586-595.

173. Valla, V., Therapeutics of diabetes mellitus: focus on insulin analogues and insulin pumps. *Experimental Diabetes Research* **2010,** 2010, 1-14.

174. Valla, V., Therapeutics of Diabetes Mellitus: Focus on Insulin Analogues and Insulin Pumps. *Experimental Diabetes Research* **2010**, 14.

175. Bell, D. S. H., Insulin therapy in diabetes mellitus - How can the currently available injectable insulins be most prudently and efficaciously utilised? *Drugs* **2007,** 67, (13), 1813-1827.

176. Mayer, J. P.; Zhang, F.; DiMarchi, R. D., Insulin structure and function. *Biopolymers* **2007,** 88, (5), 687-713.

177. Geddes, A. J.; Parker, K. D.; Atkins, E. D. T.; Beighton, E., Cross-beta conformation in proteins. *Journal of Molecular Biology* **1968,** 32, (2), 343-358

178. 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, (28), 8397-8409.

179. Brange, J.; Dodson, G. G.; Edwards, D. J.; Holden, P. H.; Whittingham, J. L., A model of insulin fibrils derived from the x-ray crystal structure of a monomeric insulin (despentapeptide insulin). *Proteins-Structure Function and Genetics* **1997,** 27, (4), 507-516.

180. Ivanova, M. I.; Thompson, M. J.; Eisenberg, D., A systematic screen of beta(2)-microglobulin and insulin for amyloid-like segments. *Proceedings of the National Academy of Sciences of the United States of America* **2006,** 103, (11), 4079-4082.

181. Jimenez, J. L.; Nettleton, E. J.; Bouchard, M.; Robinson, C. V.; Dobson, C. M.; Saibil, H. R., The protofilament structure of insulin amyloid fibrils. *Proceedings of the National Academy of Sciences of the United States of America* **2002,** 99, (14), 9196-9201.

182. Vestergaard, B.; Groenning, M.; Roessle, M.; Kastrup, J. S.; van de Weert, M.; Flink, J. M.; Frokjaer, S.; Gajhede, M.; Svergun, D. I., A helical structural nucleus is the primary elongating unit of insulin amyloid fibrils. *Plos Biology* **2007,** 5, (5), 1089-1097.

183. Choi, J. H.; May, B. C. H.; Wille, H.; Cohen, F. E., Molecular Modeling of the Misfolded Insulin Subunit and Amyloid Fibril. *Biophysical Journal* **2009,** 97, (12), 3187-3195.

184. Horn, A. H. C.; Sticht, H., Amyloid-beta 42 Oligomer Structures from Fibrils: A Systematic Molecular Dynamics Study. *Journal of Physical Chemistry B* **2010,** 114, (6), 2219-2226.

185. Tsai, H. H.; Reches, M.; Tsai, C. J.; Gunasekaran, K.; Gazit, E.; Nussinov, R., Energy landscape of amyloidogenic peptide oligomerization by parallel-tempering molecular dynamics simulation: Significant

role of Asn ladder. *Proceedings of the National Academy of Sciences of the United States of America* **2005,** 102, (23), 8174-8179.

186. Mark, A. E.; Berendsen, H. J. C.; Vangunsteren, W. F., Conformational flexibility of aqueous monomeric and dimeric insulin - a molecular-dynamics study. *Biochemistry* **1991,** 30, (45), 10866-10872.

187. Zoete, V.; Meuwly, M.; Karplus, M., Investigation of glucose binding sites on insulin. *Proteins-Structure Function and Bioinformatics* **2004,** 55, (3), 568-581.

188. Zoete, V.; Meuwly, M., Importance of individual side chains for the stability of a protein fold: Computational alanine scanning of the insulin monomer. *Journal of Computational Chemistry* **2006,** 27, (15), 1843-1857.

189. Zoete, V.; Meuwly, M.; Karplus, M., Study of the insulin dimerization: Binding free energy calculations and per-residue free energy decomposition. *Proteins-Structure Function and Bioinformatics* **2005,** 61, (1), 79-93.

190. Falconi, M.; Cambria, M. T.; Cambria, A.; Desideri, A., Structure and stability of the insulin dimer investigated by molecular dynamics simulation. *Journal of Biomolecular Structure & Dynamics* **2001,** 18, (5), 761-772.

191. Lu, B. Z.; Chen, W. Z.; Wang, C. X.; Xu, X. J., Protein molecular dynamics with electrostatic force entirely determined by a single Poisson-Boltzmann calculation. *Proteins-Structure Function and Genetics* **2002,** 48, (3), 497-504.

192. Keller, B.; Daura, X.; van Gunsteren, W. F., Comparing geometric and kinetic cluster algorithms for molecular simulation data. *Journal of Chemical Physics* 132, (7), 16.

193. Takeda, T.; Klimov, D. K., Probing Energetics of A beta Fibril Elongation by Molecular Dynamics Simulations. *Biophysical Journal* **2009,** 96, (11), 4428-4437.

194. Soto, C.; Estrada, L.; Castilla, J., Amyloids, prions and the inherent infectious nature of misfolded protein aggregates. *Trends in Biochemical Sciences* **2006,** 31, (3), 150-155.

195. Padrick, S. B.; Miranker, A. D., Islet amyloid: Phase partitioning and secondary nucleation are central to the mechanism of fibrillogenesis. *Biochemistry* **2002,** 41, (14), 4694-4703.

196. Ono, K.; Condron, M. M.; Teplow, D. B., Structure-neurotoxicity relationships of amyloid beta-protein oligomers. *Proceedings of the National Academy of Sciences of the United States of America* **2009,** 106, (35), 14745-14750.

197. Sorci, M.; Grassucci, R. A.; Hahn, I.; Frank, J.; Belfort, G., Time-dependent insulin oligomer reaction pathway prior to fibril formation: Cooling and seeding. *Proteins-Structure Function and Bioinformatics* **2009,** 77, (1), 62-73.

198. Heldt, C. L.; Sorci, M.; Posada, D.; Hirsa, A.; Belfort, G., Detection and Reduction of Microaggregates in Insulin Preparations. *Biotechnology and Bioengineering* **2010,** 108, (1), 237-241.

199. Manno, M.; Giacomazza, D.; Newman, J.; Martorana, V.; San Biagio, P. L., Amyloid Gels: Precocious Appearance of Elastic Properties during the Formation of an Insulin Fibrillar Network. *Langmuir* **2009,** 26, (3), 1424-1426.

200. Fodera, V.; Cataldo, S.; Librizzi, F.; Pignataro, B.; Spiccia, P.; Leone, M., Self-Organization Pathways and Spatial Heterogeneity in Insulin Amyloid Fibril Formation. *Journal of Physical Chemistry B* **2009,** 113, (31), 10830-10837.

201. Xue, W. F.; Homans, S. W.; Radford, S. E., Systematic analysis of nucleation-dependent polymerization reveals new insights into the mechanism of amyloid self-assembly. *Proceedings of the National Academy of Sciences of the United States of America* **2008,** 105, (26), 8926-8931.

202. Fawzi, N. L.; Okabe, Y.; Yap, E. H.; Head-Gordon, T., Determining the critical nucleus and mechanism of fibril elongation of the Alzheimer's A beta(1-40) peptide. *Journal of Molecular Biology* **2007,** 365, (2), 535-550.

203. Fawzi, N. L.; Kohlstedt, K. L.; Okabe, Y.; Head-Gordon, T., Protofibril assemblies of the arctic, dutch, and flemish mutants of the Alzheimer's A beta(1-40) peptide. *Biophysical Journal* **2008,** 94, (6), 2007-2016.

204. Chisato Nishiuraa, K. T., Katsuhiko Minouraa, Miho Sumidab, Taizo Taniguchib, Koji Tomooa, and Toshimasa Ishidaa, Importance of Tyr310 residue in the third repeat of microtubule binding domain for filament formation of tau protein. *J Biochem* **2010,** 147, 405-414.

205. M. von Bergen, P. F., J. Biernat, J. Heberle, E.-M. Mandelkow, and E. Mandelkow, Assembly of τ protein into Alzheimer paired helical filaments depends on a local sequence motif (306VQIVYK311) forming β structure. *Proc Natl Acad sci* **2000,** 97, 5129-5134.

206. Landau M, S. M., Faull KF, Laganowsky A, Jiang L, Sievers SA, Liu J, Barrio JR, Eisenberg D., Towards a Pharmacophore for Amyloid. *PLoS Biol.* **2011,** 9(6):e1001080.

207. Porat Y, A. A., Gazit E Inhibition of amyloid fibril formation by polyphenols: structural similarity and aromatic interactions as a common inhibition mechanism. *Chem Biol Drug Des* **2006,** 67, 27–37.

208. Dobson, C. M., Protein folding and misfolding. *Nature* **2003,** 426, (6968), 884-890.

209. Stefani, M.; Dobson, C. M., Protein aggregation and aggregate toxicity: new insights into protein folding, misfolding diseases and biological evolution. *Journal of Molecular Medicine-Jmm* **2003,** 81, (11), 678-699.

210. Haataja, L.; Gurlo, T.; Huang, C. J.; Butler, P. C., Islet amyloid in type 2 diabetes, and the toxic oligomer hypothesis. *Endocrine Reviews* **2008,** 29, (3), 303-316.

211. Bucciantini, M.; Calloni, G.; Chiti, F.; Formigli, L.; Nosi, D.; Dobson, C. M.; Stefani, M., Prefibrillar amyloid protein aggregates share common features of cytotoxicity. *Journal of Biological Chemistry* **2004,** 279, (30), 31374-31382.

212. Butterfield, D. A., Amyloid beta-peptide (1-42)-induced oxidative stress and neurotoxicity: Implications for neurodegeneration in Alzheimer's disease brain. A review. *Free Radical Research* **2002,** 36, (12), 1307-1313.

213. Zhu, M.; Rajamani, S.; Kaylor, J.; Han, S.; Zhou, F. M.; Fink, A. L., The flavonoid baicalein inhibits fibrillation of alpha-synuclein and disaggregates existing fibrils. *Journal of Biological Chemistry* **2004,** 279, (26), 26846-26857.

214. Zhu, J. T. T.; Choi, R. C. Y.; Chu, G. K. Y.; Cheung, A. W. H.; Gao, Q. T.; Li, J.; Jiang, Z. Y.; Dong, T. T. X.; Tsim, K. W. K., Flavonoids possess neuroprotective effects on cultured pheochromocytoma PC12 cells: A comparison of different flavonoids in activating estrogenic effect and in preventing beta-amyloid-induced cell death. *Journal of Agricultural and Food Chemistry* **2007,** 55, (6), 2438-2445.

215. Akaishi, T.; Morimoto, T.; Shibao, M.; Watanabe, S.; Sakai-Kato, K.; Utsunomiya-Tate, N.; Abe, K., Structural requirements for the flavonoid fisetin in inhibiting fibril formation of amyloid beta protein. *Neuroscience Letters* **2008,** 444, (3), 280-285.

216. Kim, H.; Park, B. S.; Lee, K. G.; Choi, C. Y.; Jang, S. S.; Kim, Y. H.; Lee, S. E., Effects of naturally occurring compounds on fibril formation and oxidative stress of beta-amyloid. *Journal of Agricultural and Food Chemistry* **2005,** 53, (22), 8537-8541.

217. Masuda, M.; Suzuki, N.; Taniguchi, S.; Oikawa, T.; Nonaka, T.; Iwatsubo, T.; Hisanaga, S.; Goedert, M.; Hasegawa, M., Small molecule inhibitors of alpha-synuclein filament assembly. *Biochemistry* **2006,** 45, (19), 6085-6094.

218. Shoval, H.; Lichtenberg, D.; Gazit, E., The molecular mechanisms of the anti-amyloid effects of phenols. *Amyloid-Journal of Protein Folding Disorders* **2007,** 14, (1), 73-87.

219. Porat, Y.; Abramowitz, A.; Gazit, E., Inhibition of amyloid fibril formation by polyphenols: Structural similarity and aromatic interactions as a common inhibition mechanism. *Chemical Biology & Drug Design* **2006,** 67, (1), 27-37.

220. He, J.; Xing, Y. F.; Huang, B.; Zhang, Y. Z.; Zeng, C. M., Tea Catechins Induce the Conversion of Preformed Lysozyme Amyloid Fibrils to Amorphous Aggregates. *Journal of Agricultural and Food Chemistry* **2009,** 57, (23), 11391-11396.

221. Ho, L.; Pasinetti, G. M., Polyphenolic compounds for treating neurodegenerative disorders involving protein misfolding. *Expert Review of Proteomics* **2010,** 7, (4), 579-589.

222. Ono, K.; Yamada, M., Antioxidant compounds have potent anti-fibrillogenic and fibril-destabilizing effects for alpha-synuclein fibrils in vitro. *Journal of Neurochemistry* **2006,** 97, (1), 105-115.

223. Bastianetto, S.; Krantic, S.; Quirion, R., Polyphenols as potential inhibitors of amyloid aggregation and toxicity: Possible significance to Alzheimer's disease. *Mini-Reviews in Medicinal Chemistry* **2008,** 8, (5), 429-435.

224. Wu, C.; Wang, Z. X.; Lei, H. X.; Duan, Y.; Bowers, M. T.; Shea, J. E., The Binding of Thioflavin T and Its Neutral Analog BTA-1 to Protofibrils of the Alzheimer's Disease A beta(16-22) Peptide Probed by Molecular Dynamics Simulations. *Journal of Molecular Biology* **2008,** 384, (3), 718-729.

225. Convertino, M.; Pellarin, R.; Catto, M.; Carotti, A.; Caflisch, A., 9,10-Anthraquinone hinders beta-aggregation: How does a small molecule interfere with A beta-peptide amyloid fibrillation? *Protein Science* **2009,** 18, (4), 792-800.

226. Takeda, T.; Chang, W. L. E.; Raman, E. P.; Klimov, D. K., Binding of nonsteroidal anti-inflammatory drugs to A beta fibril. *Proteins-Structure Function and Bioinformatics* **2010,** 78, (13), 2849-2860.

227. Liu, F. F.; Ji, L.; Dong, X. Y.; Sun, Y., Molecular Insight into the Inhibition Effect of Trehalose on the Nucleation and Elongation of Amyloid beta-Peptide Oligomers. *Journal of Physical Chemistry B* **2009,** 113, (32), 11320-11329.

228. Bruce, N. J.; Chen, D. L.; Dastidar, S. G.; Marks, G. E.; Schein, C. H.; Bryce, R. A., Molecular dynamics simulations of A beta fibril interactions with beta-sheet breaker peptides. *Peptides* **2010,** 31, (11), 2100-2108.

229. Lemkul, J. A.; Bevan, D. R., Destabilizing Alzheimer's A beta(42) Protofibrils with Morin: Mechanistic Insights from Molecular Dynamics Simulations. *Biochemistry* **2010,** 49, (18), 3935-3946.

230. Ho, L.; Yemul, S.; Wang, J.; Pasinettia, G. M., Grape Seed Polyphenolic Extract as a Potential Novel Therapeutic Agent in Tauopathies. *Journal of Alzheimers Disease* **2009,** 16, (2), 433-439.

231. Taniguchi, S.; Suzuki, N.; Masuda, M.; Hisanaga, S.; Iwatsubo, T.; Goedert, M.; Hasegawa, M., Inhibition of heparin-induced tau filament formation by phenothiazines, polyphenols, and porphyrins. *Journal of Biological Chemistry* **2005,** 280, (9), 7614-7623.

232. Wang, J.; Santa-Maria, I.; Ho, L.; Ksiezak-Reding, H.; Ono, K.; Teplow, D. B.; Pasinetti, G. M., Grape Derived Polyphenols Attenuate Tau Neuropathology in a Mouse Model of Alzheimer's Disease. *Journal of Alzheimers Disease* **2010,** 22, (2), 653-661.

233. Yang, F. S.; Lim, G. P.; Begum, A. N.; Ubeda, O. J.; Simmons, M. R.; Ambegaokar, S. S.; Chen, P. P.; Kayed, R.; Glabe, C. G.; Frautschy, S. A.; Cole, G. M., Curcumin inhibits formation of amyloid beta oligomers and fibrils, binds plaques, and reduces amyloid in vivo. *Journal of Biological Chemistry* **2005,** 280, (7), 5892-5901.

234. Amijee, H.; Scopes, D. I. C., The Quest for Small Molecules as Amyloid Inhibiting Therapies for Alzheimer's Disease. *Journal of Alzheimers Disease* **2009,** 17, (1), 33-47.

235. Ho, L.; Pasinetti, G. M., Polyphenolic compounds for treating neurodegenerative disorders involving protein misfolding. *Expert Review of Proteomics* 7, (4), 579-589.

236. Meng, F. L.; Abedini, A.; Plesner, A.; Verchere, C. B.; Raleigh, D. P., The Flavanol (-)-Epigallocatechin 3-Gallate Inhibits Amyloid Formation by Islet Amyloid Polypeptide, Disaggregates Amyloid Fibrils, and Protects Cultured Cells against IAPP-Induced Toxicity. *Biochemistry* 49, (37), 8127-8133.

237. Ladiwala, A. R. A. L., J. C.; Bale, S. S.; Marcelino-Cruz, A. M.; Bhattacharya, M.; Dordick, J. S.; Tessier, P. M., , Resveratrol Selectively Remodels Soluble Oligomers and Fibrils of Amyloid A beta into Off-pathway Conformers. *Journal of Biological Chemistry* **2010,** 285, (31), 24228-24237.

238. Bulic B, P. M., Schmidt B, Mandelkow EM, Waldmann H, Mandelkow E., Development of tau aggregation inhibitors for Alzheimer's disease. *Angew Chem Int Ed Engl.* **2009,** 48, 1740-1752.

239. Fabiola, F.; Bertram, R.; Korostelev, A.; Chapman, M. S., An improved hydrogen bond potential: Impact on medium resolution protein structures. *Protein Science* **2002,** 11, (6), 1415-1423.

240. Gilson, M. K.; Zhou, H. X., Calculation of protein-ligand binding affinities. *Annual Review of Biophysics and Biomolecular Structure* **2007,** 36, 21-42.

241. Berhanu, W. M. M., A. E., Controlling the aggregation and rate of release in order to improve insulin formulation: molecular dynamics study of full-length insulin amyloid oligomer models. *Journal of Molecular Modeling* published online: 15 June 2011; DOI 10.1007/s00894-011-1123-3.

242. Armstrong, A. H.; Chen, J.; McKoy, A. F.; Hecht, M. H., Mutations That Replace Aromatic Side Chains Promote Aggregation of the Alzheimer's A beta Peptide. *Biochemistry* **2011,** 50, (19), 4058-4067.

243. Gutierrez, L. J.; Enriz, R. D.; Baldoni, H. A., Structural and Thermodynamic Characteristics of the Exosite Binding Pocket on the Human BACE1: A Molecular Modeling Approach. *Journal of Physical Chemistry A* **2010,** 114, (37), 10261-10269.

244. Miyamoto, S. K., P. A.,, What determines the strength of noncovalent association of ligands to proteins in aqueous-solution. *Proceedings of the National Academy of Sciences of the United States of America* **1993,** 90, (18), 8402–8406.