Binding and Destabilization of Amyloid-β Protofibrils by β-sheet Breaker Molecules: A Molecular Dynamics Simulation Study

Submitted in Partial Fulfilment of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

By

PAVAN KRISHNA KANCHI
136107014

CHEMICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

March, 2021

CERTIFICATE

It is certified that the work contained in the thesis entitled “Binding and Destabilization of Amyloid β Protofibrils by β-sheet Breaker Molecules: A Molecular Dynamics Simulation Study”, by Pavan Krishna Kanchi, has been carried out under my supervision and that this work has not been submitted elsewhere for a degree.

_____

Date:

Prof. Ashok Kumar Dasmahapatra
Department of Chemical Engineering,
Indian Institute of Technology Guwahati
Guwahati 781039, Assam

Declaration

This is to declare that the thesis entitled “Binding and Destabilization of Amyloid β Protofibrils by β-sheet Breaker Molecules: A Molecular Dynamics Simulation Study” submitted by me to Indian Institute of Technology Guwahati for the award of the degree Doctor of Philosophy is a bonafide record of research work carried out by me under the supervision of Prof. Ashok Kumar Dasmahapatra. I hereby declare that the contents of this thesis, in full or in parts, have not been submitted to any other University or Institute for the award of any degree or diploma.

Dated:

Guwahati

Pavan Krishna Kanchi

Abstract

Alzheimer's disease is a fatal neurodegenerative disease which affects the elderly population leading to the loss of memory and the ability to perform activities, and eventually death. There is no cure for Alzheimer's disease at present. One of the hallmark characteristics of this disease is the presence of plaques in the brains of patients, formed mainly by the amyloid-beta peptide in the form of beta-sheets. One of the therapeutic strategies is the use of small molecules as drugs to destabilize these fibrils. In the present thesis, we employ all atom molecular dynamics simulations in order to study the destabilization of amyloid-beta protofibrils by small molecules. The extent of destabilization and the binding affinities were characterized, and the dominant interactions which influenced the binding process were identified.

The first major finding of the thesis was that oligoproline chains of various lengths could break the beta-sheet structure of the protofibrils and induce the formation of random coils. Critical interactions which are important for protofibril stability such as hydrogen bonds and salt bridges were disrupted. Proline was able to bind strongly to the protofibrils by strong electrostatic interactions, making proline an important amino acid to consider in the design of novel peptide-based drugs.

The second major finding was that a peptide $KLVFFP_5$ which was designed to exploit the properties of the self-recognition sequence of the amyloid-β peptide KLVFF and the β-sheet breaker amino acid proline could destabilize the amyloid protofibrils to a greater extent than the KLVFF peptide. In the presence of the $KLVFFP_5$ peptide the protofibrils lost their beta-sheet structure leading to the formation of coils and helices. The hydrogen bonding network of the protofibrils and the salt bridges are critical for protofibril stability. The $KLVFFP_5$ peptide disrupted the hydrogen bonding network and the salt bridges in the protofibrils to a greater extent than the KLVFF peptide. The tight interatomic packing of the protofibrils was made looser by the $KLVFFP_5$ peptide.

The third major finding was that for a model of the protofibrils which is known to be a particularly difficult target for drugs, the increased presence of aromatic amino acids can enhance the binding of a known beta-sheet breaker peptide LPFFD, when it was modified by these aromatic amino acids. The aromatic amino acids enhanced the binding affinity by forming aromatic and hydrophobic contacts. The electrostatic complementarity of the ligands and the fibrils, along with favourable orientation of these ligands played an important role in the binding. The tryptophan modified LPFFD peptides had the highest binding affinity.

The fourth important finding was that THC molecules could destabilize the amyloid-beta protofibrils. We obtained insights into the destabilization of the protofibril structure by the THC molecules. Our results show that the THC molecules bind to the hydrophobic surface formed by hydrophobic residues of the C-terminal of the protofibrils. Upon binding, the THC molecules formed strong hydrophobic contacts with the protofibril residues and weakened interchain interactions of the protofibrils. Hydrophobic interactions were the driving force for the binding of the THC molecules to the protofibrils. By competing for these hydrophobic contacts with other protofibril residues, the THC molecules disrupted the native contacts of the protofibrils and made them less stable.

Acknowledgement

I would like to express my gratitude to my supervisor Prof. Ashok Kumar Dasmahapatra for his guidance and encouragement. I would also like to express my gratitude to the doctoral committee members Dr. A. Kumar, Dr. T. Banerjee and Dr. P. Padmanabhan for their support and helpful suggestions. I wish to thank the staff of the Chemical Engineeing department for their help at various stages. I would also like to thank all the lab members, both past and present, for their help at various points in time. I would also like to thank the many friends I made over the years both inside the campus and outside. Finally, I would like to thank my father Dr. Muralidhar Kanchi for his encouragement and support.

Table of Contents

List of Figures

List of Tables

Chapter 1. Introduction

1.1 The Prion Diseases

Scrapie was first recognized in countries of Western Europe over 250 years ago [1] affecting sheep. The earliest reported occurrence of scrapie in Britain was in 1732 [1]. In more recent times, the importation of infected sheep has introduced this disease in several other countries, including even India [1]. The prevalence of scrapie can have a major economic impact, particularly for industries dealing with cattle fed with sheep offal [1].

Scrapie is an insidious, degenerative disease affecting the central nervous system of sheep and goats [2]. Scrapie is the prototype of the group of diseases known as the sub-acute, transmissible spongiform encephalopathies which affect man and some animal species, notably ruminants. In humans, Kuru, Gerstmann-Sträussler-Scheinker (GSS) syndrome and Creutzfeldt-Jakob disease (CJD), GSS and familial CJD have a hereditary cause but are also transmissible [1]. Host genes can exert a major effect on the length of the incubation period and on clinical disease occurrence in some of the diseases, including natural scrapie in sheep. The cause of the animal diseases and kuru, a geographically localised human disease, is a polymorphic transmissible agent which is yet to be characterised. All these diseases start with a prolonged incubation period of months or years, followed by a progressive, debilitating, neurological illness which is always fatal [1]. Pathological changes which are confined to the central nervous system and include vacuolation, neuronal loss, astrocytosis and, in some diseases or individuals, amyloid plaques, scrapie-associated fibrils (SAF) in detergent-treated extracts of brain tissue may be seen with an absence of detectable inflammatory or immune responses [1].

Historically, there were two different opinions on the cause of scrapie: one being that scrapie was infectious in origin, and the other being that scrapie was genetic in origin [1], with the chemical composition of this transmissible agent being unknown. Experiments to detect the presence of scrapie in mice by Pattison *et al.* [2] to examine the possibility of the transmissible agent of scrapie being a polypeptide or being involved with a polypeptide, lead them to conclude that the agent causing scrapie could be present in some or all normal tissues, under the assumption that scrapie was not infectious in nature. In response to these experiments, Mackay [3] asserted the evidence supporting the infectious nature of scrapie and the need for

these experiments to be performed in laboratories where known strains of scrapie had not been used, and on mice from colonies which had no history of the disease. There were many attempts to study scrapie to grow the infected cells from the brains of sheep and mice in a culture. [4]. Studies by Clark and Haig showed that the "scrapie agent" multiplied in a cell culture [4].

In his seminal work, Prusiner [5] had shown that the infectious scrapie agent contained a protein, and coined the term "prion" to describe the scrapie agent as a *pro*teinaceous *in*fectious particle. It was shown later that this protein had hydrophobic domains [6]. Scrapie infectivity could be reduced by procedures that hydrolyze or modify proteins [7]. Further work had shown that the prion was different from viruses and viroids [8], although it is noted that diseases caused by prions and viruses have some similar features [7]. Legname *et al.* [9] finally proved the infectious nature of prions proteins. In their notation, PrP^{C} is the normal, cellular form of prion protein and PrP^{Sc} is the pathogenic scrapie form of prion protein. Scrapie infectivity was found to be resistant to UV radiation at 250nm [7]. According to Prusiner, prions are composed of a modified protein PrP^{Sc}, without nucleic acid (being resistant to procedures that alter nucleic acids) [7]. By means of a posttranslational process PrP^{C} is converted into PrP^{Sc} during which it acquires a high β-sheet content [7].

It was suggested that a conformational change takes place in the formation of PrP^{Sc} [10] by Pan *et al.* whose experiments showed that PrP^{C} had a high α-helix content and PrP^{Sc} had both α and β structures, indicating that the conformational change from α-helices to β-sheets may be a fundamental event in the propogation of prions. PrP^{C} and PrP^{Sc} have different properties and conformations [10].

1.2 Alzheimer's Disease

Neuritic plaques are a kind of lesion observed in patients which are microscopic foci of extracellular amyloid deposition and associated axonal and dendritic injury usually populated in limbic and association cortices [11]. These plaques contain extracellular deposits of star-shaped masses of amyloid fibrils within and around which there are dilated and tortuous dystrophic neuritis with enlarged lysosomes, mitochondria, paired helical filaments and associated microglia expressing surface antigens like CD45 and HLA-DR. Astrocytes are found outside the plaque [11]. It is hypothesized that these legions form over a period of time ranging from months to years. The cross-sectional diameter of the neuritic plaques varies from 10 to >120 μm.

The other type of lesions found in brains of patients with Alzheimer's disease is the neurofibrillary tangles which are large nonmembrane-bound insoluble bundles of abnormal fibers which occupy most of the perinuclear cytoplasm composed of hyperphosphorylated tau proteins[11], termed tauopathies which tend to self-associate to form paired helical filament structures[12]. Enzymes associated with the addition and removal of phosphate residues determine the extent of tau phosphorylation [12]. Intermediate aggregates of abnormal tau molecules are cytotoxic and impair cognition [12]. It is hypothesized that Aβ accumulation precedes and drives tau aggregation.

Alzheimer's disease leads primarily to failure and loss of the hippocampal synapses [12]. The normally high levels of neutrotrophin receptors in cholinergic neurons in the basal forebrain are drastically reduced in diseased brains. Free radicals released from dysfunctional mitochondria result in oxidative stress. Glucose intolerance and type 2 diabetes are hypothesized to be risk factors for dementia, while about 60% to 90% of patients with Alzheimer's disease are affected by ischemic disease. There may also be a loss of calcium regulation which stimulates aggregation by increased cytosolic calium concentrations. [12]

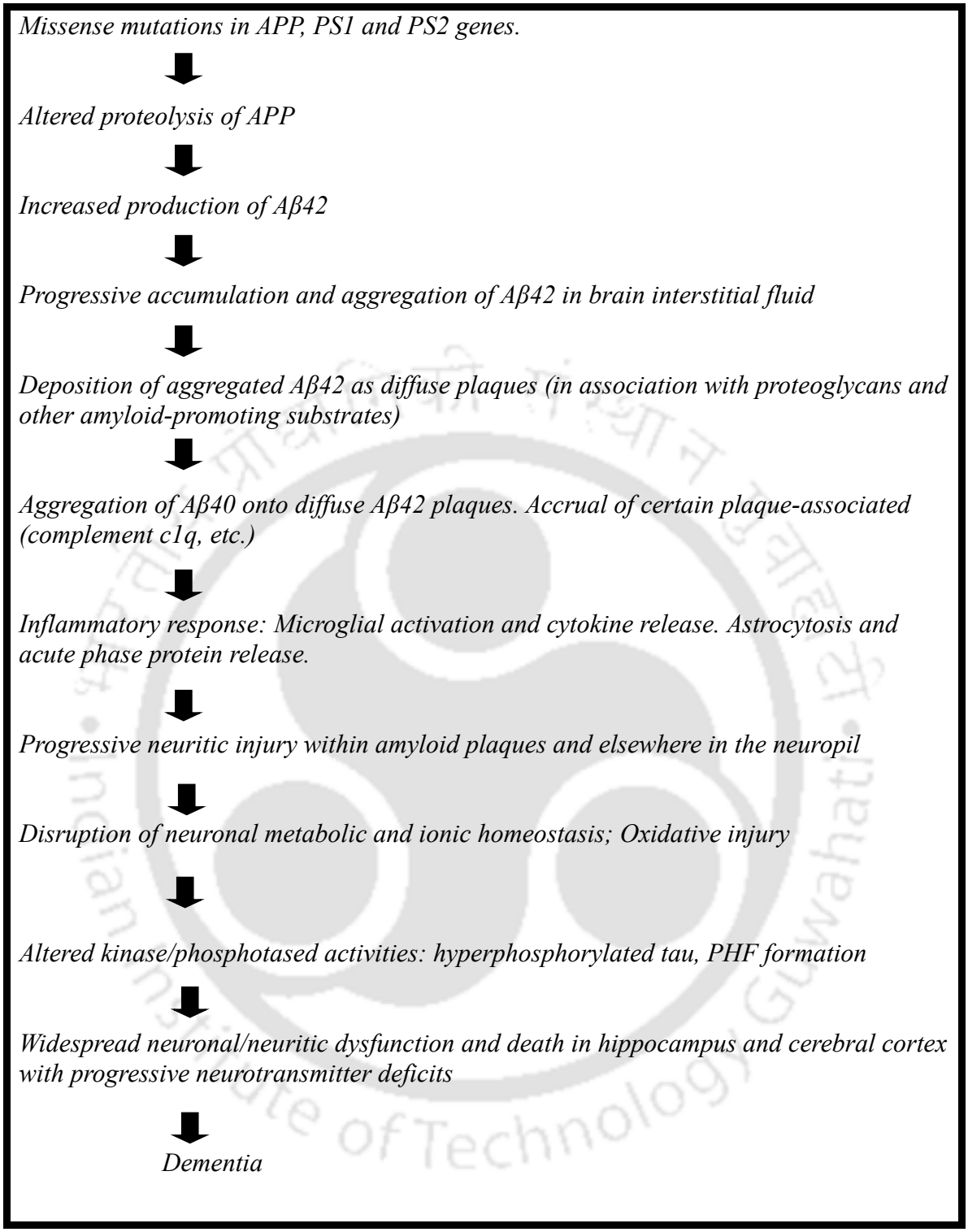


Figure 1.1: A Hypothetical sequence of the pathogenic steps of familial forms of Alzheimer's disease [11]

Table 1.1: Summary of main human amyloidoses [13]

Clinical syndrome	**Plaque components**
Alzheimer's disease	Amyloid β peptide; tau protein
Parkinson's disease	α-synuclein
Huntington's disease	Huntington
Spongiform encephalopathy	Prion protein
Type II diabetes	Islet amyloid polypeptide
Familial amyloidotic polyneuropathy I	Mutant transthyretin and fragments
Senile systemic amyloidosis	Wild-type transthyretin and fragments
Haemodialysis-related amyloidosis	β_2-microglobulin
Finnish hereditary amyloidosis	Mutant gelsolin fragments
Hereditary systemic amyloidosis	Mutant lysozyme

1.3 Protein Aggregation Kinetics and Mechanism

The number of different proteins in the human body is about 100,000 [14]. Elaborate mechanisms exist in biological systems to ensure the proper folding of proteins, and incorrectly folded proteins are detected and degraded before they can cause any harm to the organism. Under certain conditions proteins may misfold and form aggregates. Productive aggregation of proteins also occurs in nature, such as the fibrillation of globular actin molecules to actin fibrils; and the self-assembly of bovine liver glutamate dehydrogenase [15], [16]. Aggregation in these cases is natural and not toxic.

There are several physical methods to study aggregation of proteins: absorbance studies, atomic force microscopy, calorimetry, circular dichroism, studies using dyes, electron microscopy, electron paramagnetic resonance spectroscopy, flow birefringence, fluorescence spectroscopy with an extrinsic/intrinsic fluorophore, Fourier transform infrared spectroscopy, light scattering studies, mass spectrometry, nuclear magnetic resonance spectroscopy, quartz crystal oscillator measurements, turbidity studies, viscosity studies and x-ray diffraction [15].

The process of aggregation begins with protein monomers. It has been shown that in some cases, like the α-synuclein aggregation leading to Parkinson's disease, protein monomers which were unfolded had a greater tendency to aggregate. Aggregation leads to the amyloid fibrils being formed, which are insoluble, and in some cases, amorphous aggregates. The

amorphous aggregates are hypothesized as off-pathway products in the reaction: these have been reported to be toxic and may be soluble or insoluble. It has been observed that amyloid fibrils share a similar structure – cylindrical β-sheets with β-strands in the anti-parallel configuration [17]. The protofilament has a helical array of β-sheets twisted around the main axis of the protofilament and the β-strands are perpendicular to the main axis. When the protofilaments comprise of more than about 37 to 40 residues, there is a greater probability of disease occurring, because of greater stability arising from the result of such fibres having at least two concentric β-sheets which are held in place by hydrogen bonds between their amides and which can also become nuclei for further growth [17].

Protein aggregation is an entropy-driven process with a positive enthalpy and entropy. The increased entropy arises from water molecules being released during aggregation. [18]

Broadly there are three approaches to determine the kinetics and mechanism for protein aggregation depending on the information required from analysis: the kinetic and thermodynamic approach, empirical approaches, and the prion mechanism approach [15].

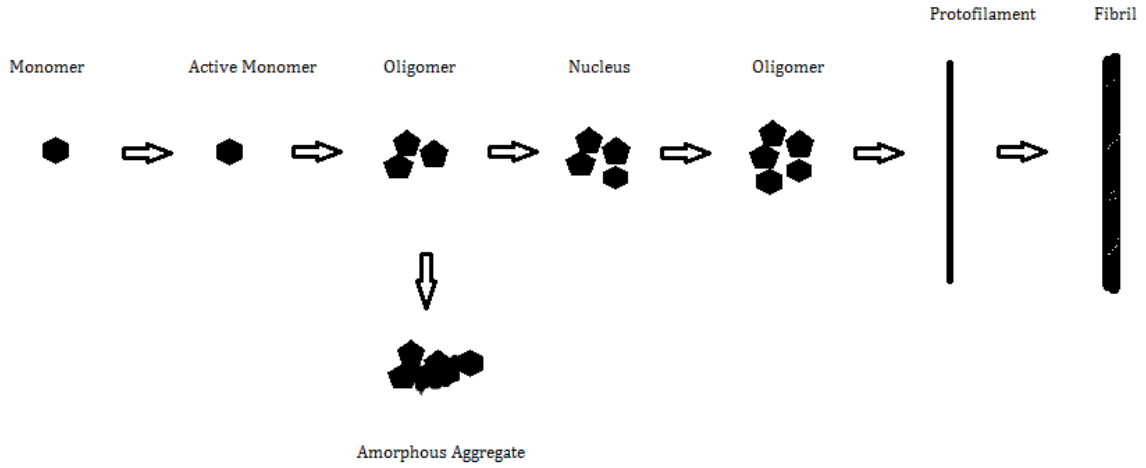


Figure 1.2: General pathway of the formation of fibrils [15]

Morris *et al.* [15] have classified the postulated mechanisms of the protein aggregation process into five classes: 1) the subsequent monomer addition mechanism, 2) the reversible association mechanism, 3) prion aggregation mechanisms, 4) the Finke-Watzky 2-step model, and 5) quantitative structure activity relationship models.

1.4 Amyloid Fibrils

The term amyloid was coined by Rudolph Virchow in 1854 when he used iodine to stain cerebral corpora amylacea (with an abnormal appearance) to pale blue and adding sulphuric acid to this made it turn violet, leading him to conclude that the abnormality was made of cellulose [19]. In 1859, it was demonstrated that the amyloid plaques contained protein due to its high nitrogen content, suggesting that there was no carbohydrate content. Historically, light microscopy and histopathological dyes like Thioflavin T and Congo red were used in the characterization of amyloid plaques [13]. Positive birefringence of the deposits dyed with Congo red suggested an ordered submicroscopic structure whereas electron microscopy studies demonstrated the fibrillar structure of the plaques leading to studies on isolating these fibrils and sequencing the proteins [13]. Amyloid deposits of various origins exhibit a similar fibrillar microscopic structure consisting of fibrils about 60 to 130 Å wide and 1000 to 16000 Å long [13].

Initially it was assumed that proteolysis determined the conversion of amyloidogenic proteins into fibrils although now it is widely accepted that protein misfolding is the molecular basis of all amyloid-based disorders. Proteolysis generates aberrant protein fragments highly prone to forming plaques, for example the Aβ peptide, which is the proteolysis product of the amyloid precursor protein.

Aβ is produced by endoproteolysis of the parental amyloid precursor protein (APP) which occurs by the sequential cleavage of APP by the α-, β-, and γ-secretases [20]. APP can be cleaved and processed in either a non-amyloidogenic pathway or an amyloidogenic pathway [20]. In the latter, the initial proteolysis is mediated by the β-secretase at the position 99 residues away from the C terminus. As a result of this cleavage, the 99 residue C-terminal stub is left inside the membrane, which is further cleaved by γ-secretase to yield an intact Aβ peptide [20]. Most of the segments produced are 40 residues in length (Aβ40) although the 42 residue variant Aβ42 is more hydrophobic and more prone to aggregation [20].

A proteinaceous fibrillar structure is amyloid if it shows the following characteristics [13]:

1. Straight unbranched fibrillar morphology detected by EM, with widths of about 7-12 nm and indeterminate length.
2. Cross-β X-ray diffraction pattern.
3. Binding to dyes like Congo Red and Thioflavin T.
4. Protein solutions display a polymeric β-sheet CD signature.

It was hypothesized that the ability to form amyloid fibrils is a generic property of all polypeptide chains. It has been shown that the SH3 module of PI3 kinase and acylphosphotase formed fibrils in denaturing conditions [13].

It is hypothesized that certain sequence motifs are more prone to self assembly than others [13]. It has been shown that for the case of the $A\beta_{1-40}$ peptide, the sequence of residues 16-21 is the shortest sequence able to form fibrils in vitro but is not toxic in cell culture.

According to Mezzenga [21], our current understanding of the process of the formation of fibrils identifies three main critical steps:

i At sufficient contour lengths, single protofilaments align on approaching each other due to liquid crystalline interactions;

ii attractive interactions of hydrophobic or Lennard- Jones nature overcome electrostatic repulsion among likewise charged protofilaments and lead to a merging of individual protofilments into multistranded fibrils precursors;

iii The fibrils start to twist and develop with time a well-defined twisting pitch, whose handiness is settled by chirality and further amplified by electrostatic iteractions.

1.5 Properties of Fibrils [21]

Based on how the protofilaments pack, amyloid fibrils show different morphologies and properties. At molecular length scales, NMR is a valuable tool. TEM and AFM provide structural details at the single fibril length scale and via statistical analysis can yield averages over populations of hundreds to thousands of fibrils. TEM, cryogenic TEM (cryo-TEM), and scanning transmission electron microscopy (STEM) can yield also other very valuable information, which are not accessible to other techniques, such as, for example the mass per unit length of amyloid fibrils. Then, if the exact peptidic sequence forming the amyloid fibrils is known, this quantity can be ideally exploited to identify and reconstruct topological models of the fibrils. Two main packing mechanisms are the close-packed model (CP) and a ribbon-like packing model. Post-packing there may be many transient polymorphic states, such as twisted ribbons, helical ribbons, or nanotubes. It may be possible that the driving force behind this packing process is the minimization of energy. Amyloid fibrils have high elasticity, stiffness, and resistance, and their Young's moduli are in the order of several GPa, comparable

to the most rigid proteinaceous materials found in nature. Amyloid fibrils may be characterised by their persistence length, which is the typical length at which thermal fluctuations begin to bend the polymer in different directions. If this persistence length is less than contour length, then the fibril is considered flexible. From AFM images, we can find l_p via the bond correlation function for semiflexible polymers in 2D conformations:

$$\langle \cos \theta(s) \rangle = \exp\left(\frac{-s}{2l_p}\right) \qquad \text{(Equation 1.1)}$$

where θ is the angle between the tangent vectors to the chain at two points separated by a contour distance s. The factor 2 is used to rescale the exponential decay accounting for the two-dimensional nature of fibrils absorbed on a substrate.

Persistence length is related to Young's modulus, E, by:

$$l_p = \frac{EI}{k_B T} \qquad \text{(Equation 1.2)}$$

where I is the moment of inertia of the fibril cross section, k_B is the Boltzmann constant, and T is temperature. E depends directly on the specific peptide sequence by which the individual protofilaments of the amyloid fibrils are formed, and I depends on the form and the dimensions of the fibril cross section. Larger cross-sections therefore lead to large I and large persistence length.

The effects of water in protein aggregation are poorly understood, and it is also experimentally difficult to monitor water activity in the fibrillation process [22]. Water alters substantially the hydrophobic and electrostatic interactions [22]. Water influences the energy landscape of Aβ monomers: for example, a stable intermolecular salt bridge can only form if the intervening water molecules are expelled in the corresponding region which requires overcoming a large barrier to desolvation [22]. Water molecules surround the surface of folded proteins with the thickness of the hydration layer being about 7 Å and are in equilibrium with the bulk water in solution. Upon crystallization, these molecules are released into the bulk causing an increase in overall entropy and hypothesized to drive the crystallization reaction [22].

β-lactoglobulin fibrils have shown to form amyloid aggregates when subjected to a Couette shear flow by generating precursors which act as seeds for aggregation. Physiological shear rates range from 100 to 8000 s^{-1} in blood vessels and in exocellular matrix, and can be significant [23]. Nanomechanical stretching of the silk-elasin-like peptide polymer (SELP) extending its conformation led to the formation of nucleation sites where other molecules could assemble revealing a correlation between molecular stretching and amyloid nucleation where the fibre growth is perpendicular to the direction of the stretching [23].

Molecular dynamics studies on the aggregation of Aβ16-22 monomers found a transient intermediate oligomer state with high α-helical content with increased inter-peptide interactions which eventually converts to a β-structure, a surprising result as both the monomers and fibrils have no α-helical content [13]. Side-chain interactions play an important role in the formation of an oligomer [13]. It was also found that the interior of Aβ oligomers is dry and that hydrophobic interactions between monomers are the driving force in the association of $A\beta_{10-35}$ peptides into dimers [13].

The high concentration of macromolecules present in all cells has the consequence that most of the interior space of the cells in unavailable and this has shown to dramatically accelerate fibrillation [24].

1.6 Motivation and Objectives

1.6.1 Motivation

There is presently no cure for Alzheimer's disease. Small molecule drugs which can destabilize and prevent the aggregation of the misfolded amyloid-beta proteins are a therapeutic strategy for Alzheimer's disease [25]. Drug discovery and development for Alzheimer's disease is difficult and new drugs have not been approved since 2003 - no disease-modifying treatments, which are approved exist [26]. Small molecules-based drugs which can cross the blood-brain barrier have attracted attention in treating Alzheimer's disease. Several efforts had been made over the past two decades to identify different classes of small molecules which could inhibit the aggregation of amyloid-beta fibrils into neurotoxic species and to destabilize the amyloid-beta fibrils. Conventional small molecule drugs are limited by their low selectivity indicated by side-effects in humans and a low affinity to the amyloid fibrils [27] [28]. These drawbacks can be overcome by peptide-based drugs which may serve as an alternative option to these chemical compounds. Protein therapeutics have a greater specificity for their targets because of their ability to have more interactions with them [28]. Peptide-based drugs are highly potent and selective with a broad range of targets. The motivation of the present study is to investigate the destabilization of the amyloid-beta fibrils by beta-sheet breaker peptides and molecules. In the present thesis, we present a study associated with the destabilization of the amyloid-beta fibrils by beta-sheet breaker molecules and their binding affinities by employing protein docking and all-atom molecular dynamics simulations. For this purpose, novel peptide sequences were

designed in order to study their destabilizing effect and their binding affinities. The destabilization of the amyloid-beta fibrils by the THC molecule was also studied.

1.6.2 Objectives

The objectives of the thesis are outlined as follows:

1. Identifying small molecules which can destabilize the amyloid-beta fibrils.
2. Characterizing the destabilization mechanism of the amyloid-beta fibrils.
3. The design of novel peptide sequences based on the properties of constituent amino acids which can destabilize the amyloid-beta fibrils.
4. Characterizing the mechanism of binding of these peptides and small molecules.
5. Investigating the relation between binding affinity and the extent of destabilization of the fibrils.

1.7 Organization of the thesis

A brief introduction to the various prion diseases and Alzheimer's disease is provided in **Chapter 1**, along with the motivation and objectives of the thesis. In **Chapter 2**, we discuss the methodology of the molecular dynamics simulation protocols. In **Chapter 3**, we present a study of the destabilization of the amyloid protofibrils by oligoproline chains and their associated binding affinities. Based on these findings, in **Chapter 4** we present a study where we determine whether the destabilization and binding properties of the amino acid proline can be combined with an existing beta-sheet breaker peptide KLVFF. We designed a novel peptide with the sequence KLVFFPPPPP and showed that the presence of proline improved the destabilization and binding ability of the KLVFF sequence. Based on our observations that proline binds well to the fibrils, we were then interested in the amino acids with the highest binding affinities to the fibrils: the aromatic amino acids. In **Chapter 5**, we present a study in which we demonstrated that when the beta-sheet breaker peptide LPFFD was modified by aromatic amino acids, its binding to the fibrils was enhanced, particularly to a model of the fibrils which is known to be a difficult target for drugs. In **Chapter 6**, we demonstrate that the THC molecule is able to destabilize the amyloid-beta protofibrils and can bind to them by hydrophobic interactions. Finally, in **Chapter 7**, we present a summary of the thesis with a discussion on the scope for future studies.

References

1. L.A. Detwiler, Scrapie., Rev. Sci. Tech. 11 (1992) 491–537.
2. I.H. Pattison, K.M. Jones, Detection of the scrapie agent in tissues of normal mice and in tumours of tumour-bearing but otherwise normal mice., Nature. 218 (1968) 102–104.
3. J.M. Mackay, Detection of the scrapie agent in tissues of normal mice., Nature. 219 (1968) 182–183.
4. M.C. Clarke, D.A. Haig, Evidence for the multiplication of scrapie agent in cell culture., Nature. 225 (1970) 100–101.
5. S.B. Prusiner, Novel proteinaceous infectious particles cause scrapie., Science. 216 (1982) 136–144.
6. S.B. Prusiner, M.P. McKinley, D.F. Groth, K.A. Bowman, N.I. Mock, S.P. Cochran, F.R. Masiarz, Scrapie agent contains a hydrophobic protein., Proc. Natl. Acad. Sci. U. S. A. 78 (1981) 6675–6679.
7. S.B. Prusiner, Prions, Proc. Natl. Acad. Sci. 95 (1998) 13363 LP – 13383.
8. T.O. Diener, M.P. McKinley, S.B. Prusiner, Viroids and prions, Proc. Natl. Acad. Sci. U. S. A. 79 (1982) 5220–5224.
9. G. Legname, I. V Baskakov, H.-O.B. Nguyen, D. Riesner, F.E. Cohen, S.J. DeArmond, S.B. Prusiner, Synthetic mammalian prions., Science. 305 (2004) 673–676.
10. K.M. Pan, M. Baldwin, J. Nguyen, M. Gasset, A. Serban, D. Groth, I. Mehlhorn, Z. Huang, R.J. Fletterick, F.E. Cohen, *et al.*., Conversion of alpha-helices into beta-sheets features in the formation of the scrapie prion proteins, Proc. Natl. Acad. Sci. U. S. A. 90 (1993) 10962–10966.
11. D.J. Selkoe, Alzheimer's disease: genes, proteins, and therapy., Physiol. Rev. 81 (2001) 741–766. doi:10.1152/physrev.2001.81.2.741.
12. H.W. Querfurth, F.M. LaFerla, Alzheimer's disease., N. Engl. J. Med. 362 (2010) 329–344.
13. V. Muñoz, ed., Protein Folding, Misfolding and Aggregation, The Royal Society of Chemistry, 2008.
14. C.M. Dobson, Protein misfolding, evolution and disease., Trends Biochem. Sci. 24 (1999) 329–332.
15. A.M. Morris, M.A. Watzky, R.G. Finke, Protein aggregation kinetics, mechanism, and curve-fitting: A review of the literature, Biochim. Biophys. Acta - Proteins Proteomics. 1794 (2009) 375–397.
16. D. Thusius, Mechanism of bovine liver glutamate dehydrogenase self-assembly: II. Simulation of relaxation spectra for an open linear polymerization proceeding via a sequential addition of monomer units, J. Mol. Biol. 94 (1975) 367–383.
17. M.F. Perutz, J.T. Finch, J. Berriman, A. Lesk, Amyloid fibers are water-filled nanotubes., Proc. Natl. Acad. Sci. U. S. A. 99 (2002) 5591–5595.
18. S.S. Licht, C.C. Lawrence, J. Stubbe, Thermodynamic and kinetic studies on carbon-cobalt bond homolysis by ribonucleoside triphosphate reductase: the importance of entropy in catalysis., Biochemistry. 38 (1999) 1234–1242.
19. J.D. Sipe, A.S. Cohen, Review: history of the amyloid fibril., J. Struct. Biol. 130 (2000) 88–98.
20. F.M. LaFerla, K.N. Green, S. Oddo, Intracellular amyloid-beta in Alzheimer's disease., Nat. Rev. Neurosci. 8 (2007) 499–509.

21. J. Adamcik, R. Mezzenga, Proteins Fibrils from a Polymer Physics Perspective, Macromolecules. 45 (2012) 1137–1150.
22. D. Thirumalai, G. Reddy, J.E. Straub, Role of water in protein aggregation and amyloid polymorphism., Acc. Chem. Res. 45 (2012) 83–92.
23. N. Varongchayakul, S. Johnson, T. Quabili, J. Cappello, H. Ghandehari, S.D.J. Solares, W. Hwang, J. Seog, Direct observation of amyloid nucleation under nanomechanical stretching., ACS Nano. 7 (2013) 7734–7743.
24. K.W. Plaxco, M. Gross, Cell biology. The importance of being unfolded., Nature. 386 (1997) 657,659.
25. C.A. Ross, M.A. Poirier, Protein aggregation and neurodegenerative disease., Nat. Med. 10 Suppl (2004) S10-7.
26. J. Cummings, G. Lee, A. Ritter, M. Sabbagh, K. Zhong, Alzheimer's disease drug development pipeline: 2019, Alzheimer's Dement. (New York, N. Y.). 5 (2019) 272–293.
27. D. Goyal, S. Shuaib, S. Mann, B. Goyal, Rationally Designed Peptides and Peptidomimetics as Inhibitors of Amyloid-β (Aβ) Aggregation: Potential Therapeutics of Alzheimer's Disease., ACS Comb. Sci. 19 (2017) 55–80.
28. D.J. Craik, D.P. Fairlie, S. Liras, D. Price, The future of peptide-based drugs., Chem. Biol. Drug Des. 81 (2013) 136–147.

Chapter 2. Methodology

There are two techniques which may be used in the development of compounds that may inhibit Aβ aggregation: molecular dynamics simulations and docking [1]. GROMACS [2] was used for conducting the molecular dynamics simulations, while ClusPro [3] for docking. The following part of this chapter is derived from the GROMACS User Manual [4].

Edge effects in a finite system are minimized by applying periodic boundary conditions. There are translated copies of the atoms of the system surrounding a set of atoms in a space-filling box thereby removing the boundaries of the system and thus replacing the edge effects by different conditions, i.e., periodic boundary conditions. The periodic boundary conditions are not suitable for non-periodic systems such as liquids and solutions.

Of the possible shapes to fill a unit cell, the rhombic dodecahedron and the truncated octahedron are best suited for systems which are approximately spherical as these shapes use less solvent molecules to fill the box. These shapes are special cases of triclinic unit cells. According to the minimum image convention, only the nearest image of each particle is considered for short-range non-bonded interaction terms. For long-range electrostatic interaction lattice sum methods such as Ewald sum, PME and PPPM are used.

The unit cell is defined by 3 box vectors $\vec{a}$, $\vec{b}$, and $\vec{c}$. The box vectors must satisfy the following conditions:

$$a_y = a_z = b_z = 0$$
$$a_x > 0, b_y > 0, c_z > 0$$
$$|b_x| \leq \frac{a_x}{2}, |c_x| \leq \frac{a_x}{2}, |c_y| \leq \frac{b_y}{2} \qquad \text{(Equation 2.1)}$$

In order for the expressions in Equation 1 to be satisfied in the most efficient way, GROMACS places the particles in a brick-shaped volume for a 2-dimensional system. The trajectory can be converted to a different unit-cell representation after the simulation.

A consequence of the minimum image convention is that the cut-off radius which is used to truncate non-bonded interactions may not exceed half the shortest box vector:

$$R_c < \frac{1}{2} min\left(||\vec{a}||, ||\vec{b}||, ||\vec{c}||\right) \qquad \text{(Equation 2.2)}$$

This is done to prevent more than one image from being within the cut-off distance of the force. For the case of proteins however, due to the effects of the solvent, the length of each

box must exceed the length of the macromolecule in the direction of that edge plus two times the cut-off radius R_c.

Table 2.1. The global MD Algorithm

Input initial conditions Potential interaction V as a function of atom positions Positions r of all atoms in the system Velocities v of all atoms in the system
Repeat 2,3,4 for the required number of steps:
Compute forces The force on any atom $F_i = \frac{-\partial V}{\partial r_i}$ is computed by calculating the force between non-bonded atom pairs, $F_i = \sum_j F_{ij}$, along with the forces due to bonded interactions, restraining and external forces. The potential and kinetic energies and the pressure tensor are computed.
Update configuration The movement of the atoms is simulated by numerically solving Newton's equations of motion: $$\frac{d^2 r_i}{dt^2} = \frac{F_i}{m_i}$$ Or $$\frac{dr_i}{dt} = v_i \; ; \frac{dv_i}{dt} = \frac{F_i}{m_i}$$
Output step Write positions, velocities, energies, temperature, pressure, etc.

Initially, the topology and the force field are read in. The box size and the coordinates and velocities of all particles are required before a run starts. The box size and shape is determined by the three basis vectors of the periodic box. If velocities are not available, the program can generate initial atomic velocities with a Maxwell-Boltzmann distribution at a given temperature. The update algorithm introduces a very slow change in the centre-of-mass velocity, and therefore in the total kinetic energy of the system – especially when temperature coupling is used. Such changes are quenched in order to prevent the temperature from being misinterpreted.

Particles that are close to a given particle must be found, and this searching may be done with periodic boundary conditions (if a grid search is used, of the order O (N)) or otherwise (O (N^2)).

The temperature is given by the total kinetic energy of the N-particle system:

$$E_{kin} = \frac{1}{2}\sum_{i=1}^{N} m_i v_i^2 \quad \text{(Equation 2.3)}$$

From this the absolute temperature T can be computed by:

$$\frac{1}{2} N_{df} kT = E_{kin} \quad \text{(Equation 2.4)}$$

where k is the Boltzmann's constant and N_{df} is the number of degrees of freedom.

The pressure tensor **P** is calculated from the expression:

$$P = \frac{2}{V}(E_{kin} - \Xi) \quad \text{(Equation 2.5)}$$

Here kinetic energy is expressed as a tensor, which is required when shear forces are imposed, V is the volume and **Ξ** is the virial tensor. The scalar pressure may be obtained from:

$$P = \frac{1}{3} trace(\vec{P}) \quad \text{(Equation 2.6)}$$

The total potential energy is summed for various contributions: the Lennard-Jones, Coulomb and bonded terms. The default MD integrator in GROMACS is the leap-frog algorithm. Velocity Verlet integrators may be preferable when extremely accurate integration is needed with temperature and pressure coupling. However, the velocity verlet algorithm is associated with an increase in computational cost and slightly less accurate kinetic energies. In the present study, the leap frog algorithm was employed.

The leap-frog integrator:

$$v\left(t + \frac{1}{2}\Delta t\right) = v\left(t - \frac{1}{2}\Delta t\right) + \frac{\Delta t}{m} F(t) \quad \text{(Equation 2.7)}$$

$$r(t + \Delta t) = r(t) + \Delta t v\left(t + \frac{1}{2}\Delta t\right) \quad \text{(Equation 2.8)}$$

$$r(t + \Delta t) = 2r(t) - r(t - \Delta t) + \frac{1}{m} F(t)\Delta t^2 + O(\Delta t^4) \quad \text{(Equation 2.9)}$$

The Velocity verlet integrator:

$$v\left(t + \frac{1}{2}\Delta t\right) = v(t) + \frac{\Delta t}{2m} F(t) \quad \text{(Equation 2.10)}$$

$$r(t + \Delta t) = r(t) + \Delta t v\left(t + \frac{1}{2}\Delta t\right) \quad \text{(Equation 2.11)}$$

$$v(t + \Delta t) = v\left(t + \frac{1}{2}\Delta t\right) + \frac{\Delta t}{2m} F(t + \Delta t) \quad \text{(Equation 2.12)}$$

or equivalently,

$$r(t + \Delta t) = r(t) + \Delta t v + \frac{\Delta t^2}{2m} F(t) \quad \text{(Equation 2.13)}$$

$$v(t + \Delta t) = v(t) + \frac{\Delta t}{2m}[F(t) + F(t + \Delta t)] \quad \text{(Equation 2.14)}$$

Temperature coupling is used when an NVT ensemble is studied. The weak coupling scheme of Berendsen, the extended ensemble Nose-Hoover and the velocity rescaling scheme may be used in GROMACS. The Berendsen algorithm is based on a weak coupling to an

external heat bath at a temperature T_0. A deviation of the system from T_0 is slowly corrected according to:

$$\frac{dT}{dt} = \frac{T_0 - T}{\tau} \qquad \text{(Equation 2.15)}$$

where τ is the time constant. The Berendsen thermostat does not generate rigorously the canonical ensemble, as fluctuations in the kinetic energy are suppressed and fluctuation properties such as the heat capacity may be inaccurate. This error scales as 1/N. The velocity rescaling thermostat generates a correct ensemble with a slight modification of the Berendsen method according to the following expression:

$$dK = (K_0 - K)\frac{dt}{\tau_T} + 2\sqrt{\frac{KK_0}{N_f}}\frac{dW}{\sqrt{\tau_T}} \qquad \text{(Equation 2.16)}$$

where K is the kinetic energy, N_f the number of degrees of freedom and dW is a Wiener process.

Pressure coupling may be done by the Parrinello-Rahman, Berendsen and the Martyna-Tuckerman-Tobias-Klein schemes. The Berendsen pressure coupling algorithm has the effect of a first-order kinetic relaxation of the pressure towards a reference pressure according to:

$$\frac{dP}{dt} = \frac{P_0 - P}{\tau_p} \qquad \text{(Equation 2.17)}$$

Table 2.2: The MD update algorithm

Given: **Positions of all atoms at a given time** **Velocities of all atoms at a given time** **Accelarations on all atoms at a given time** **Total kinetic energy and virial at a given time**
Compute the scaling factors
Update and scale velocities
Compute new unconstrained coordinates
Apply constraint algorithm to coordinates
Correct velocities for constraints
Scale coordinates and box

Constraints may be implemented by using LINCS or the SHAKE method. The SHAKE algorithm changes a set of unconstrained coordinates to a set of coordinates that fulfil a list of distance constraints; while the LINCS algorithm resets bonds to their correct lengths after an unconstrained update. In the present study, the constraints were applied to all bonds.

Energy minimization may be done using steepest descent, conjugate gradients, or the limited-memory Broyden-Fletcher-Goldfarb-Shanno quasi-Newtonian minimizer.

The potential functions may be classified as non-bonded functions (Lennard-Jones or Buckingham, and Coulomb or modified Coulomb), bonded functions (covalent bond-stretching, angle bending, improper dihedrals, proper dihedrals), and restraints (position restraints, angle restraints, distance restraints, orientation restraints and dihedral restraints.)

The total electrostatic energy of N particles and the periodic images is given by

$$V = \frac{f}{2}\sum_{n_x}\sum_{n_y}\sum_{n_{z*}}\sum_{i}^{N}\sum_{j}^{N}\frac{q_i q_j}{r_{ij,n}} \qquad \text{(Equation 2.18)}$$

where $(n_x, n_y, n_z) = \vec{n}$ is the box index vector. The asterisk indicates that the terms with the same indices i and j should not be included when $\vec{n} = 0$. $r_{ij,\mathbf{n}}$ is the distance between the charges. The sum converges very slowly.

By means of Ewald summation, the sum may be converted to two terms, which converge fast, and another constant term:

$$V = V_{dir} + V_{rec} + V_0 \qquad \text{(Equation 2.19)}$$

where

$$V_{dir} = \frac{f}{2}\sum_{i,j}^{N}\sum_{n_x}\sum_{n_y}\sum_{n_{z*}} q_i q_j \frac{erfc(\beta r_{ij,n})}{r_{ij,n}} \qquad \text{(Equation 2.20)}$$

$$V_{rec} = \frac{f}{2\pi V}\sum_{i,j}^{N} q_i q_j \sum_{m_x}\sum_{m_y}\sum_{m_{z*}} \frac{exp\left(\left(\frac{-\pi m}{\beta}\right)^2 + 2\pi i m\cdot(r_i - r_j)\right)}{m^2} \qquad \text{(Equation 2.21)}$$

$$V_0 = \frac{-f\beta}{\sqrt{\pi}}\sum_{i}^{n} q_i^2 \qquad \text{(Equation 2.22)}$$

Here β measures the relative weight of the direct and reciprocal sums and the wave vector $\vec{m} = (m_x, m_y, m_z)$. The Ewald summation method is not suitable for large systems as the computational cost of the reciprocal term increases as N^2.

An alternate to the Ewald summation method is the particle-mesh Ewald method where charges are assigned to a grid by interpolation. The grid is then subject to a Fourier transformation and the reciprocal energy term is obtained by a single sum over the grid in k-space. The computational cost is proportional to N log (N) and is therefore significantly faster than Ewald summation. Another method is the particle-particle particle-mesh method.

The simulation can be run in parallel over more than one processor. In GROMACS there exist two types of parallelization: particle decomposition and domain decomposition. Of these, domain decomposition scales faster. In particle decomposition, particles and forces between particles are assigned to processors, and each processor requires the coordinates of at least half of the particles in the system. For N processors, N(N/2) coordinates are required, as a result of

which the particle decomposition method does not scale well. It is used in cases where domain decomposition does not work, for instance systems with long-range bonded interactions. In domain decomposition, a spatial domain is assigned to each processor which then integrates the equations of motion for the particles in that domain.

For the case of atomistic simulations, the coordinates of the proteins are taken from the Protein Data Bank for simulation [5]. From the pdb file, the topology and the structure files are generated. The topology file contains information needed to define the molecules. From here, it is possible to choose from the force fields in GROMACS. The OPLS-AA, GROMOS96 53A6 and GROMOS96 54A7 force fields are well suited for the simulation of the fibrils [6]. Heavy atoms are constrained by position restraints. After the protein is placed in an appropriate box the box is filled with water. In the present study, explicit water models were employed. Ions may be added to neutralize the overall charges of the system, if any exist. In order that the geometry of the structure is represented accurately, and to prevent steric clashes, the structure is relaxed by energy minimization. In order to bring the solvent and ions to the temperature specified to simulate, and to orient them correctly around the protein, equilibration is conducted in two stages: an NVT ensemble and then an NPT ensemble. The position restraints are then removed and the system is simulated. In the present study, a timestep of 2fs was used.

Coarse-grained simulations are done using the Native structure-based models, also called the Gō-type models [7] or by using the MARTINI force field [8]. A structure based model is generated from a pdb file using a set of python tools [7]. Residues are represented as a bead at the position of the C-alpha atoms. From the model, the required configuration files for GROMACS simulations are generated which are then used for the simulations. Another set of python tools similarly exist for setting up files for the MARTINI force field simulations.

The simulations may be analysed by a number of parameters, depending on the aspect of interest. Some of these are:

- Potential energy, as a function of the reaction pathway
- Radial distribution function: The local density of a species around another species relative to the average density.
- Root mean square deviation and root mean square fluctuation of atom postions.
- Radius of gyration: measure for the compactness of a structure.
- The distance between atoms and the minimum distance between groups of atoms: Changes in the structure may be analysed by plotting these distances.
- The average lifetime of a hydrogen bond, and the total number of hydrogen bonds in each time frame.

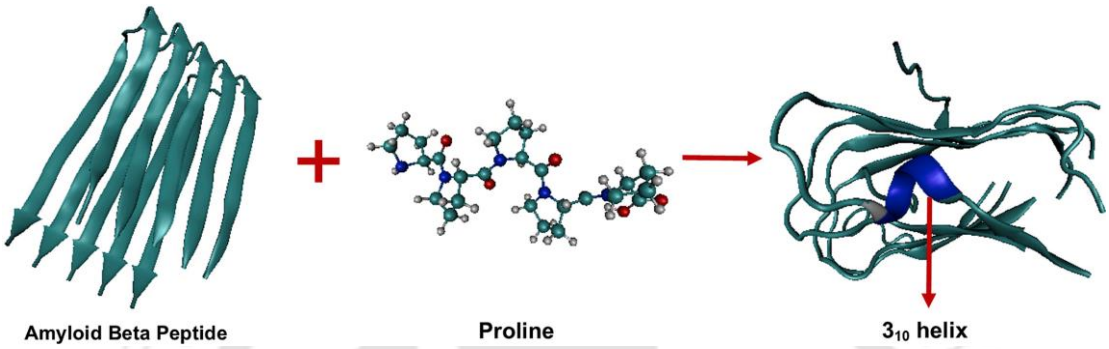
- Secondary structure analysis.
- Ramachandran plots.
- Principal component analysis: to obtain modes of collective motion and their amplitudes.
- Potential of mean force: by steered molecular dynamics and umbrella sampling.
- The solvent accessible surface area.
- The number of formed native contacts in a structure.
- The heat capacity over a temperature range to describe changes in the structure.

.

References

1. J.A. Lemkul, D.R. Bevan, The role of molecular simulations in the development of inhibitors of amyloid-β peptide aggregation for the treatment of Alzheimer's disease., ACS Chem. Neurosci. 3 (2012) 845–856.
2. B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation., J. Chem. Theory Comput. 4 (2008) 435–447.
3. S.R. Comeau, D.W. Gatchell, S. Vajda, C.J. Camacho, ClusPro: an automated docking and discrimination method for the prediction of protein complexes., Bioinformatics. 20 (2004) 45–50.
4. GROMACS 4.5.4 User Manual
5. H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The Protein Data Bank., Nucleic Acids Res. 28 (2000) 235–242.
6. S.R. Gerben, J.A. Lemkul, A.M. Brown, D.R. Bevan, Comparing atomistic molecular mechanics force fields for a difficult target: a case study on the Alzheimer's amyloid β-peptide., J. Biomol. Struct. Dyn. 32 (2014) 1817–1832. doi:10.1080/07391102.2013.838518.
7. B. Lutz, C. Sinner, G. Heuermann, A. Verma, A. Schug, eSBMTools 1.0: enhanced native structure-based modeling tools., Bioinformatics. 29 (2013) 2795–2796.
8. S.J. Marrink, H.J. Risselada, S. Yefimov, D.P. Tieleman, A.H. de Vries, The MARTINI force field: coarse grained model for biomolecular simulations., J. Phys. Chem. B. 111 (2007) 7812–7824.

Chapter 3. Polyproline chains destabilize the Alzheimer's amyloid-β protofibrils: A molecular dynamics simulation study



3.1. Introduction

Alzheimer's is a fatal neurodegenerative disease and a primary source of dementia. More than 35 million people worldwide are affected by Alzhemier's disease (as of 2010) which leads to the progressive loss of memory, the ability to perform activities, speech, recognition of people and objects, and ultimately death (within 3–9 years of diagnosis) [1,2]. The neuritic plaques consisting of extracellular aggregates of amyloid fibrils and intracellular aggregates called neurofibrillary tangles, formed by tau proteins are found in the brains of patients with Alzheimer's disease [2][3]. Therapeutic strategies include small molecules as drugs to prevent aggregation of the misfolded amyloid-beta proteins [2].

Small molecules-based drugs have attracted attention in treating various diseases, as they may be able to cross the blood-brain barrier. The blood-brain barrier is highly selective and does not allow the entry to large molecules (viz., molecular mass more than 400 Da) and 98% of small molecules [4]. The only small molecule drugs discovered so far, which can cross the blood-brain barrier, are largely for the treatment of depression, schizophrenia, chronic pain and epilepsy [4]. Some molecules that have been studied thus far which have had some success in preventing aggregation, binding to the fibrils or disrupting the beta sheet structures are Congo red, LPFFD (a designed pentapeptide), myricetin, melatonin, 9,10 anthraquinone, sugar

trehalose, Thioflavin T, N-methylated peptides, polyphenol (−)-epigallocatechin-3-gallate (EGCG), ibuprofen, naproxen, morin [5], wine-related polyphenolic compounds [6], piceid [7], fullerene [8], a fullerene derivative [9] and the peptide FFVLK [10]. An inhibitor wgx-50 [11] was found to break amyloid fibrils into monomers, inhibit their aggregation, and inhibit Aβ-induced neuronal apoptosis and apoptotic gene expression with the reduction of neuronal calcium toxicity. It was also found that wgx-50 could penetrate the blood-brain barrier and improve the cognitive abilities of mice [12]. The destabilization of Aβ fibrils by wgx-50 was confirmed by molecular dynamics (MD) studies [13]. The Aβ peptide-induced transmembrane ion conductance in fibril barrels was also blocked by wgx-50 [14]. The wgx-50 complexed with gold nanoparticles had an enhanced effect on the inhibition of aggregation of Aβ fibrils [15]. In vitro studies showed that a peptide, Humanin, is capable of protecting neurons from being killed by toxic Aβs [16].

Over 100 peptide-based drugs are in the ethical pharmaceutical market, with a growing share of about 10% [17]. Peptides of length 8 to 10 amino acids are more common, while some can even be as long as 45 amino acids [17]. The peptide-based drugs with molecular weight greater than 500 Da are usually delivered by injection [17]. The advantages of peptide-based drugs are their high potency, high selectivity, and broad range of targets, potential lower toxicity than small molecules, low accumulation in tissues, high chemical and biological diversity, and being discoverable at peptide and/or nucleic acid levels [17]. However, peptide-based drugs have several disadvantages: poor metabolic stability, poor membrane permeability, poor oral availability, high cost, and poor solubility in some cases [17]. If the challenges of using peptide-based drugs are overcome, they can potentially be used to treat diseases effectively and reliably [17].

Proline-rich native antibacterial peptides can enter bacterial and host cells. Hence, these molecules can potentially be used to transport peptide-based drugs into the target cells [18]. Clinical and experimental studies have demonstrated that proline-rich polypeptides can potentially be therapeutic in treating neurodegenerative diseases such as Alzheimer's disease [19]. These polypeptides displayed neuroprotective, immunomodulatory and antioxidant activity [19]. It has also been shown that Colostrinin, which contains proline-rich polypeptides, improved cognition in patients with Alzheimer's disease [19]. It has been postulated that the immunomodulatory action stimulated by Colostrinin may play a vital role in the treatment of Alzheimer's disease [19]. It has also been shown that Colostrinin successfully prevents the

aggregation of amyloid-β (1–40) in vitro [19]. In clinical studies, Leszek *et al*. had shown that when a set of patients with Alzheimer's disease were treated with Colostronin, 50% of them showed improvement and 50% showed stabilization of their health condition [20]. In another study 33% of the patients showed stabilization or improvement of the disease [21]. The PRP-1 (proline-rich polypeptide-1) is a 15 residue hypothalamic proline-rich polypeptide which contains four proline residues and can potentially be a preventive or therapeutic agent against Alzheimer's disease and protect brain neurons as shown by a study by Yenkoyan *et al*. [22].

The amino acid proline has some unique properties due to which they are rarely found in β-sheets [23]. One of these is the conformation of the peptidyl-prolyl bond, which is incompatible with the geometries of peptide bond in β-sheets [24]. Sterically, the ring of the proline residue cannot fit into the hydrogen bonding network of the β-sheet [24]. In addition, the nitrogen of the peptide bond in the peptidyl-prolyl group does not contribute to the hydrogen bonding network of β-sheets [24]. The side chains of amino acids influence the freedom of internal rotations of the protein backbone. Therefore, the entropic contribution to the free energy of the unfolded protein depends on the type of amino acids present in the protein [25]. The pyrrolodine ring of proline restricts proline to access fewer available conformations [26]. In a study by Wood *et al*., it was shown that replacing any residue with proline in the amyloidogenic sequence LVFFAED (residues 17–23 in the Alzheimer's amyloid peptide) leads to the complete loss of fibril formation and makes the peptide soluble [24]. Elastomeric proteins possess rubber-like elasticity and can be subjected to a high deformation force without rupturing [27]. In a molecular dynamics simulation study, sequences similar to those in elastin, an elastomeric protein, were found to form amorphous and disordered aggregates, while those without proline residues formed amyloid-like structures with β-sheets [26].

It has been postulated that there are reverse-turn and 3_{10} helix intermediates in the folding/unfolding of alpha helices [[28], [29], [30], [31], [32]]. Most 3_{10} helices found in proteins are 3–4 residues long, and usually exist as binding surfaces or active sites on the surface of proteins [28]. The side chain interactions may explain how the stability of 3_{10} helices is maintained. Energetically, for short helical segments, a 3_{10} helix may be preferred to an alpha helix [33][34]. In another experimental study, a segment of the Aβ peptide, 17–23, which was constrained as a helix, made branched fibrils shorter with less beta sheet content [35].

In our present work, we study the mechanism of the beta-sheet breaker proline on the Alzheimer's amyloid beta protofibrils with oligoproline chains of different lengths by atomistic molecular dynamics (MD) simulations. We seek to understand the impact of different number of proline residues on the protofibrils, particularly whether more proline residues cause more loss of beta-sheet structure and more binding. We also wish to know which regions proline prefers to bind to and which residues it interacts with. Proline is a hydrophobic residue, and we wish to investigate if hydrophobic interactions play a role in binding. We also wish to know the overall suitability of proline in drug design. In all the systems, proline destabilizes the protofibrils, tends to break the beta-sheet structure, and causes the protofibrils to adopt random coil structures and tends to reduce the number of hydrogen bonds in the chains, which are at the peripheries, causing them to move away from the beta sheets. Proline binds well to the protofibrils by electrostatic interactions. This indicates that proline may be a reliable candidate as drug molecules and can be employed in therapeutics as indicated by experimental studies [19, 20, 21] of polypeptides, which contain significant proline content.

3.2. Methodology

The protofibril system we have considered, obtained from the Protein Data Bank (PDB ID: 2BEG), has five chains [36], which we will refer to as A, B, C, D and E. The system is shown in Fig. 3.1. N-terminal residues 1–16 are missing in this model. However, it has been hypothesized that the remaining residues contribute more to the stability of the fibrils [37, 38]. The region spanned by the residues 18–26 is denoted as β1 and the region spanned by the residues 31–42 is denoted as β2. Two intermolecular and parallel β-sheets are formed by β1 and β2 which are connected by a bend region [36].

Fig. 3.1. The five chain $A\beta_{17-42}$ system considered for the simulations represented as a cartoon. The structure consists of parallel β-sheets and a bend region. In this figure, different colors indicate different chains.

The proline chains considered for docking with the protofibrils were a 2mer, 5mer, 6mer and a 7mer. Three control runs were performed: these were of the amyloid-beta protofibrils alone. One replica simulation each was also done for all the systems, using the same starting velocities. For convenience, we will use the nomenclature described in Table 3.1 to describe the systems considered.

Table 3.1. Nomenclature for the systems considered.

System Index	Name	System
1	Control 1	$A\beta_{17-42}$
2	Control 2	$A\beta_{17-42}$
3	Control 3	$A\beta_{17-42}$
4	Pro_2-1	Proline 2mer + $A\beta_{17-42}$
5	Pro_2 – 2 (Replica)	Proline 2mer + $A\beta_{17-42}$
6	Pro_5-1	Proline 5mer + $A\beta_{17-42}$
7	Pro_5-2 (Replica)	Proline 5mer + $A\beta_{17-42}$
8	Pro_6-1	Proline 6mer + $A\beta_{17-42}$
9	Pro_6-2 (Replica)	Proline 6mer + $A\beta_{17-42}$
10	Pro_7-1	Proline 7mer + $A\beta_{17-42}$
11	Pro_7-2 (Replica)	Proline 7mer + $A\beta_{17-42}$

Proline chains were constructed using Avogadro 1.1.1 [39]. The proline chains were then docked to the receptor amyloid- β fibrils by using the automated protein docking server ClusPro [40], [41], [42], [43], and [44]. Since 2000, a community-wide experiment called the critical assessment of predicted interactions (CAPRI) began in which various docking algorithms are compared to the unpublished structures shared by structural biologists, only with the evaluators [45]. There have been more than 34 rounds of CAPRI so far. The ClusPro was found to yield acceptable or medium accuracy models in many rounds of CAPRI, and can therefore identify best near-native conformations [41]. Hence, ClusPro was used to generate the amyloid fibrils – proline complex. The recommended default procedure was followed in ClusPro. Based on favorable desolvation energies and electrostatics, the top 2000 models were clustered and ranked by the ClusPro algorithm [43]. Out of the top 10 models generated by the ClusPro algorithm, the top model was chosen as there was no prior information about the complex to be able to be biased towards any other model. In the case of the Pro5 systems, we chose the third highest ranked model, which was a complex in which the 5mer proline chain was in the inner region of the inner hydrophobic core of the amyloid fibrils, i.e. the concave edge, since this region is not accessible to most molecules. This shows that proline may be able to access this inner region over a long period of time associated with an energy penalty. In addition, this proline chain is also able to interact with the Asp 23-Lys 28 salt bridges. The starting structures are shown in Fig. 3.2.

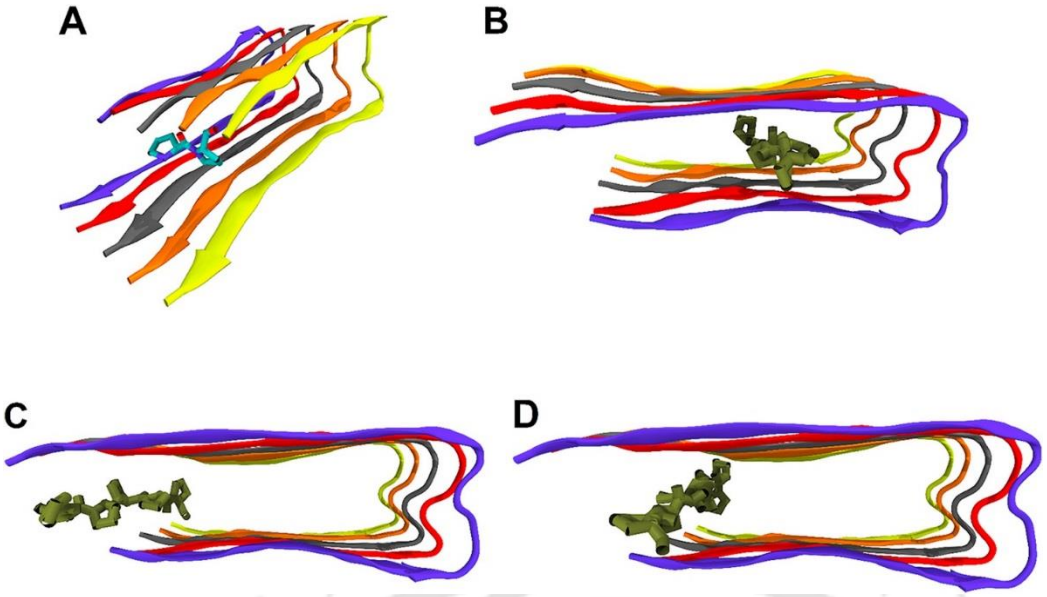


Fig. 3.2. Results of the docking algorithm. The protofibrils are shown in a cartoon representation and the proline chains are shown as bonds. **A.** The Pro_2 system. The proline chain is nearer to the N-termini region. **B.** The Pro_5 system. Here the proline chains are in the inner concave hydrophobic region. **C.** The Pro_6 system. The proline chains are near both the N- and C-termini. **D.** The Pro_7 system. The proline chains are again near both the N- and C-termini.

Each system was placed in a cubic box of size 6.84 nm of simple point charge (SPC) water. The simulations employed the OPLS all-atom force field as it has been shown that it is well suited for the simulation of this structure of the protofibrils [46]. Five sodium ions were added to each system to make the system neutral in charge. Following energy minimization using the steepest descent algorithm, the systems were equilibrated in two steps with position restraints applied to heavy atoms throughout. An equilibration was carried out under NVT conditions at 300 K using coupling to the temperature bath (Nose Hoover [47, 48]). Following this, an equilibration was carried out under NPT conditions with the system coupled to a barostat (Parinello-Rahman [49] [50] and a thermostat at 300 K. All the equilibrations were done for 100ps. Finally, the MD simulation with position restraints removed was run for 400ns using GROMACS 5.1.1 [51]. The periodic boundary conditions were applied in all the directions. The constraint algorithm used for bond lengths was P-LINCS [52]. The neighbor search cut-off was at approximately 1.4 nm. The fast smooth particle mesh Ewald summation method was used to calculate the long-range electrostatic interactions and the Fourier grid spacing was 0.12 nm [53].

In our analysis, the first 100ns were not considered. The minimum distance of a protein system to its periodic image was measured to check for artefacts. The secondary structure was calculated using the do_dssp [54, 55] tool in GROMACS. The fraction of residues which

formed a secondary structure was then calculated for every time frame. Hydrogen bonds were calculated using the hbon

d [56] tool in GROMACS. Salt bridges were calculated using VMD [57], hydrophobic contacts with LigPlot [58]. Python 3.7.3 was used to calculate running averages over 100 windows, and for plotting results. Solvent accessible surface areas (SASA) were calculated using the GROMACS tool sasa [59].

The MM/PBSA method was used to calculate the binding free energies of the amyloid protofibrils with the oligoproline chains. In MM/PBSA, the binding energy is calculated as follows [60, 61].

$$\Delta G^{bind} = \Delta EMM + \Delta G^{psolv} + \Delta G^{npsolv} - T\Delta S \qquad \text{(Equation 3.1)}$$

Where ΔEMM is the molecular mechanics contribution to the binding free energy in vacuum, ΔG^{psolv} is the polar contribution to the solvation energy (calculated by solving the Poisson-Boltzmann equation) and ΔG^{npsolv} is the non-polar contribution to the solvation energy (calculated by using the solvent-accessible surface area (SASA) model.) Entropic terms were not included in calculating the binding energies of the different lengths of oligoproline chains.

The g_mmpbsa tool [62, 63] was used to calculate the binding free energies using the single trajectory protocol. Snapshots from the last 10ns of the production run extracted every 100ps were used for the analysis. Energetic components are expressed as an average with standard deviations. The contributions to the binding energy made by individual residues were also calculated using 1000 bootstraps in the g_mmpbsa program. The solute and solvent dielectric constants were taken as 4 and 80 respectively. The non-linear Poisson-Boltzmann equation was solved to calculate the polar solvation energy.

3.3 Results and discussion

We begin with describing the effect of proline on the secondary structure, hydrogen bonding and salt bridges of the protofibrils. We then describe the binding mode of proline and the residues which contribute favorably to binding energy.

3.3.1. Proline disrupted the secondary structure, hydrogen bonds, and salt bridges of the amyloid protofibrils

Secondary Structure. Local and tertiary environment interactions influence the secondary structure of a given segment of a protein backbone [64]. We found that the interaction with proline chains destabilizes the amyloid fibrils. There is a significant loss of β-sheet content in all the systems considered with the formation of turns and random coils. The final structures of the all the systems considered are shown in Fig. 3.3.

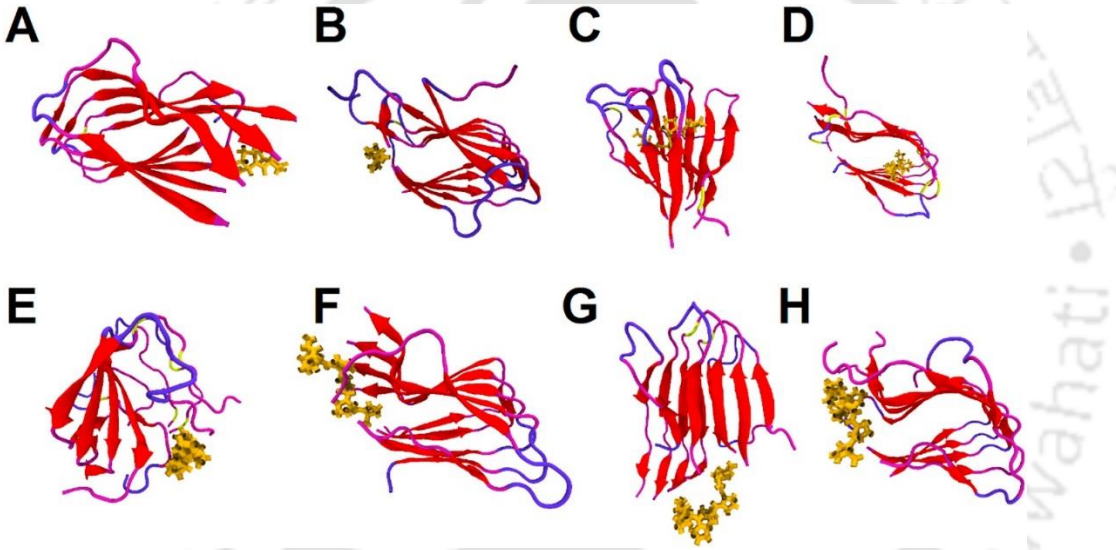


Fig. 3.3. Final structures (after 400ns) of the systems considered with the protofibrils represented as a cartoon. Extended β-sheets are shown in red color, bridge-β in yellow, turns in blue, coils in purple. Proline chains are shown in orange represented by bonds. A. Pro_2 -1 system B. Pro_2-2 system C. Pro_5-1 system D. Pro_5-2 system E. Pro_6-1 system F. Pro_6 -2 system G. Pro_7 -1 system H. Pro_7 -2 system

In the Pro_2 - 1 system, the C-terminal residue of the proline chain made contacts with residues in the N-terminal regions of chains C and D and 40 Val D. This affected the hydrogen bonding network of the last four residues in the N terminus region of chain C with the N terminus region in chain B leading to the formation of random coils in these residues of chain B. Apart from these few residues, the β-sheet structure of the protofibrils remained somewhat preserved. In the Pro_2 - 2 system, proline made contacts with residues in the N terminal of chain C and D and the C-terminal of chains D and E. The disruption of hydrogen bonds by proline caused the loss in β-sheet structure in the regions 36 Val – 41 Ile in chains D and E, and 35 Met

– 41 Ile in chains A, B and C. This loss in β-sheet structure is reflected in Fig. 3.4 (system index 5).

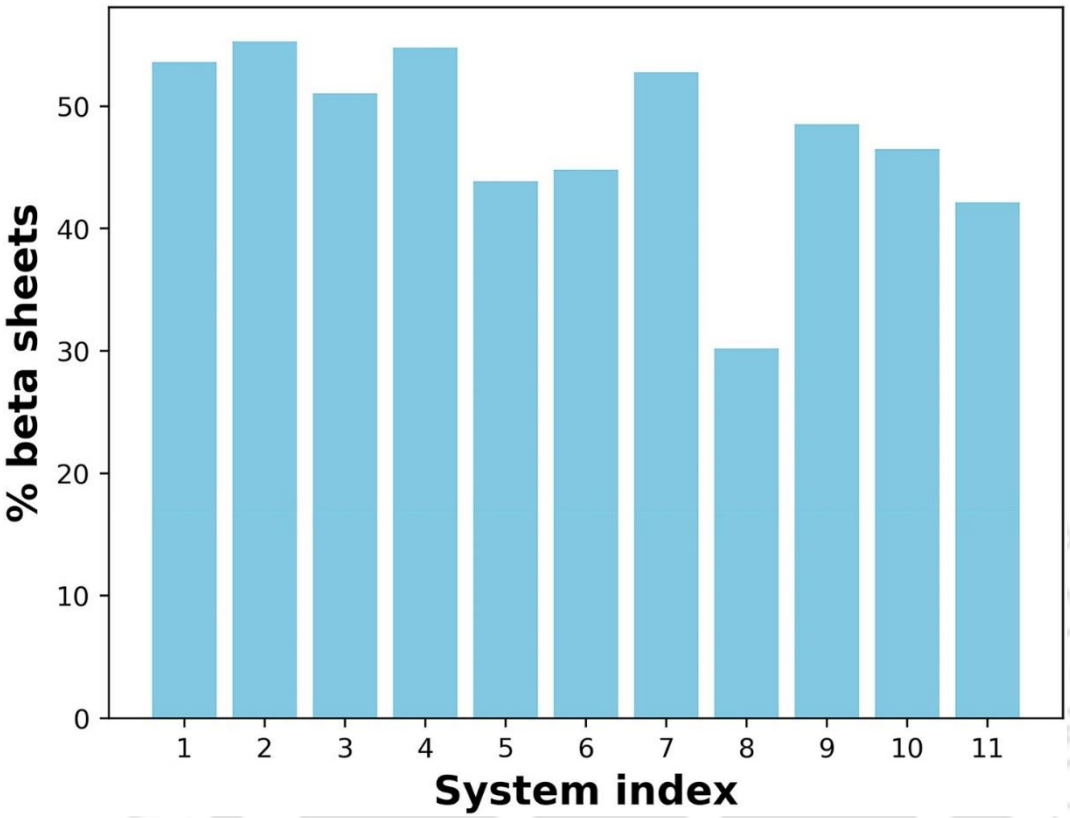


Fig. 3.4. Percentage of residues forming β-sheets in the last 300ns. According to this graph, there is a considerable reduction in β-sheet in all systems containing proline except systems 4 and 7. The system indices are, according to Table 3.1, as follows: 1 – 3: Control systems; 4: Pro_2 -1 system; 5: Pro_2 -2 system; 6: Pro_5 -1 system; 7: Pro_5 -2 system; 8: Pro_6 -1 system, 9: Pro_6 -2 system, 10: Pro_7 -1 system, and 11: Pro_7 -2 system.

In the Pro_5 systems, proline made several contacts as it was in the inner concave hydrophobic core of the protofibrils and disrupted several hydrogen bonds. In the Pro_5 -1 system the prominent loss of β-sheet structure and formation of random coils was in the 21 Ala – 26 Ser region in chain A and 18 Val – 26 Ser region in chain E. This can be seen in Fig. 3.4 (system index 6). Similarly, in the Pro_5 -2 system the 21 Ala – 26 Ser region formed random coils.

The Pro_6 – 1 system showed the maximum loss of β-sheet structure. In the Pro_6 systems, the C-terminal proline residues interacted with several residues in the N-terminal regions whereas, the N-terminal proline residues interact with the C-terminal regions of chains C and D. Disruption of hydrogen bonds due to several contacts made by proline lead to the 17 Leu –

Ser 26 regions in chains A, D and E forming random coils. In the Pro_6 – 2 system the 35 Met A – 41 Ile A region formed random coils. In the Pro_7 systems, the N-terminus proline residue interacted with the C-terminal residues 42 Ala C and 42 Ala D of the protofibrils. This weakened the hydrogen bonding network of residues in these terminal regions with the corresponding terminal regions in chains B and E respectively. Similarly, the C-terminus proline residues interacted with the N-terminal 17 Leu residues of chains C and D weakening their hydrogen bonds with chains B and E respectively. Due to the disruption in the hydrogen bonds, these regions in chains B and E lost their β-sheet structure and formed random coils.

As can be seen in Fig. 4, the Pro_2 -2, Pro_5 -1, Pro_6 -2 and the Pro_7 systems showed comparable average percentage beta sheet structure. While the Pro_2 – 1 and Pro_5 – 2 systems preserved the most β-sheets, the Pro_6 – 1 system had lost the most β-sheet content. It appears that while proline chains can reduce the β-sheet content of the protofibrils, different lengths of the proline oligomers considered in our study had a similar effect on the β-sheet content, with the Pro_2 -1, Pro_5–2 and Pro_6 -1 systems being outliers. The contrast in the values observed in these systems may be due to a difference in the nature of contacts established by the ligands with the protofibril residues and resolved by increasing the duration of the simulations and number of replicates in order to improve the observed statistics and draw better inferences. This is also reflected in Fig. 3.5, in which the average percentages of residues forming random coils are comparable for all the chains except Pro_2 -1 and Pro_6 - 1.

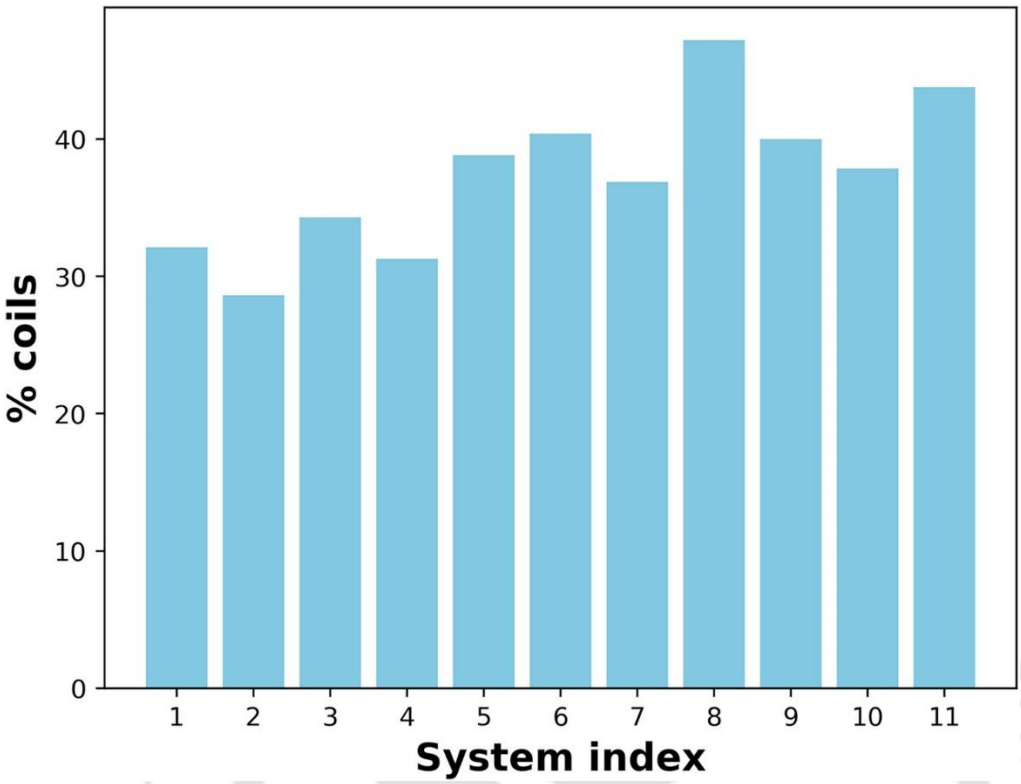


Fig. 3.5. Percentage of residues forming coils in the last 300ns. The system indices are, according to Table 3.1, as follows: 1 – 3: Control systems; 4: Pro_2 -1 system; 5: Pro_2 -2 system; 6: Pro_5 -1 system; 7: Pro_5 -2 system; 8: Pro_6 -1 system, 9: Pro_6 -2 system, 10: Pro_7 -1 system, and 11: Pro_7 -2 system. According to this graph, there is a considerable increase in random coil content in all systems containing proline except system 4.

Transient 3_{10} helices were observed in control system 2 (16.9% of the total number of frames), the Pro_2 -1 (4.55%), Pro_5- 1 (20.45%), Pro_6- 1 (1.52%), Pro_7 -1 (7.85%) and Pro_7 -1 (2.5%) systems. Control system 2 (4.55%) and the Pro_5- 1 system (1.02%) also had residues forming transient alpha helices. In the Pro_2 -1 system, residues 39 Val, 40 Val, and 41 Ile of chain A adopt a 3_{10}-helix conformation for a very short period of time. In the Pro_5 -1 system, 3_{10} helix conformation was adopted by residues 22 Glu, 23 Asp, 24 Val and 25 Gly of chain E and also by 26 Ser, 27 Asn, and 28 Lys of the other peripheral chain, A. Fig. 3.6 shows the presence these two helices at an instant in the Pro_5-1 system. In the Pro_6- 1 system, the residues 39 Val, 40 Val, and 41 Ile of chain D adopt the 3_{10}-helix conformation. In the Pro_7 -1 system, 31 Ile, 32 Ile and 33 Gly formed the helix. In the Pro_7- 2 system, the residues 39 Val, 40 Val, and 41 Ile of chain B formed the 3_{10} helix. Minimal alpha helical content was also observed in the control system 2 (0.05%) and the Pro_5- 1 system (1.02%). Helix formation was preceded by the formation of random coils. We note that these helices were formed in regions containing charged residues.

Fig. 3.6. The Pro_5 -1 system at an instance when two transient 3_{10} helices were formed. The helices were formed by 22 Glu, 23 Asp, 24 Val and 25 Gly of chain E (bottom helix) and by 26 Ser, 27 Asn, and 28 Lys of chain A (top helix). The helices are shown in green color, extended β-sheets are shown in red color, bridge-β in yellow, turns in blue, coils in purple.

Previous studies on amyloid-β fibrils indicated that the existence of a 3_{10} helix is possible. Vivekanandan *et al*. found 3_{10} helices in $A\beta_{1\text{-}40}$ fibrils that were partially unfolded in an aqueous medium, with residues 13 to 23 being most likely to be found in this conformation [65]. They also found that the residues in the region 20–26 are prone to forming turns and bends [65]. In their study, hydrophobic surfaces were formed on both the sides of the helix by the clustering of hydrophobic residues leading to the interaction between the ends of the fibrils (which contain turns) with these clusters [65]. A replica exchange molecular dynamics (REMD) study on the structure of a peptide fragment of the residues 21 to 30 of the Aβ fibrils found that the residues 26 to 28 adopt a 3_{10} helix conformation [66].

Hydrogen Bonding. Hydrogen bonds between the β-sheets stabilize the protofibril structure. Fig. 3.7 shows show the number of hydrogen bonds between neighboring chains. In general, the total number of hydrogen bonds between the peripheral chains and their neighbors decreases

after interaction with the oligoproline chains, with the exception of the Pro_2 – 1 system. At least one of these peripheral chains is pulled away from the other chains as a result, as can be expected since these have only one neighboring chain each and the side chains are freer to move. Fig. 3.8 shows the average number of hydrogen bonds between neighboring chains.

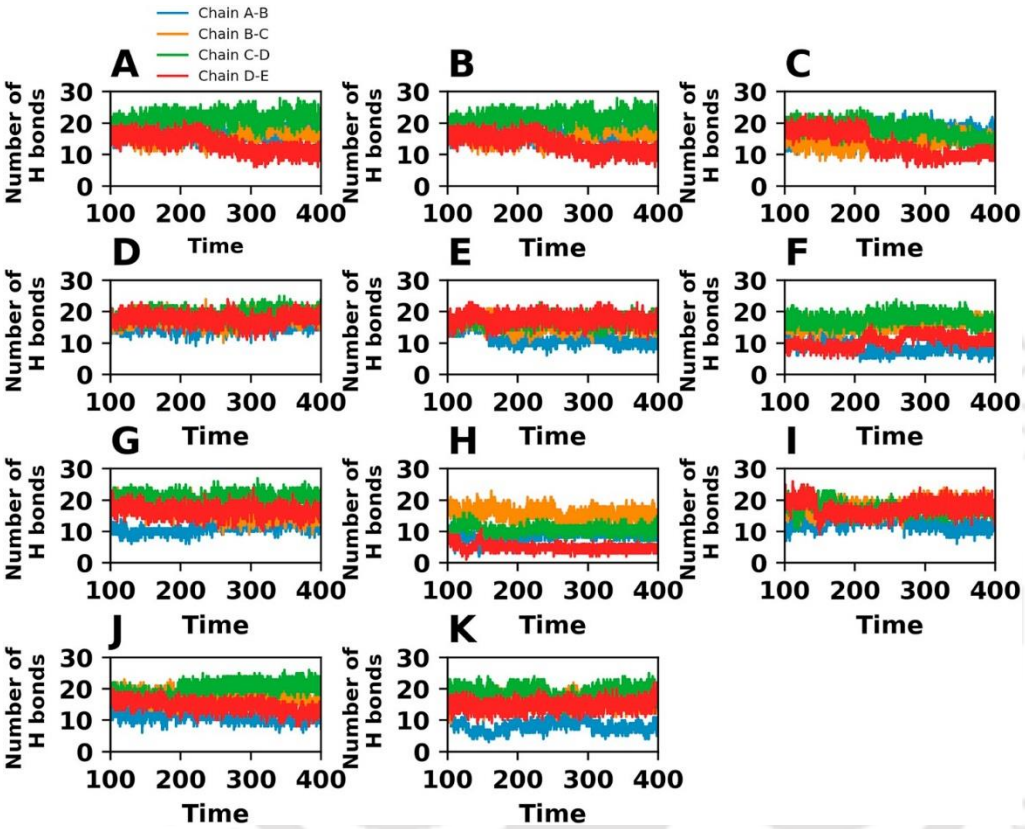


Fig. 3.7. The number of hydrogen bonds formed between neighboring chains as a function of time. **A.** Control system 1. **B.** Control system 2. **C.** Control system 3. **D.** Pro_2 -1 system. **E.** Pro_2 -2 system. **F.** Pro_5 -1 system. **G.** Pro_5 -2 system. **H.** Pro_6 -1 system. **I.** Pro_6 -2 system. **J.** Pro_7 -1 system. **K.** Pro_7 -2 system.

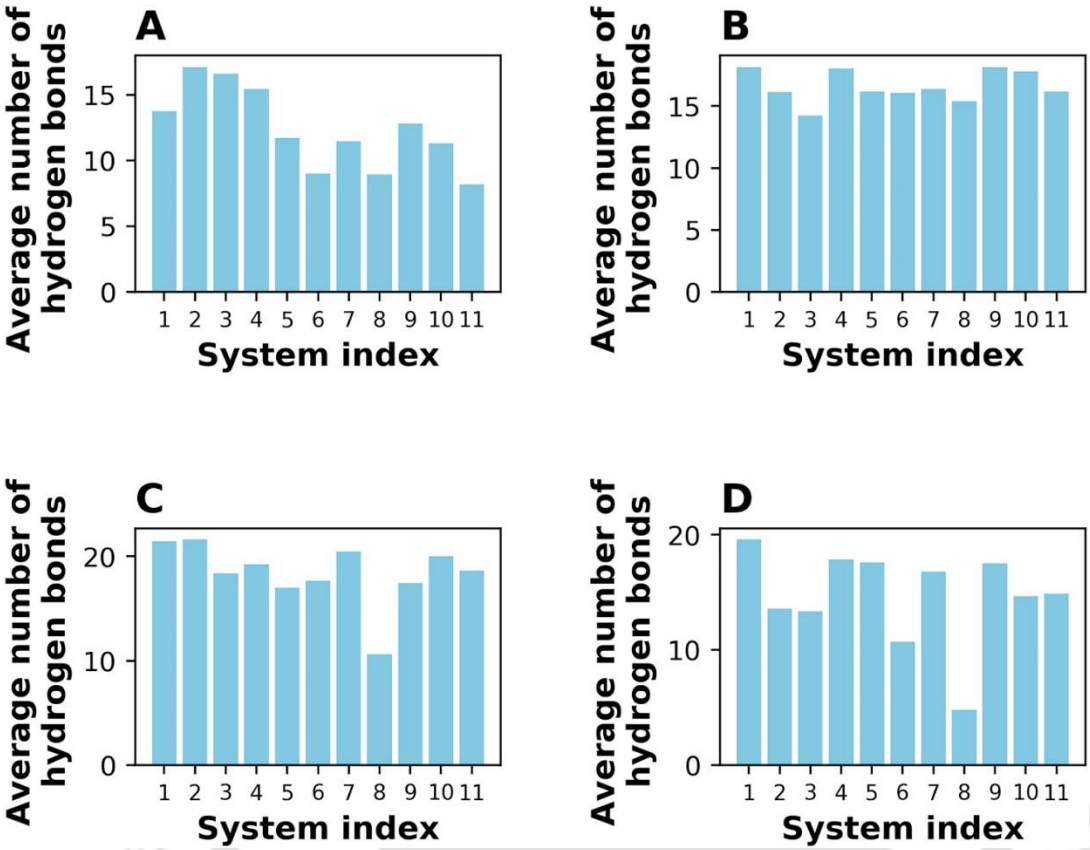


Fig. 3.8. Average number of hydrogen bonds between neighboring chains formed by all the systems in the last 300 ns. The system indices are, according to Table 3.1, as follows: 1–3: Control systems; 4: Pro_2 -1 system; 5: Pro_2 -2 system; 6: Pro_5 -1 system; 7: Pro_5 -2 system; 8: Pro_6 -1 system, 9: Pro_6 -2 system, 10: Pro_7 -1 system, and 11: Pro_7 -2 system. **A.** Hydrogen bonds between chains A and B. **B.** Hydrogen bonds between chains B and C. **C.** Hydrogen bonds between chains C and D. **D.** Hydrogen bonds between chains D and E.

In the Pro_2 -1 system, the interpeptide hydrogen bonds between neighboring chains were preserved, as can be seen in Fig. 3.8 (system index 4). There were reduced hydrogen bonds between chains A and B in the Pro_2 -2 system. In the Pro_5 – 1 system, the maximum loss of β-sheet structure occurred in the 21 Ala – 30 Ala region in chain A and 17 Leu – 31 Ile region in chain E, reflected by reduced hydrogen bonds between chains A-B and D-E. This is reflected in Fig. 8 (system index 6). In the Pro_5 -2 system, the loss of β-sheet structure in the 21 Ala – 31 Ile regions of chain A is reflected by reduced hydrogen bonds between chains A and B.

In the Pro_6 -1 system the loss of β-sheet strucure in the 17 Leu – Ser 26 region in chains A, D and E is reflected by a reduction in the number of hydrogen bonds between chains A-B, C-D and D-E. In the Pro_6 -2 system the loss of β-sheet strucure in the 35 Met A – 41 Ile A region is

can be seen by a reduction in the number of hydrogen bonds between chains A and B. The disruption of the hydrogen bonding network in the terminal regions of chains B and E is reflected by the reduced A-B and D-E hydrogen bonds in the Pro_7 systems.

From Fig. 3.8, it appears that with the exception of the Pro_2 -1 and the Pro_6 -1 systems, different lengths of the proline chains considered in this study had a similar effect on disrupting the hydrogen bonding network of neighboring chains of the protofibrils.

Salt Bridges. The presence of a salt bridge increases the stability of a protein [67]. It has been hypothesized that the Asp 23 – Lys 28 salt bridge contributes strongly to the stability of the bend region and the interaction between the fibrils, and increases significantly the stability of the protofibrils especially as these grow [68]. When this salt bridge between an Asp 23 of a chain and a Lys 28 of a neighboring chain is lost, the chain dissociates from the protofibril structure [68]. We use a cut-off of 4 Å for all salt bridges [69]. Fig. 3.9 shows the average number of interpeptide Asp 23 – Lys 28 salt bridges between neighboring chains.

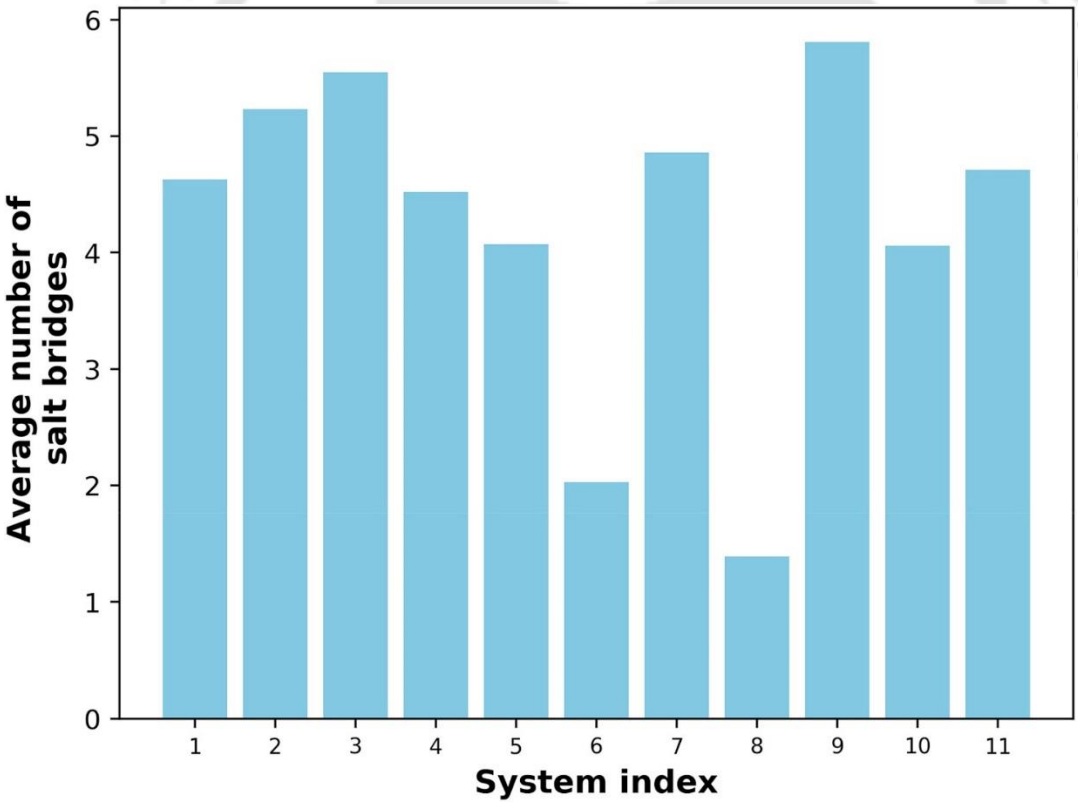


Fig. 3.9. Average number of interpeptide Asp 23 – Lys 28 salt bridges between neighboring chains. The system indices are, according to Table 3.1, as follows: 1–3: Control systems; 4: Pro_2 -1 system; 5: Pro_2 -2 system; 6: Pro_5 -1 system; 7: Pro_5 -2 system; 8: Pro_6 -1 system, 9: Pro_6 -2 system, 10: Pro_7 -1 system, and 11: Pro_7 -2 system.

In the Pro_5 systems, the proline chains were able to interact with the residues in the region of the salt-bridge forming residues 23 Asp and 28 Lys. As a result, both interpeptide and intrapeptide salt bridges were affected. The disruption or weakening of salt bridges led to the loss of backbone hydrogen bonds. The loss of these hydrogen bonds caused the loss of β-sheet structure in these systems. Loss of these interactions made it difficult for re-formation of native interactions.

In the Pro_5 – 1 system, due to electrostatic repulsion between the N-terminus proline and 28 Lys E, the salt bridge which the latter formed with 23 Asp D was lost causing chain E to move slightly away from the protofibrils. 28 Lys E re-established electrostatic interaction with 23 Asp E causing chain E to move even further apart. The N-terminus proline competed with 28 Lys D for electrostatic interaction with 23 Asp D, reducing the strength of that salt bridge. There was no electrostatic interaction between 23 Asp D and 28 Lys C. However, the salt bridge between 28 Lys D and 23 Asp C remained fairly stable in the course of the simulation. The C-terminus proline residue made strong hydrogen bonds with 28 Lys C and 28 Lys B. As a result, the salt bridge between 23 Asp C and 28 Lys B was lost and the 23 Asp C – 28 Lys B salt bridge was lost and the 23 Asp B – 28 Lys B salt bridge was weakened. The 23 Asp A – 28 Lys B salt bridge was also lost due to repulsion between the C terminus proline residue and 23 Asp A. As a result, the residues near the 23 Asp A region moved further away from the protofibril structure. These residues consequently formed the transient 3_{10} helix described previously. The loss of salt bridges also caused chain E to move slightly away from the protofibrils to form random coils and another transient 3_{10} helix. Therefore, the loss or weakening of salt bridges destabilized the protofibrils leading to the loss of hydrogen bonds in the 23 Asp – 28 Lys regions of all the chains and the loss of β-sheet structure in these regions.

In the Pro_5 – 2 system, the C terminal proline residue formed a hydrogen bond with 28 Lys A while 23 Asp A is repulsed by the N-terminal proline residue. This proline residue weakens the interaction between 23 Asp A and 28 Lys A. The region between 23 Asp A and 28 Lys A formed a turn after losing its hydrogen bonds with the corresponding region in chain B, thereby losing its β-sheet structure. Due to electrostatic repulsion between 23 Asp A and the C-terminal proline residue, the former could only make contact with 28 Lys A when it was sufficiently far away from the C-terminal proline residue, moving this region further away from the protofibril assembly. The 23 Asp C – 28 Lys D salt bridge remained fairly stable in the course of the simulations. The N-terminal proline residue competed with 28 Lys D for interactions with 28 Lys D, weakening the 23 Asp D – 28 Lys D and 23 Asp C – 28 Lys D salt bridges.

The Pro_2 systems, Pro_6 – 1 and Pro_7 systems showed reduced number of salt bridges even though they were not in the vicinity of the salt bridge forming residues. This indicates that the destabilization of the protofibrils caused by proline also destabilizes the salt bridges in these systems.

Figure 3.10 shows the maximum solvent-accessible surface areas of the systems. It can be seen that proline induces the formation of relatively more loosely packed states. In summary, the mechanism of the destabilization of the protofibrils was by the disruption of the hydrogen bonding network of the protofibrils. The disruption of hydrogen bonds lead to residues losing their β-sheet structure and consequently forming random coils. In the next section, we discuss the binding mode of the proline chains with the amyloid protofibrils.

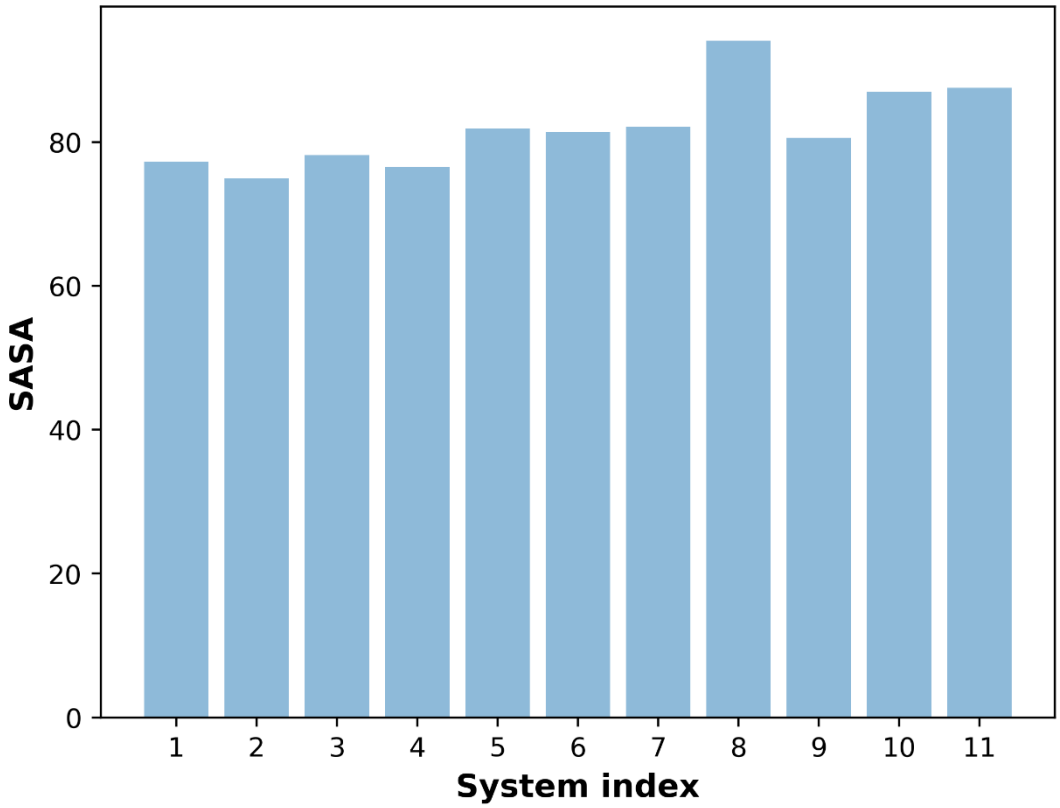


Figure 3.10. Maximum solvent accessible surface area in all the systems. The system indices are as follows: 1 – 3: Control systems; 4: Pro_2 -1 system; 5: Pro_2 -2 system; 6: Pro_5 -1 system; 7: Pro_5 -2 system; 8: Pro_6 -1 system, 9: Pro_6 -2 system, 10: Pro_7 -1 system, and 11: Pro_7 -2 system.

3.3.2. Binding mode

Proline ligands preferably bind to charged regions in Aβ protofibrils. Docking produced complexes in which the proline ligands were close to both the N- and C- termini of the protofibrils or close to the charged residues Asp 23 and Lys 28. The oligoproline chains showed strong binding affinity to the amyloid protofibrils and did not move away from the protofibrils in the course of the simulations. Because of the polarization effect in water, the N-terminus of both the protofibrils and ligands have positive charges and the C-termini have negative charge. The first and last oligoproline residues bind to regions corresponding to charged regions in the amyloid protofibrils: the terminii or the charged residues Glu 22, Asp 23 and Lys 28. Proline also makes a few hydrophobic contacts with protofibril residues. These observations were confirmed by the MM-PBSA calculations. Table 3.2 shows a summary of the binding energies.

Table 3.2. MMPBSA results. Energies are expressed in kJ mol^{-1}.

System	van der Waal energy	Electrostatic energy	Polar solvation energy	SASA energy	Binding energy
Pro_2-1	−30.419 ± 1.267	−550.669 ± 5.803	294.485 ± 5.399	−7.819 ± 0.080	−294.565 ± 2.155
Pro_2-2	−46.563 ± 1.219	−403.599 ± 4.348	242.158 ± 2.241	−9.173 ± 0.051	−217.033 ± 3.342
Pro_5-1	−225.078 ± 1.691	−897.166 ± 5.883	635.643 ± 2.938	−26.191 ± 0.085	−512.953 ± 3.754
Pro_5-2	−235.493 ± 1.673	−703.045 ± 16.693	578.654 ± 7.942	−27.184 ± 0.102	−386.927 ± 9.188
Pro_6 -1	−190.327 ± 1.371	−712.929 ± 5.598	552.153 ± 3.983	−20.363 ± 0.108	−371.798 ± 3.435
Pro_6 -2	−153.949 ± 1.955	−1085.651 ± 9.062	755.308 ± 6.343	−21.153 ± 0.153	−504.831 ± 4.303
Pro_7-1	−69.504 ± 1.438	−1191.693 ± 9.487	790.157 ± 5.601	−16.456 ± 0.150	−487.396 ± 4.656
Pro_7-2	−193.132 ± 2.334	−698.083 ± 21.263	629.498 ± 13.140	−26.154 ± 0.153	−288.323 ± 8.277

Table 3.2 summarizes the results of the MMPBSA calculations. Clearly, the energy contributions are overestimated. There are a few reasons for this. Firstly, the amyloid protofibrils have a net charge of −5. They contain charged termini 17 Leu and 42 Ala, in addition to the charged salt-bridge forming residues 22 Glu, 23 Asp, and 28 Lys. Also the oligoproline N- and C-termini are charged. Secondly, there are large conformational changes in the amyloid protofibrils as a result of interaction with the oligoproline chains, which contributes to inaccurate MMPBSA predictions [62]. Subsequently, the use of a higher and non-uniform dielectric constant for the protein is required to account for the charged residues and large conformational changes [70]. This is a drawback of the MMPBSA method, which is sensitive

to input parameters and may overestimate the energy contributions by 2–5 times the real value [62]. In previous works, values of dielectric constants used for protein solutes ranged from 1 to 40, depending on the system being studied [70]. The standard dielectric constant used is epsilon = 4 [70]. We have used the standard value of the dielectric constant for a relative comparison of the binding energies of the systems being considered.

Fig. 3.11 shows the electrostatic potential map of the residues of the amyloid beta protofibrils, which are within 5 Å of the oligoproline chains and the electrostatic potential map of the oligoproline chains. Proline residues make many contacts with charged protofibril residues as can be seen in Figures 3.12- 3.19. The electrostatic interaction dominates in all the systems because of the interaction between charged residues of the amyloid protofibrils and the oligoproline chains. A plot of the contribution energy of individual residues to the binding energy is shown in Fig. 3.20 illustrates this. The electrostatic interactions dominate over the polar solvation energies, and contribute the most to the $\Delta G_{binding}$ being negative. Apolar solvation energies were fairly uniform. The vdW energies were the highest for the Pro_5 systems because the proline chains were in the inner concave hydrophobic region of the protofibrils making the most contacts of the systems considered. The vdW energies were also high for the Pro_6 and Pro_7 -2 systems because they made 4 hydrophobic contacts with the protofibrils. The entropic contributions to the binding energies were not calculated, although studies suggest that this may be low due to the inherent rigidity of the proline residues [71].

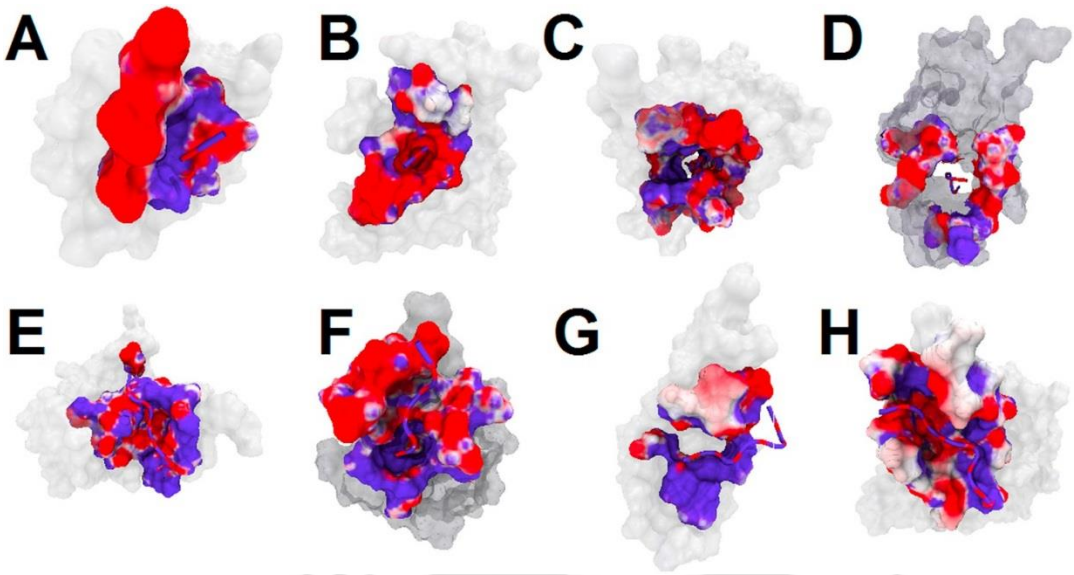


Fig. 3.11. Electrostatic contact map of the residues of the protofibrils which contain atoms that are within 5 Å of the oligoproline chains and the electrostatic contact map of the oligoproline chains. These residues are colored red if negatively charged and blue if positively charged. The protofibrils are represented as a surface and the proline chains as tubes. Residues which contain atoms beyond 5 Å of the proline residues are shown as a transparent surface. **A.** Pro_2 -1 system **B.** Pro_2-2 system **C.** Pro_5-1 system **D.** Pro_5-2 system **E.** Pro_6-1 system **F.** Pro_6 -2 system **G.** Pro_7 -1 system **H.** Pro_7 -2 system.

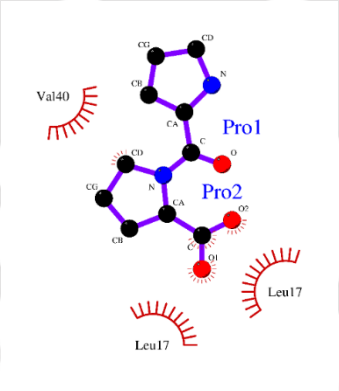


Figure 3.12. Contacts made by the proline chain in the Pro_2 -1system. Residues which make contact with proline are shown with red spokes.

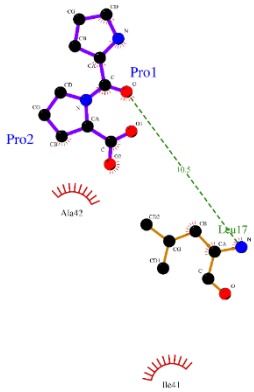


Figure 3.13. Contacts made by the proline chain in the Pro_2 -2 system. Residues which make contact with proline are shown with red spokes.

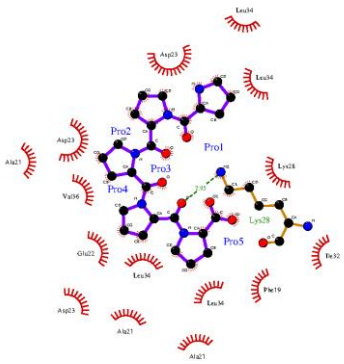


Figure 3.14. Contacts made by the proline chain in the Pro_5 -1 system. Residues which make contact with proline are shown with red spokes.

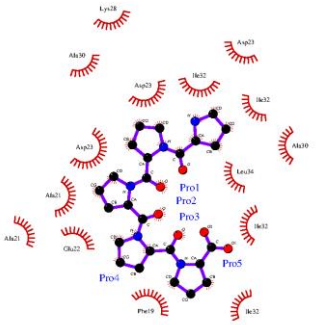


Figure 3.15. Contacts made by the proline chain in the Pro_5 -2 system. Residues which make contact with proline are shown with red spokes.

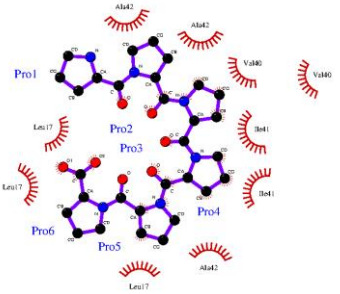


Figure 3.16. Contacts made by the proline chain in the Pro_6 – 1 system. Residues which make contact with proline are shown with red spokes.

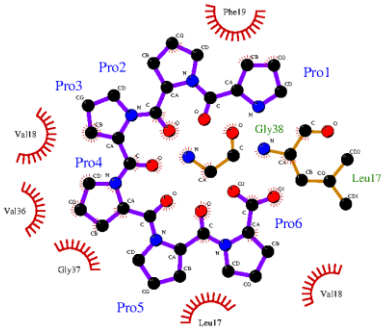


Figure 3.17. Contacts made by the proline chain in the Pro_6 – 2 system. Residues which make contact with proline are shown with red spokes.

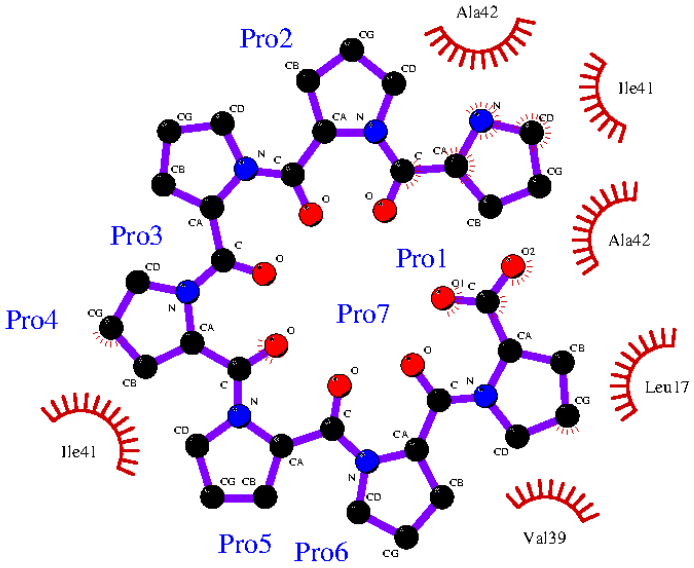


Figure 3.18. Contacts made by the proline chain in the Pro_7 -1 system. Residues which make contact with proline are shown as red spokes.

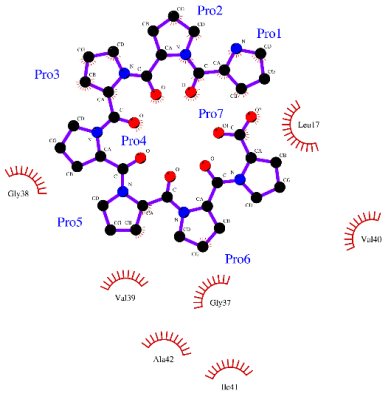


Figure 3.19. Contacts made by the proline chain in the Pro_7 -1 system. Residues which make contact with proline are shown as red spokes.

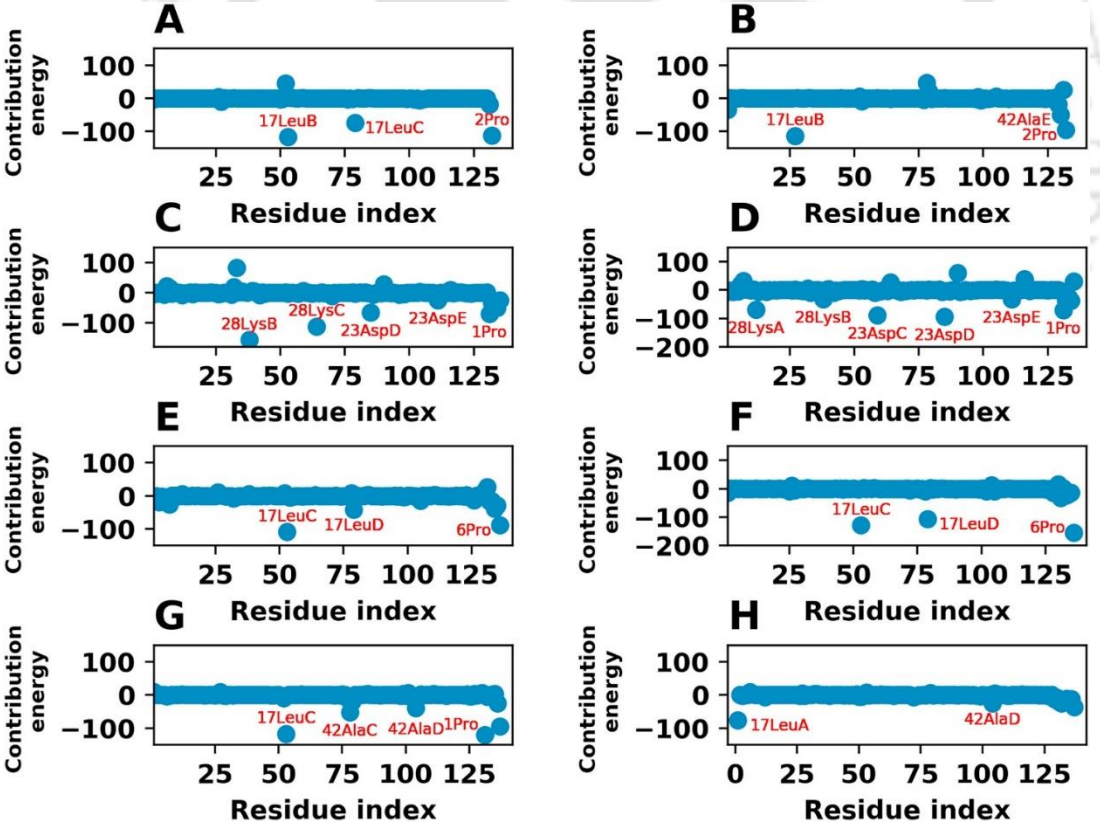


Figure 3.20. Individual residue contribution to the total binding energy expressed in kJ mol^{-1}. Charged residues which contribute to the binding have negative values. Some residues also have positive values, these are the instances when there are repulsive forces between similarly charged residues. Key residues which contribute the most to binding are indicated with a label

in red **A.** Pro_2 -1 system **B.** Pro_2-2 system **C.** Pro_5-1 system **D.** Pro_5-2 system **E.** Pro_6-1 system **F.** Pro_6 -2 system **G.** Pro_7 -1 system **H.** Pro_7 -2 system.

It can be seen in Table 3.2 that the binding energies for the Pro_5, Pro_6 and Pro_7 -1 systems do not vary significantly, indicating that increasing the length of the proline chains may not affect the binding affinity of the proline. The proline chains in these systems are sufficiently long enough to make many favorable contacts with the protofibrils, which was not possible for the proline chains in the Pro_2 systems.

Hydrogen bonds formed between the oligoproline chains and the amyloid protofibrils. Hydrogen bonds between receptor and drug molecules are an important criterion for binding affinities [72]. We observed that hydrogen bonds were formed between charged residues of the oligoproline and protofibril chains. Since hydrogen bonds between charged residues are stronger than those between one or two uncharged residues, hydrogen bonds between these protofibril residues with those in a neighboring protofibril chain were lost leading to a loss in beta sheet structure. We note that the N-termini of the oligoproline and amyloid protofibril chains are positively charged and the C-termini are negatively charged. Fig. 3.21 shows the running average of the number of hydrogen bonds over 100 windows formed between oligoproline chains and the $A\beta_{17-42}$ protofibrils.

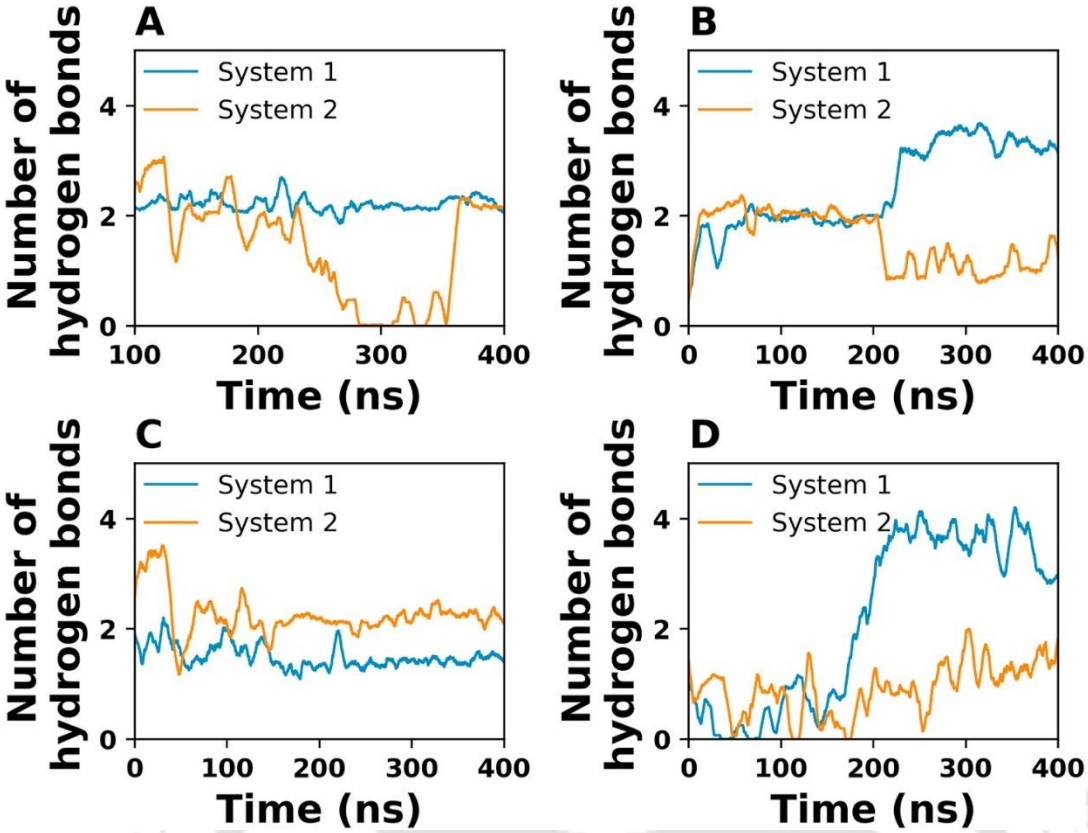


Fig. 3.21. Running average of the number of hydrogen bonds over 100 windows formed between oligoproline chains and the $A\beta_{17-42}$ protofibrils **A.** Pro_2 systems. **B.** Pro_5 systems. **C.** Pro_6 systems. **D.** Pro_7 systems.

The negatively charged C-terminal residue in the Pro_2 -1 system formed strong hydrogen bonds with the positively charged N-terminal 17 Leu of chains C and D at an average of 1.07 and 0.942 respectively, indicating that these hydrogen bonds were very stable in the course of the simulation due to the close proximity of the C-terminal Pro_2 residue with 17 Leu C (average distance 0.19 nm) and 17 Leu D (average distance 0.17 nm). In the Pro_2 -2 system, the proline chain made weak hydrogen bonds with 17 Leu C (average 0.153), 17 Leu D (average 0.366), 42 Ala D (average 0.23) and 42 Ala E (average 0.16).

In the Pro_5- 1 system, the C-terminal proline residue made strong hydrogen bonds with 28 Lys C (average 0.991 bonds) and 28 Lys B (average 0.761 bonds). The N-terminal residue of proline made hydrogen bonds with 23 Asp D with the average number of hydrogen bonds at 0.943, and their distance was less than 0.20 nm in the course of the simulation indicating a very strong hydrogen bond. We note that 28 Lys is positively charged and 23 Asp is negatively charged and together these contribute to the formation of salt bridges in the protofibrils. By forming strong hydrogen bonds with proline these salt bridges are lost thereby reducing the stability of the

protofibrils. Initially, hydrogen bonds existed between the Asp 23 residue in chain E (Asp 23 E) and the N-terminus proline residue, and between Asp 23 D and Asp 23 E. Asp 23 D was closer to the first proline residue, and for most of the course of the simulation was at a minimum distance of less than 0.20 nm away. The hydrogen bond formed between the charged residues Asp 23 D and the first proline residue was hence very strong, and stable throughout the course of the simulation. After the formation of this strong hydrogen bond, the 23 Asp D - 23 Asp E hydrogen bond was lost. After about 66ns, the residues Asp 23 E, 24 Val, 25 Gly and 26 Ser of chain E lost their β-sheet structure, and also lost their hydrogen bonds with the corresponding region in chain D. The loss of these hydrogen bonds led to increased gain in mobility of this region, and Asp 23 E gradually moved away from the N-terminus proline residue and lost its hydrogen bond with it. This charged residue Asp 23 E, along with 24 Val, 25 Gly and 26 Ser formed a turn and eventually a 3_{10} helix after the formation of 23 Asp E − 26 Ser E and 22 Glu E − 25 Gly E hydrogen bonds. The formation of the helix brought the Asp 23 E residue again in close contact with the first proline residue briefly, and hydrogen bonds were again formed between them albeit for a short period of time until the 3_{10} helix collapsed into a random coil along with residues 18–22. The 22–25 residues of chain E then formed a beta hairpin with the residues 32–36 of the same chain with the region 17–22 making only two hydrogen bonds with its corresponding region in chain D.

In the Pro_5- 2 system, hydrogen bonds were formed between 23 Asp D and the N-terminal proline residue (average 0.542); and between 28 Lys A and the C-terminal proline residue (average 0.5).

In the Pro_6- 1 system, a strong hydrogen bond was formed between 17 Leu C and the C-terminal proline residue (average 0.834) and a relatively less stable hydrogen bond between 38 Gly E and the second proline residue (average 0.332). In the Pro_6- 2 system, the C-terminal proline residue formed strong hydrogen bonds with 17 Leu C (average 1.024 bonds) and 17 Leu D (average 0.972 bonds).

In the Pro_7 systems, as many as four hydrogen bonds were formed between the proline ligand and the protofibrils. The N-terminal proline residue formed hydrogen bonds with 42 Ala C (average 0.704 bonds) and 42 Ala D (average 0.847 bonds) which are the C-termini of chains C and D respectively. The C-terminus proline residue formed hydrogen bonds with 17 Leu C (average 0.721 bonds) and 17 Leu D (average 0.357 bonds).

Hydrophobic contacts. The oligoproline chains also made hydrophobic contacts with the protofibrils. The residues with hydrophobic contacts were made are 40 Val, 34 Leu, 36 Val, 21 Ala, 19 Phe, 30 Ala, 32 Ile, 37 Gly, 18 Val, 41 Ile, and 39 Val. These interactions are illustrated

in Fig. 3.12 – 3.19. As can be expected, the Pro_5 systems had the most number of hydrophobic contacts as the proline chains were in the inner hydrophobic core of the protofibrils.

3.4. Conclusions

Our results are meaningful because protofibrils lose their beta sheet structure and bind well to proline, which was expected. The loss of β-sheet structure in residues of the protofibrils lead to the formation of random coils and turns. This is shown by reduced percentage of β-sheets and increased percentage of random coils. Increased numbers of random coils and turns leads to increased detachment of individual chains. Loss of hydrogen bonds of charged residues with residues in a neighboring, followed by loss of β-sheet structure and the formation of random coils and turns meant that these residues could form new hydrogen bonds with residues in the same chain. This accounts for the random coils forming transient 3_{10} helices in regions containing charged residues. Loss of beta-sheet structure is by loss of hydrogen bonds. Proline aids in the loss of these hydrogen bonds by itself making more favorable ionic hydrogen bonds and contacts with the charged residues of the protofibrils. The loss of critical salt bridges between 23 Asp and 28 Lys of neighboring chains lead to the further loss of protofibril stability, even in cases where proline chains were not in the vicinity of these residues (Pro_6 and Pro_7 systems). Interactions, which involve residues 41 and 42 are necessary for the formation of stable fibrils [73]. By interacting with these residues, proline further destabilizes the protofibrils. Amyloid protofibrils become more loosely packed on account of increased maximum SASA.

A study on the binding energy of all possible combinations of amino acid tripeptides with the Aβ fibrils showed that proline always enhances binding [69]. Therefore it was expected that proline binds well to the protofibrils. MMPBSA results show that electrostatic interactions dominate over all other interactions which can be expected because the docking algorithm produced structures in which the oligoproline chains were within 5 Å of several charged residues, and the subsequent ionic interactions between oppositely charged residues lead to very strong binding. The binding energies of the Pro_5, Pro_6 and Pro_7 -1 systems do not vary significantly indicating that increasing the length of the proline chain may not significantly improve the binding.

To summarize, using docking, MD simulations and MMPBSA, we showed the mechanism in which the β-sheet breaker oligoproline chains break the β-sheet structure of amyloid-β protofibrils and induce the formation of random coils. Critical interactions between protofibril

chains such as hydrogen bonds and salt bridges were disrupted. Proline binds strongly to the protofibrils by electrostatic interactions. These factors make proline an important amino acid to consider in the design of novel peptide-based drugs.

References

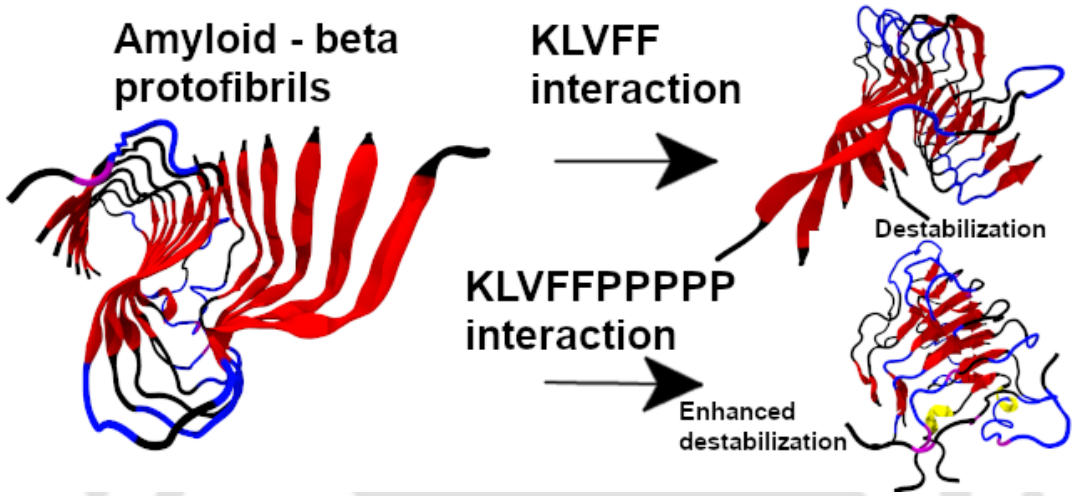
1. H. Querfurth, F. LaFerla, Alzheimer's disease: Mechanism of disease, N. Engl. J. Med. 362 (2010) 329-344.
2. C.A. Ross, M.A. Poirier, Protein aggregation and neurodegenerative disease. Nat. Med. 10 (2004) S10.
3. D. Selkoe, Alzheimer's disease: Genes, proteins and therapy, Physiol. Rev. 81 (2001) 66.
4. W. Pardridge, The blood-brain barrier: Bottleneck in brain drug development, NeuroRx. 2 (2005) 3–14.
5. J. Lemkul, D. Bevan, The role of molecular simulations in the development of inhibitors of amyloid β-peptide aggregation for the treatment of Alzheimer's disease, ACS Chem. Neurosci. 3 (2012) 845– 856.
6. K. Ono, Y. Yoshiike, A. Takashima, K. Hasegawa, H. Naiki, M. Yamada, Potent Anti-Amyloidogenic and Fibril-Destabilizing Effects of Polyphenols in Vitro: Implications for the Prevention and Therapeutics of Alzheimer's Disease, J. Neurochem. 87(1) (2003) 172– 181
7. C. Riviere, J.C. Delaunay, F. Immel, C. Cullin, J.P. Monti, The Polyphenol Piceid Destabilizes Preformed Amyloid Fibrils and Oligomers in Vitro: Hypothesis on Possible Molecular Mechanisms, Neurochem. Res. 34(6) (2009) 1120– 1128
8. J.E. Kim, M. Lee, Fullerene inhibits beta-amyloid peptide aggregation, Biochem Biophys Res Commun. 303(2) (2003) 576–9.
9. X. Zhou, W. Xi, Y. Luo, S. Cao, G. Wei, Interactions of a water-soluble fullerene derivative with amyloid-β protofibrils: dynamics, binding mechanism, and the resulting salt-bridge disruption, J Phys Chem B. 118(24) (2014 Jun 19) 6733-41
10. G. Zhang, M.J. Leibowitz, P.J. Sinko, S. Stein, Multiple-peptide conjugates for binding beta-amyloid plaques of Alzheimer's disease, Bioconjug Chem. 14(1) (2003 Jan-Feb) 86-92
11. R.X. Gu, H. Gu, Z.Y. Xie, J.F. Wang, H.R. Arias, D.Q. Wei, K.C. Chou, Possible Drug Candidates for Alzheimer's Disease Deduced from Studying Their Binding Interactions with Alpha7 Nicotinic Acetylcholine Receptor Med. Chem. 5(3) (2009) 250– 262
12. M. Tang, Z. Wang, Y. Zhou, W. Xu, S. Li, L. Wang, D. Wei, Z. Qiao, A Novel Drug Candidate for Alzheimer's Disease Treatment: Gx-50 Derived from Zanthoxylum Bungeanum, J. Alzheimer's Dis. 34(1) (2013) 203– 213
13. H.M. Fan, R.X. Gu, Y.J. Wang, Y.L. Pi, Y.H. Zhang, Q. Xu, D.Q. Wei, Destabilization of Alzheimer's Aβ42 Protofibrils with a Novel Drug Candidate wgx-50 by Molecular Dynamics Simulations, J. Phys. Chem. B. 119 (34) (2015) 11196 -202
14. S. Hou, R.X. Gu, D.Q. Wei, Inhibition of β-amyloid Channels with a Drug Candidate wgx-50 Revealed by Molecular Dynamics Simulations, J Chem Inf Model. 57(11) (2017) 2811-2821
15. A.C. Kaushik, A. Kumar, Zhennan Peng, A. Khan, M. Junaid, A. Ali, S. Bharadwaj, D.Q. Wei, Evaluation and validation of synergistic effects of amyloid-beta inhibitor–gold nanoparticles complex on Alzheimer's disease using deep neural network approach, J. Mater. Res. 34 (2019) 1845 -1853

16. M. Matsuoka, Humanin; a defender against Alzheimer's disease? Recent Pat CNS Drug Discov. 4(1) (2009) 37-42.
17. D. Craik, D. Fairlie, S. Liras, D. Price, The future of peptide-based drugs, Chem Biol Drug Des. 81 (2013) 136–147.
18. L. Otvos Jr., M. Cudic, B.Y. Chua, G. Deliyannis, D.C. Jackson, An insect antibacterial peptide-based drug delivery system, Mol Pharm. 1 (3) (2004) 220-232.
19. A. Gladkevich, F. Bosker, J. Korf, K. Yenkoyan, H. Vahradyan, M. Aghajanov, Prog. Neuro-Psychopharmacol. Biol. Psychiatry (2007) 31, 1347– 1355.
20. A. Bilikiewicz, W. Gaus, Colostrinin (a naturally occurring, proline-rich, polypeptide mixture) in the treatment of Alzheimer's disease, J Alzheimers Dis. 6(1) (2004)17-26.
21. J. Leszek, A. Inglot, M. Janusz, J. Lisowski, K. Krukowska, J. Georgiades, Colostrinin: A proline-rich polypeptide (PRP) complex isolated from ovine colostrums for treatment of Alzheimer's disease, Arch Immunol Ther Exp. 47 (1999) 377-85.
22. K. Yenkoyan, K. Safaryan, V. Chavushyan, I. Meliksetyan, G. Navasardyan, J. Sarkissian, A. Galoyan, M. Aghajanov, Neuroprotective action of proline-rich polypeptide-1 in β-amyloid induced neurodegeneration in rats, Brain Res. Bull. 86 (2011) 262-271.
23. P. Chou, G. Fasman, Prediction of protein conformation, Biochemistry 13 (1973) 222-45.
24. S. Wood, R. Wetzel, J. Martin, M. Hurle, Prolines and amyloidogenicity in fragments of the Alzheimer's peptide beta/A4, Biochemistry 34 (1995) 724-730.
25. T. Herning, K. Yutani, K. Inaka, R. Kuroki, M. Matsushima, M. Kikuchi, Role of proline residues in human lysozyme stability: A scanning calorimetric study combined with X-ray structure analysis of proline mutants, Biochemistry 31 (1992) 7077-85.
26. S. Rauscher, S. Baud, M. Miao, F.W. Keeley, R. Pomès, Proline and glycine control protein self-organization into elastomeric or amyloid fibrils, Structure 14 (2006) 1667-1676.
27. A. Tatham, P. Shewry, Elastomeric proteins: Biological roles, structures and mechanisms. Trends Biochem Sci. 25 (2000) 571.
28. M. Karpen, P. De Haseth, K. Neet, Differences in the amino acid distributions of 3_{10}-helices and α-helices. Protein Sci. 1(10) (1992) 1333–1342.
29. M. Sundaralingam, Y. Sekharudu, Water-inserted alpha-helical segments implicate reverse turns as folding intermediates, Science 244 (1989) 1333–1337.
30. J. Tirado-Rives, W. Jorgensen, Molecular dynamics simulations of the unfolding of an alpha-helical analog of ribonuclease A S-peptide in water, Biochemistry 30(24) (1991) 3864–3871.
31. D. Tobias, C. Brooks, Thermodynamics and mechanism of α helix initiation in alanine and valine peptides, Biochemistry 30(24) (1991) 6059–6070.
32. R. Armen, D. Alonso, V. Daggett, The role of α-, 3_{10}-, and π-helix in helix→coil transitions, Protein Sci. 12 (2003) 1145–1157.
33. C. Rohl, A. Doig, Models for the 3(10)-helix/coil, pi-helix/coil, and alpha-helix/3(10)-helix/coil transitions in isolated peptides, Protein Sci. 5(8) (1996)1687–1696.
34. R. Harrison, P. Sharpe, Y. Singh, D. Fairlie, Amyloid peptides and proteins in review, Rev Physiol Biochem Pharmacol. 159 (2007) 1-77.
35. Z. Zhang, Z. Hu, D. Zhao, Y. Chen, Y. Li, Helices with rational residues conduct different modulations towards Aβ aggregation, Chem. Lett. 46(7) (2017) 979-982.
36. T. Lührs, C. Ritter, M. Adrian, D. Riek-Loher, B. Bohrmann, H. Döbeli, D. Schubert, R. Riek, 3D structure of Alzheimer's amyloid-β(1-42) fibrils, Proc. Natl. Acad. Sci. U.S.A 102 (48) (2005) 17342-17347.

37. J. Lemkul, D. Bevan, Assessing the stability of Alzheimer's amyloid protofibrils using molecular dynamics, J. Phys. Chem. B. 114 (2010) 1652–1660.
38. M. Masman, U. Eisel, I. Csizmadia, B. Penke, R. Enriz, S. Marrink, P. Luiten, In silico study of full-length amyloid-β 1-42 tri- and penta-oligomers in solution, J. Phys. Chem. B. 113 (2009) 11710–11719.
39. M. Hanwell, D. Curtis, D. Lonie, T. Vandermeerschd, E. Zurek, G. Hutchison, Avogadro: An advanced semantic chemical editor, visualization, and analysis platform, *J.* Cheminform. 4:17 (2012)
40. D. Kozakov, D. Hall, B. Xia, K. Porter, D. Padhorny, C. Yueh, D. Beglov, S. Vajda, The ClusPro web server for protein-protein docking, Nat. Protoc. 12(2) (2017) 255-278.
41. D. Kozakov, D. Beglov, T. Bohnuud, S. Mottarella, B. Xia, D. Hall, S. Vajda, S, How good is automated protein docking? Proteins 81(12) (2013) 2159-66.
42. D. Kozakov, R. Brenke, S. Comeau, S. Vajda, PIPER: An FFT-based protein docking program with pairwise potentials, Proteins: Struct., Funct., Bioinf. 65 (2006) 392-406.
43. S. Comeau, D. Gatchell, S. Vajda, C. Camacho, ClusPro: An automated docking and discrimination method for the prediction of protein complexes, Bioinformatics 20 (2004) 45-50.
44. S. Comeau, D. Gatchell, S. Vajda, C. Camacho, ClusPro: A fully automated algorithm for protein-protein docking, Nucleic Acids Res. *32* (2004) W96-W99.
45. J. Janin, Assessing predictions of protein-protein interaction: The CAPRI experiment, Protein Sci. 14 (2005) 278-283.
46. S. Gerben, J. Lemkul, A. Brown, D. Bevan, Comparing atomistic molecular mechanics force fields for a difficult target: A case study on the alzheimers amyloid β-peptide. J Biomol Struct Dyn. 32 (2014) 1817-1832.
47. S. Nosé, A unified formulation of the constant temperature molecular dynamics methods. J. Chem. Phys. 81 (1984) 511-519.
48. W.G. Hoover, Canonical dynamics: Equilibrium phase-space distributions. Phys. Rev. A: At., Mol., Opt. Phys. 31 (1985) 1695-1697.
49. M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method, J. Appl. Phys. 52 (1981) 7182-7190.
50. S. Nosé, M.L. Klein, Constant pressure molecular dynamics for molecular systems. Mol. Phys. 50. (1983) 1055-1076
51. M. Abraham, T. Murtola, R. Schulz, S. Páll, J. Smith, B. Hess, E. Lindah, Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX 1-2 (2015) 19-25.
52. B. Hess, P-LINCS: A parallel linear constraint solver for molecular simulation, J. Chem. Theory Comput. 4 (2008) 116-122.
53. U. Essmann, L. Perera, M. Berkowitz, T. Darden, H. Lee, L. Pedersen, A smooth particle mesh ewald method. J. Chem. Phys. 103 (1995) 8577-8593.
54. W.G. Touw, C. Baakman, J. Black, T.A. te Beek, E. Krieger, R.P. Joosten, G. Vriend, A series of PDB related databases for everyday needs, Nucleic Acids Res. 43 (2015) D364-D368.
55. W. Kabsch, C. Sander, Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features, Biopolymers 22 (1983) 2577-2637.
56. D. van der Spoel, P.J. van Maaren, P. Larsson, N. Timneanu, Thermodynamics of hydrogen bonding in hydrophilic and hydrophobic media. J. Phys. Chem. B 110 (2006) 4393-4398
57. W. Humphrey, A. Dalke, K. Schulten, VMD: Visual molecular dynamics. J Mol Graph. 14 (1996) 33-38.

58. Wallace A C, Laskowski R A, Thornton J M (1996). LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. Protein Eng., 8, 127-134.
59. Eisenhaber F, Lijnzaad P, Argos P, Sander C, & Scharf M (1995) J. Comput. Chem. 16, 273-284
60. J. Srinivasan, T. E. Cheatham, P. Cieplak, P. A. Kollman, D. A. Case, Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate - DNA helices, J. Am. Chem. Soc. 120 (1998) 9401–9409.
61. P. A. Kollman, I Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D.A. Case, T. E. Cheatham 3rd, Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models, Acc Chem Res. 33(12) (2000) 889-97.
62. R. Kumari, R. Kumar, A. Lynn, g_mmpbsa - A GROMACS tool for high-throughput MM-PBSA calculations. J. Chem. Inf. Model. 54 (2014) 1951-1962.
63. N.A. Baker, D. Sept, S. Joseph, M.J. Holst, J.A. McCammon, Electrostatics of nanosystems: Application to microtubules and the ribosome, Proc. Natl. Acad. Sci. USA 98 (2001) 10037-10041.
64. M. Karpen, P. De Haseth, K. Neet, Differences in the amino acid distributions of 3_{10}-helices and α-helices, Protein Sci. 1(10) (1992) 1333–1342.
65. S. Vivekanandan, J. Brender, S. Lee, A. Ramamoorthy, A partially folded structure of amyloid-beta(1-40) in an aqueous environment, Biochem. Biophys. Res. Commun. 411 (2011) 312-316.
66. A. Baumketner, S. Bernstein, T. Wyttenbach, N. Lazo, D. Teplow, M. Bowers, J. Shea, Structure of the 21-30 fragment of amyloid β-protein. Protein Sci. 15 (2006) 1239-1247.
67. S. Kumar, C. Tsai, B. Ma, R. Nussinov, Contribution of salt bridges toward protein thermostability, J Biomol Struct Dyn. 17 (2000) 79-85.
68. J. Lemkul, D. Bevan, Assessing the stability of Alzheimer's amyloid protofibrils using molecular dynamics, J. Phys. Chem. B. 114 (2010) 1652–1660.
69. D. Barlow, J. Thornton, Ion-pairs in proteins, J. Mol. Biol. 168 (1983) 867-885.
70. L. Li, C. Li, Z. Zhang, E. Alexov, On the Dielectric "Constant" of Proteins: Smooth Dielectric Function for Macromolecular Modeling and Its Implementation in DelPhi, J Chem Theory Comput. Apr 9;9(4) (2013) 2126-2136.
71. M.H. Viet, K. Siposova, Z. Bednarikova, A. Antosova, T.T. Nguyen, Z. Gazova, M.S. Li, In Silico and in Vitro Study of Binding Affinity of Tripeptides to Amyloid β Fibrils: Implications for Alzheimer's Disease, J. Phys. Chem. B. Apr 23 119(16) (2015) 5145-55.
72. M. Nocker, S. Handschuh, C. Tautermann, K.R. Liedl, Theoretical Prediction of Hydrogen Bond Strength for Use in Molecular Modeling, J. Chem. Inf. Model 49 (9) (2009) 2067-2076.
73. R. Tycko, Alzheimer's disease: Structure of aggregates revealed, Nature 537 (2016) 492-493.

Chapter 4. Destabilization of the Alzheimer's Amyloid-β Peptide by a Proline-rich β-sheet Breaker Peptide: A Molecular Dynamics Simulation study



4.1 INTRODUCTION

Alzheimer's disease belongs a category of neurodegenerative diseases which is a leading cause of dementia affecting millions of people worldwide, leading to progressive memory loss and cognitive deterioration [1, 2]. This disease is induced by the irregular processing of neuronal proteins [3]. A distinctive feature of this disease is the presence of neuritic plaques in infected brains which are primarily composed of the amyloid fibrils [2, 4]. Neurofibrillary tangles composed of tau proteins are also found in these brains [2, 4]. As a remedy for this disease, molecules with low molecular weights which can cross the blood-brain barrier had been proposed to destabilize the amyloid fibrils and prevent their aggregation [2]. However, since 2003 there have been no new drugs which had been approved, implying that the therapy for this disease is a challenge [5].

In the quest for small molecules which can potentially be therapeutic for Alzhemier's disease, several types of molecules were considered. Chelates and metal coordinations complexes such as iridium (III) and rhodium (III) metal complexes, clioquinol – zinc ion

complex, platinum phenanthroline derivatives and Cobalt (III) Schiff bases could inhibit the aggregation of fibrils and diminish their toxicity [6 - 9]. However, in clinical trials the metal chelators PBT1 and PBT2 failed to demonstrate favourable effects [10].

Synthetic compounds such as hexahydropyrroloindoles (HPI), flavone hybrids, triazole-based derivatives, and epigallocatechin-3-gallate (EGCG) could inhibit the aggregation of the fibrils [11 – 13]. Another class of molecules considered for therapy were antibodies, which displayed insubstantial outcomes due to their large molecular weight as a result of which they were unable to cross the blood-brain barrier [14]. Bapineuzumab and Solanezumab failed in clinical trials [15 – 17]. However, in patients with mild symptoms, Aducanumab displayed some therapeutic potential [18].

However, most small molecule drugs have some drawbacks. In general, they have a low selectivity for targets, and a poor affinity for the fibrils [19, 20]. Peptide-based drugs have been proposed as an alternative, due to their ability to have an increased number of interactions with their targets, thereby improving their specificity [20]. Presently, over 100 peptide-based drugs have a market share of about 10% in the ethical pharmaceutical market [20]. They occur frequently as a sequence of 8 - 10 residues [20]. Apart from their superior specificity, peptide-based drugs are less toxic, possess chemical and biological variety and tend to not gather in tissues [20]. There are some challenges associated with peptide-based drugs. Apart from being metabolically unstable, they have trouble crossing membranes easily [20]. They are mostly delivered by injection as they have poor oral availability [20]. Also, they are associated with inferior solubility and are expensive [20]. If these disadvantages can be prevailed over, peptide-based drugs can show great promise in treating diseases effectively [20].

Peptide sequences obtained from the $A\beta_{1-42}$ peptide can inhibit the aggregation of the $A\beta_{1-42}$ peptides, such as $A\beta_{15-22}$, $A\beta_{16-23}$, and $A\beta_{17-24}$ [21]. The sequence $A\beta_{16-20}$ was recognized as the shortest sequence which initiated the nucleation of the fibrils [22]. This sequence is known as the self-recognition sequence of the amyloid fibrils, and can inhibit the fibrillation process. It was also observed that when the sequence $A\beta_{17-21}$ was modified by the substitutions V17P and A22D, and was injected in the brains of infected rats, it could destabilize the fibrils and arrest their aggregation [23]. Based on these observations, diverse sequences based on the $A\beta_{16-20}$ sequence (KLVFF) were investigated such as adding lysine residues to the KLVFF sequence [24], the $A\beta_{1-28}$ sequence in which the side chains of the residues 17-21 were linked by a lactam linkage and conformationally restrained [25], and the sequences RGKLVFFGR and RGKLVFFGR-NH2 [26]. Other examples of such modifications are the molecules PI-368, PPI-

433, PPI-457 [27], polyamine modifications [28, 29], N-methylated sequences [30-34], substitution by D-amino acids [35, 36], linkage to aminoethoxy amide and aspartate [37], the molecules SEN 606, AMY1, AMY2, and K4 [38 – 40]. The proteolytic stability of the KLVFF peptide was improved by its conjugation with PEG [41].

Amino acids which contain aromatic rings bind to the fibrils with the highest affinity, according to a study by Viet *et al.* [42]. There may be an association between the binding affinity of peptide sequences to amyloid fibrils and their capacity to destabilize amyloid fibrils [42]. Although proline has a five-membered aromatic ring, it has a high binding affinity to the fibrils [42]. Proline has attracted interest in destabilizing amyloid fibrils. Proline-rich polypeptides such as Colostrinin have shown promise in the treatment of Alzheimer's disease by preventing the aggregation of the fibrils in vitro and in clinical studies [43 – 45]. Another peptide, PRP – 1, which contains four proline residues had shown promise [46]. Proline is a unique residue in the sense that it rarely occurs in beta-sheets [47]. The peptidyl-prolyl bond in proline has a conformation which is not complementary to the geometries of the peptide bond in beta-sheets [48]. The aromatic ring in proline is unable to participate in the hydrogen bonding network of beta-sheets [48]. The internal rotations of the protein backbone depend on the side chains of residues [49]. Proline is a rigid amino acid owing to its pyrrolodine ring which restricts it to accessing fewer conformations [50]. It was shown that when any amino acid in the amyloidogenic sequence LVFFAED was replaced by proline, the sequence became soluble and unable to form fibrils [48].

In the present study, we investigated the ability of a peptide KLVFFPPPPP to destabilize the amyloid protofibrils. In this peptide, denoted as KLVFFP5, the KLVFF sequence was linked to a proline pentamer. This peptide seeks to combine the self-identification and inhibitory property of the KLVFF sequence with the beta-sheet breaker ability of the proline sequence. Five repeat units of proline were selected based on the observation of Murphy *et al*. who in their study of the lysine-modified KLVFF peptide found that three or more repeat units of lysine were more effective [24]. The ability of the KLVFFP5 peptide to destabilize the fibrils and bind to them was compared to the KLVFF peptide. Our results indicate that the KLVFFP5 could disrupt the amyloid protofibrils to a greater extent than the KLVFF peptide and bind more strongly to them.

METHODS

The structure of the amyloid-β protofibrils was obtained from the Protein Data Bank (pdb id: 2MXU) [51]. This structure comprises of a triple-parallel S-shaped β-sheet. The three extended β-sheet regions are connected by two loop regions. This model was used in many recent studies [52-58]. The protofibril structure is shown in Figure 1. The S-shaped structure of the protofibrils used in the present study has greater conformational and mechanical stability [54,59,60] compared to the U-shaped structure of the protofibrils used in previous studies [61,62,63] due to a more robust network of hydrogen bonds [59].

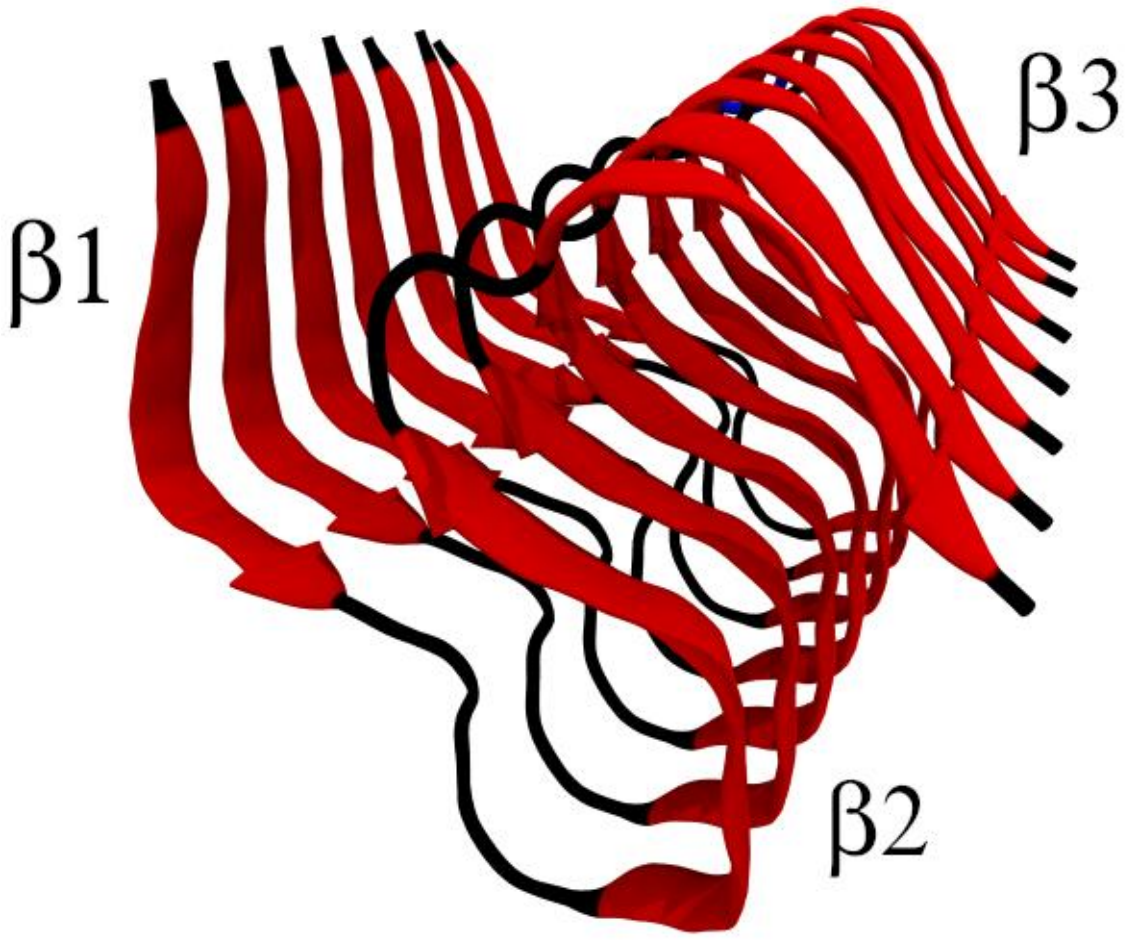


Figure 4.1. The protofibril model consists of three regions of extended beta-sheet structure shown in red color denoted by β1, β2, and β3. These regions are connected by two loop regions shown in black color.

In order to decrease the cost of computation, seven chains of the protofibril structure were considered. Xi *et al*. found that for this model, a protofibril with six chains was the

minimum critical size for protofibril stability [64]. They also found that the absence of the first ten residues did not have an impact on the overall protofibril structure. Hence, the choice of the present protofibril structure with seven chains of $A\beta_{11-42}$ is valid for the study.

Neurotoxic oligomers occur in many different sizes. At the begining of the oligomerization of amyloid-β fibrils, the paranuclei consist of pentamers and hexamers which eventually form larger oligomers which later form fibrils [65]. A coarse-grained simulation study by Cheon *et al*. showed that pentamers and hexamers are on-pathway intermediates in fibril formation [66]. Another study by Kahler showed that the hexamer is a favourable size for the oligomerization of longer protofibrils [67]. Thus, the heptamer model of the protofibrils considered in the present study is a valid target for the design of therapeutic molecules for Alzheimer's disease. Figure 1 shows the model of the protofibrils. For convenience, we refer to the regions shown in the figure as β-1, β-2, and β-3; and loop-1 and loop-2. The β-1 region is spanned by the residues Val 12 – Phe 20, the b-2 region by Asn 27 – Ile 32, and the β-3 region by Val 36 – Ile 41. The chains of the protofibrils are labeled chains A-G; and the ligand as chain H, for convenience.

The systems considered in this study are summarized in Table 1. Three sets of simulations were performed for the following systems: $A\beta_{11-42}$ protofibrils in water, KLVFF-$A\beta_{11-42}$, and $KLVFFP_5$ - $A\beta_{11-42}$. The KLVFF and $KLVFFP_5$ ligands were constructed by PyMol 2.3.1. These ligands were docked to the amyloid protofibrils using the automated protein docking server ClusPro [68-72]. The recommended default procedure was followed in ClusPro. Based on favorable desolvation energies and electrostatics, the top 2000 models were clustered and ranked by the ClusPro algorithm [71]. The top ranked model was chosen for the present study. The two models chosen for the KLVFF and $KLVFFP_5$ ligands were identically docked to the protofibrils, enabling a basis for the comparison of their effects on the protofibril structure.

Table 4.1. Summary of systems studied.

System	Composition	Abbreviation
Control System, Replica 1	$A\beta_{11-42}$	C – 1
Control System, Replica 2	$A\beta_{11-42}$	C – 2
Control System, Replica 3	$A\beta_{11-42}$	C – 3
KLVFF – protofibrils, replica 1	KLVFF - $A\beta_{11-42}$	K – 1
KLVFF – protofibrils, replica 2	KLVFF - $A\beta_{11-42}$	K – 2
KLVFF – protofibrils, replica 3	KLVFF - $A\beta_{11-42}$	K – 3
$KLVFFP_5$ – protofibrils, replica 1	$KLVFFP_5$-$A\beta_{11-42}$	KP – 1
$KLVFFP_5$ – protofibrils, replica 2	$KLVFFP_5$-$A\beta_{11-42}$	KP – 2
$KLVFFP_5$ – protofibrils, replica 3	$KLVFFP_5$-$A\beta_{11-42}$	KP – 3

The systems were placed in a cubic box of sizes 7.42543nm (control systems), 7.36071nm (KLVFF systems) and 7.36284nm (KLVFFFP5 systems). The water model used was the TIP3P model. The AMBER99SB-ILDN force field was used for the simulations. In order to neutralize the systems, 7 sodium ions were added to the control systems and 6 to the KLVFF and KLVFFFP5 systems. The steepest descents algorithm was used for energy minimization, followed by equilibriation in two steps with position restraints on heavy atoms. First, an NVT equilibritaion was done for 500ns at 300K, and then an NPT equilibriation was done for 500ns using the Nosé–Hoover temperature bath [73,74] and the Parrinello-Rahman barostat [75,76]. Finally, a molecular dynamics production run was performed for 1000ns (1μs). The GROMACS 5.1.4 package was used for the simulations with periodic boundary conditions applied in all directions [77]. The P-LINCS constraint algorithm was used for bond lengths [78]. The neighbor search cut-off was at approximately 1 nm. The fast-smooth particle mesh Ewald summation method was used to calculate the long-range electrostatic interactions and the Fourier grid spacing was 0.16 nm [79].

For analysis, we considered the last 300ns of the simulations. In order to check for artefacts, the minimum distance of a protein system to its periodic image was evaluated. The root mean squared deviation (RMSD), the secondary structure, hydrogen bonds, and the solvent

accessible surface area (SASA) were evaluated using the GROMACS tools gmx rms, do_dssp [80, 81], gmx hbond [82], and gmx sasa [83] respectively. In order to evaluate the sustained contacts, the VMD script contactFreq.tcl was used.

The MM/PBSA method was used to evaluate the binding free energies of the amyloid protofibrils with the ligands. According to the MM/PBSA method, the binding energy is given by the following equation [84, 85] :

$$\Delta G^{bind} = \Delta EMM + \Delta G^{psolv} + \Delta G^{npsolv} - T\Delta S \qquad (1)$$

In Equation (1), ΔEMM is the molecular mechanics contribution to the binding free energy in vacuum. The polar solvation energy was evaluated by solving the non-linear Poisson-Boltzmann equation and is expressed by the term ΔG^{psolv}. The non-polar contribution to the solvation energy was evaluated by using the solvent-accessible surface area (SASA) model and is given by the term ΔG^{npsolv}. Entropic terms were not considered in the evaluation of the binding energies of the ligands. The tool g_mmpbsa was used for the binding energy calculations employing the single trajectory protocol [86, 87]. Snapshots were extracted every 100ps from the last 30ns of the simulations for analysis. Using 1000 bootstraps in the g_mmpbsa program, individual contributions to the binding energies by the residues were also evaluated. The solute dielectric constant was 4 and the solvent dielectric constant was 80.

RESULTS AND DISCUSSION

We describe the binding modes of the KLVFF and $KLVFFP_5$ ligands to the amyloid-β protofibrils in terms of sustained contacts. Sustained contacts are defined as those contacts in which the protofibril residues which were within 4Å of the ligand residues for more than 40% of the last 300ns of simulations. The binding energies of the systems were evaluated along with the contribution of the most important protofibril residues.

Binding Modes

The highest ranked docked structures of the KLVFF and $KLVFFP_5$ systems were chosen. In these structures, the ligands docked to a binding pocket formed by the S-shape of the protofibrils. The binding pocket was evaluated by the web-servers CASTp 3.0, Bitenet and Active Site Prediction [88-90] which use different methods in their calculations. A study by Grasso *et al.* examined the docking of various natural compounds to this particular structure of the protofibrils which was used in the present study [91]. In their study, ligands such as curcumin, gossypin and piceatannol were able bind to this particular binding pocket [91]. Based

on these observations, the highest-ranked docked structures produced by the ClusPro algorithm was an appropriate choice for the starting structures of the complexes. The docking of both the ligands was identical, enabling a comparison of their effects on the protofibril structure. The starting structures of the systems and the ligand residues which were within 5Å of the protofibril residues are shown in Figure 2. KLVFF and $KLVFFP_5$ made contacts with the protofibril residues Val 12, His 13, His 14, Leu 17, Ile 32, Gly 33, and Leu 34. These residues are in the β-1 region and the loop region Gly 33 – Met 35. These residues correspond to the two binding regions observed by Grasso *et al*., who found that most ligands tend to bind to the regions E11–F19 and I32–L34 [91]. The KLVFF ligand made contacts with the residues in chains C-G of the protofibrils, while the increased length of the $KLVFFP_5$ ligand enabled it to make contacts with these residues in all the protofibril chains. These residues are highlighted in Figure 2. The contacts were hydrophobic in nature.

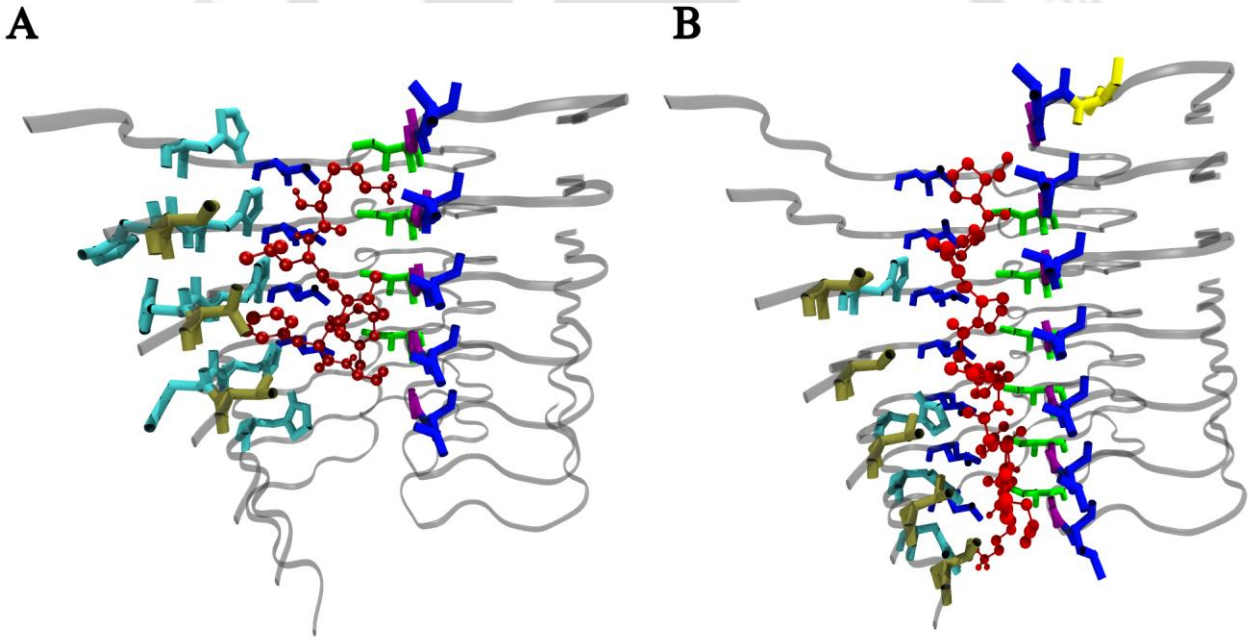


Figure 4.2. Initial docked structure of the **A.** K systems and **B.** KP systems. The Aβ protofibrils are shown as ribbons. The ligands are shown in the CPK representation in red. The protofibril residues which were within 5 of the ligands are shown as bonds. Val 12 is shown in tan color, His 13 and His 14 in cyan, Leu 17 and Leu 34 in blue, Ile 32 in green, Gly 33 in purple and Met 35 in yellow color.

The KLVFF peptide made hydrophobic, hydrogen bonding, and aromatic interactions with the protofibrils. In the last 300ns of the simulations, the KLVFF ligand made eight sustained contacts with the protofibril residues in the K – 1 system. In the K – 1 system, the

average number of hydrogen bonds between the KLVFF ligand and the protofibril residues was 3.24. The strong hydrogen bonds between His 14C – Phe 4H lasted for 88% of the time. There were aromatic interactions between Phe 20G – Phe 4H. The prominent contributions to the binding were from the hydrophobic contacts. These residues are shown in Figure 3 along with their contribution to the binding energy based on the per-residue decomposition of the binding energy term in the last 30ns. The prominent contributions to the binding energy were from Val 12C and His 14C. The contributions to the binding energy are shown in Table 2. The binding energy term was dominated by the van der Waals interactions. The electrostatic interactions were favourable for binding.

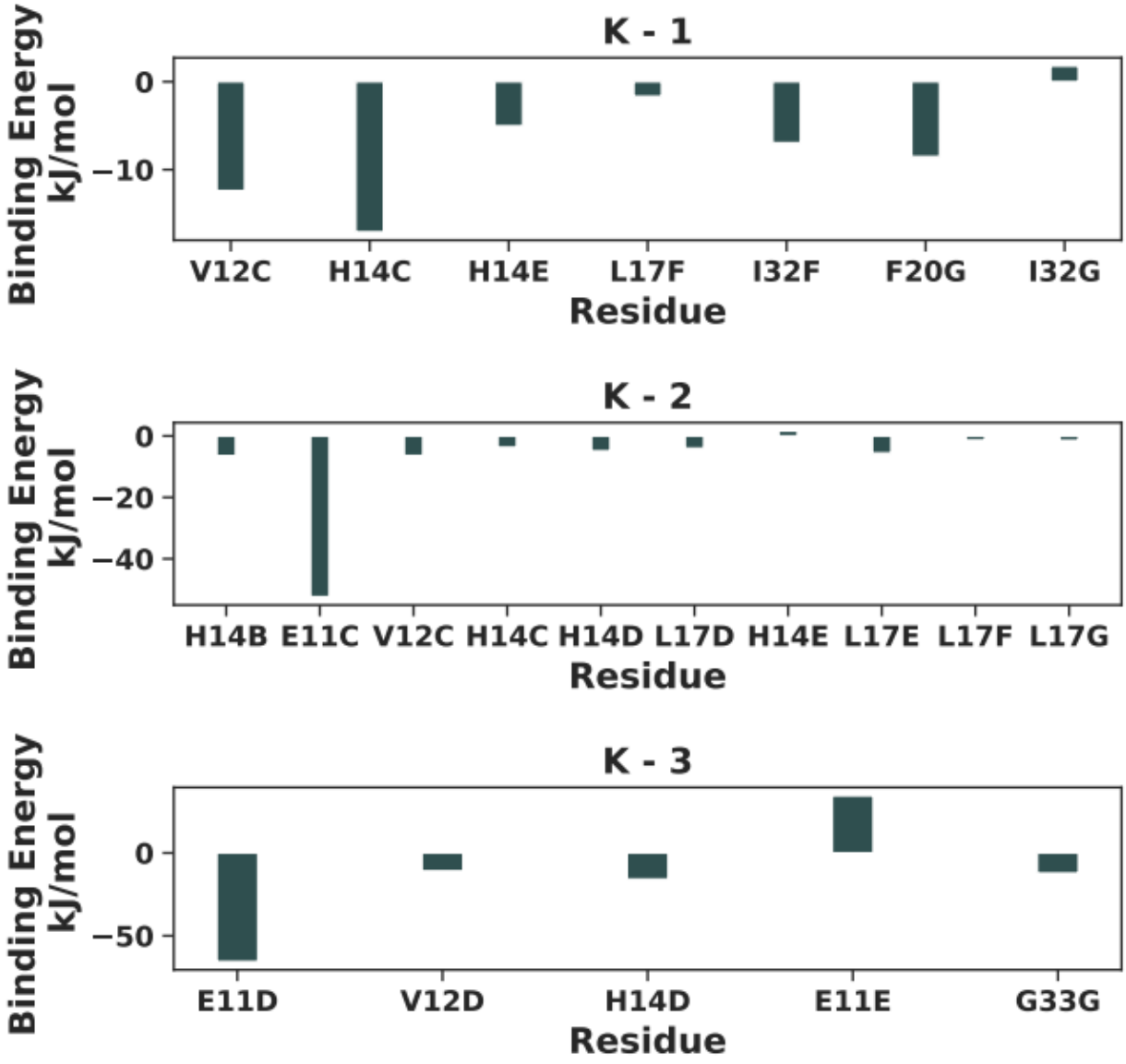


Figure 4.3. The residues which made sustained contacts with the protofibril residues in the last 300ns were the most important for binding. The energetic contribution of these residues to the binding energy in the last 30ns of the K – 1 and K – 2 systems is shown here.

Table 4.2. MMPBSA results expressed in kJ mol^{-1}.

System	van der Waals energy	Electrostatic energy	Polar solvation energy	SASA energy	Binding Energy
K – 1	-186.825 ± 0.958	-943.763 ± 7.976	857.058 ± 6.456	-24.328 ± 0.099	-297.919 ± 3.616
K - 2	-175.002 ± 1.003	-769.843 ± 5.549	762.754 ± 5.191	-23.145 ± 0.101	-205.313 ± 3.285
K – 3	-213.709 ± 1.132	-1501.012 ± 12.067	1347.511 ± 9.156	-29.088 ± 0.095	-396.634 ± 3.707
KP – 1	-213.656 ±1.160	-1239.946 ± 11.436	1247.314 ± 8.200	-28.845 ± 0.128	-234.845 ± 4.473
KP – 2	-374.728 ± 1.210	-1160.455 ± 8.772	1187.126 ± 5.345	-39.990 ± 0.095	-388.177 ± 4.545
KP – 3	-419.601 ± 1.079	-439.377 ± 4.046	820.437 ± 2.878	-43.269 ± 0.076	-81.957 ± 2.465

In the K – 2 system the KLVFF peptide made 12 sustained contacts with the protofibril residues. The average number of hydrogen bonds between the ligand and the protofibrils in the last 300ns in the K – 2 system was 1.78. The hydrogen bonds between Glu 11C – Val 3H lasted for 82.80% of the time. In the last 150ns, aromatic interactions between Phe 19F – Phe 5H developed. The important contributions to the binding were from the hydrophobic contacts. These are shown in Figure 3. The Glu 11C residue made the highest contribution to the binding energy. The contributions to the binding energy are shown in Table 2. The electrostatic interaction and polar solvation energy terms were comparable, implying that the driving force for the binding were the van der Waals interactions. Due to the decrease in favourable electrostatic interactions, this ligand had a lower affinity compared to the K – 1 ligand.

In the K – 3 system the KLVFF ligand made 15 sustained contacts with the fibrils, and made an average of 4.05 hydrogen bonds with the fibrils. Due to a greater number of contacts made with the fibril residues, an increased number of hydrogen bonds, and more favourable electrostatic interactions, the ligand in this replicate had a higher binding affinity for the fibrils than in the K – 1 and K – 2 systems.

In the last 300ns, the $KLVFFP_5$ ligand made 15 sustained contacts with the protofibril residues in the KP – 1 system. The average number of hydrogen bonds between the ligand and the protofibrils was 2.75. The prominent hydrogen bonds were between Glu 22A – Lys 1H (50.10%), Glu 22A – Leu 2H (27.20%), and Asp 23A – Lys 1H (23.60%). As these residues were charged, they contributed significantly to the binding energy, as shown in Figure 4. The hydrophobic contacts made an important contribution to the binding, the Ile 32D - Pro 9H contact in particular. The contributions to the binding energy are shown in Table 2. The

electrostatic interaction and polar solvation energy terms were comparable, implying that the binding was driven by the van der Waals interactions.

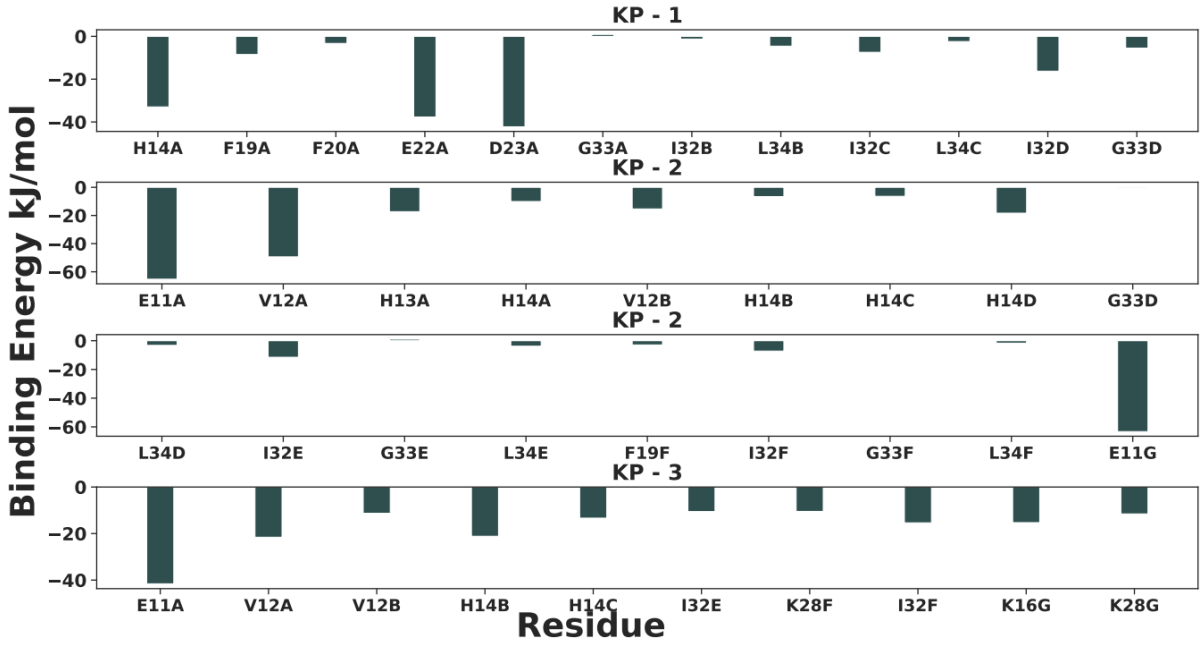


Figure 4.4. The residues which made sustained contacts with the protofibril residues in the last 300ns were the most important for binding. The energetic contribution of these residues to the binding energy in the last 30ns of the K – 1 and K – 2 systems is shown here.

In the KP – 2 system, the KLVFFP$_5$ ligand made 27 sustained contacts with the protofibril residues. This was significantly more than the other systems. The average number of hydrogen bonds between the ligand and the protofibrils was 3.73. The important hydrogen bonds were those between Glu 11A – Lys 1H (54.80%), Val 12A – Lys 1H (84.80%), Val 12A – Leu 2H (93.70%), and Glu 11G – Pro 9H (39.20%). Most of the sustained contacts were hydrophobic in nature. The prominent aromatic interaction was between Phe 19F - Phe 5H. The Glu 11A – Lys 1H contact was electrostatic in nature. Important contributions to the binding energy were made by the protofibril residues Glu 11A, Val 12A, His 13A, His 14A, Val 12B, His 14D, Ile 32E, and Glu 11G, as shown in Figure 4. The contributions to the binding energy are shown in Table 2. This ligand had the highest binding affinity compared to the other systems. The electrostatic interactions were not favourable for binding, and hence the binding was driven by the van der Waals interactions. The high number of hydrophobic contacts made by the ligand contributed significantly to the binding of the KLVFFP$_5$ ligand to the protofibrils. This ligand made more contacts with the protofibril residues than the ligand in the KP – 1 system, and hence had an increased binding affinity. In the case of the KP – 3 system, the ligand residues made 21

sustained contacts with the fibril residues. Due to the relatively low magnitude of the electrostatic interaction energy term and the high magnitude of the polar solvation energy term, the binding affinity was low.

The proline residues in the $KLVFFP_5$ ligands were able to form hydrophobic interactions with the protofibril residues. The proline residues contributed to 5 sustained contacts in the KP – 1 system and 7 sustained contacts in the KP – 2 system. These favourable interactions contributed to the overall binding affinity of the ligand. The high affinity of proline residues to the amyloid protofibrils had been observed in a previous study [42]. In the K – 1 and K – 2 systems the Phe 4H residue made the maximum sustained contacts with the protofibril residues. In comparison, the residues other than the Phe 4H in the KLVFF sequence in the $KLVFFP_5$ ligands in the KP – 1 and KP – 2 systems bound more tightly to the protofibrils by making more contacts. Thus, the proline modification to the KLVFF peptide enhanced its binding to the protofibrils.

Destabilization of the protofibrils

In the presence of the KLVFF and $KLVFFP_5$ ligands the protofibrils became destabilized. The final snapshots of the simulations are shown in Figure 5. In the absence of the ligands, the protofibrils were stable and maintained their β-sheet structure, hydrogen bonding network and their salt bridges. The extent of destabilization of the protofibrils was the maximum in the presence of the $KLVFFP_5$ ligand in the KP – 1 and KP – 2 systems. The KP – 3 system was an outlier due to the poor binding of the ligand to the fibrils.

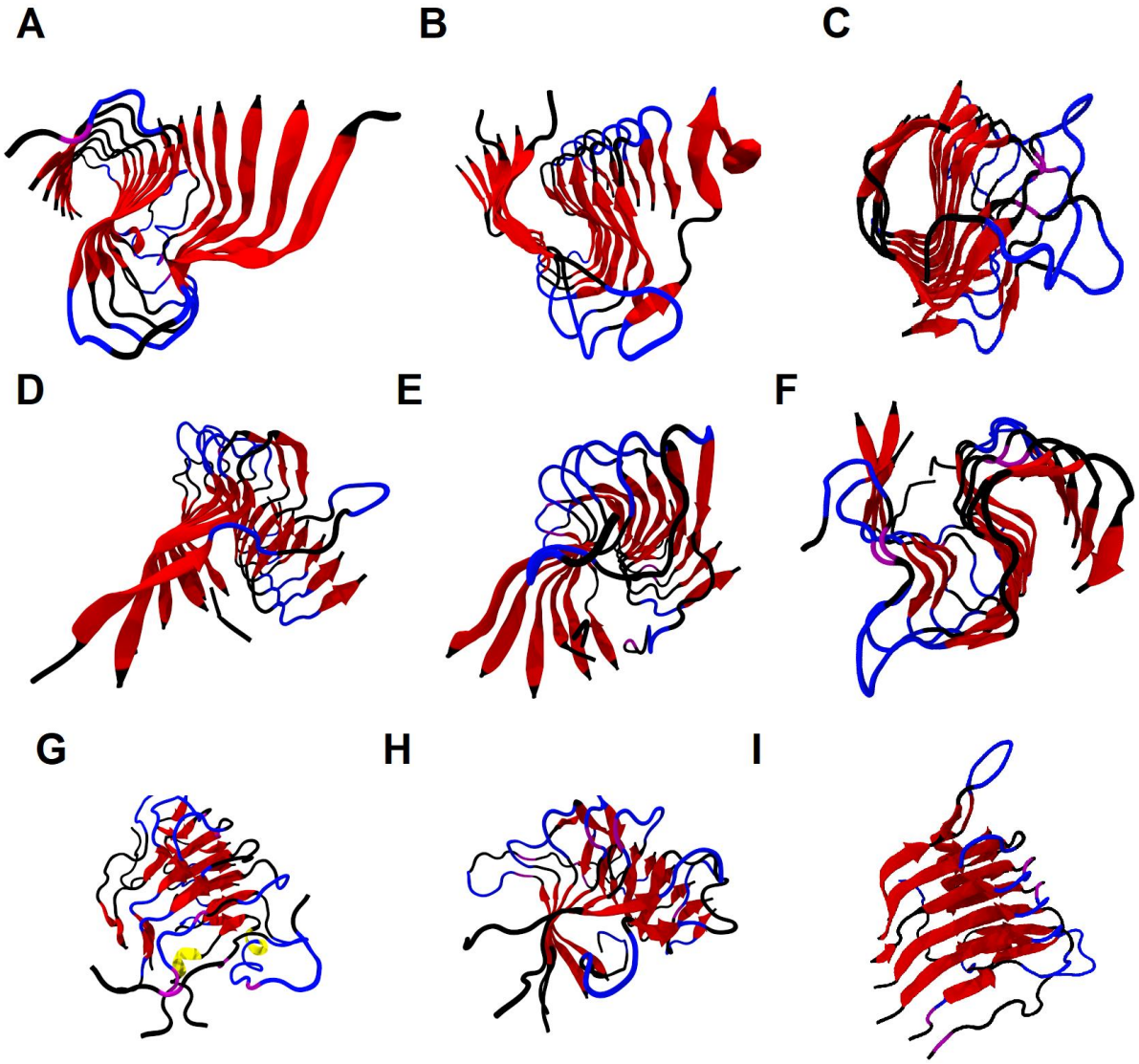


Figure 4.5. Snapshots of the systems at the end of the simulations, i.e., at 1μs. **A.** C – 1 system. **B.** C – 2 system. **C.** C – 3 system. **D.** K – 1 system. **E.** K – 2 system. **F.** K – 3 system. **G.** KP – 1 system. **H.** KP – 2 system. **I.** KP – 3 system. In the presence of the $KLVFFP_5$ peptide, the extent of destabilization of the protofibrils was the maximum (the KP – 3 system being an outlier). Extended β-sheet regions are shown in red color, turn regions in blue, random coils in black, helices in yellow and bridge-β in purple color.

The all-atom root mean squared deviation (RMSD) of the Cα atoms of the Aβ protofibrils was evaluated in order to monitor the time evolution of their structural stability. The Aβ protofibril control systems were stable and had average RMSD values of 0.49 ± 0.01nm, 0.48 ± 0.01nm, and 0.58 ± 0.02nm respectively in the last 300ns. These are shown in Figure 6.

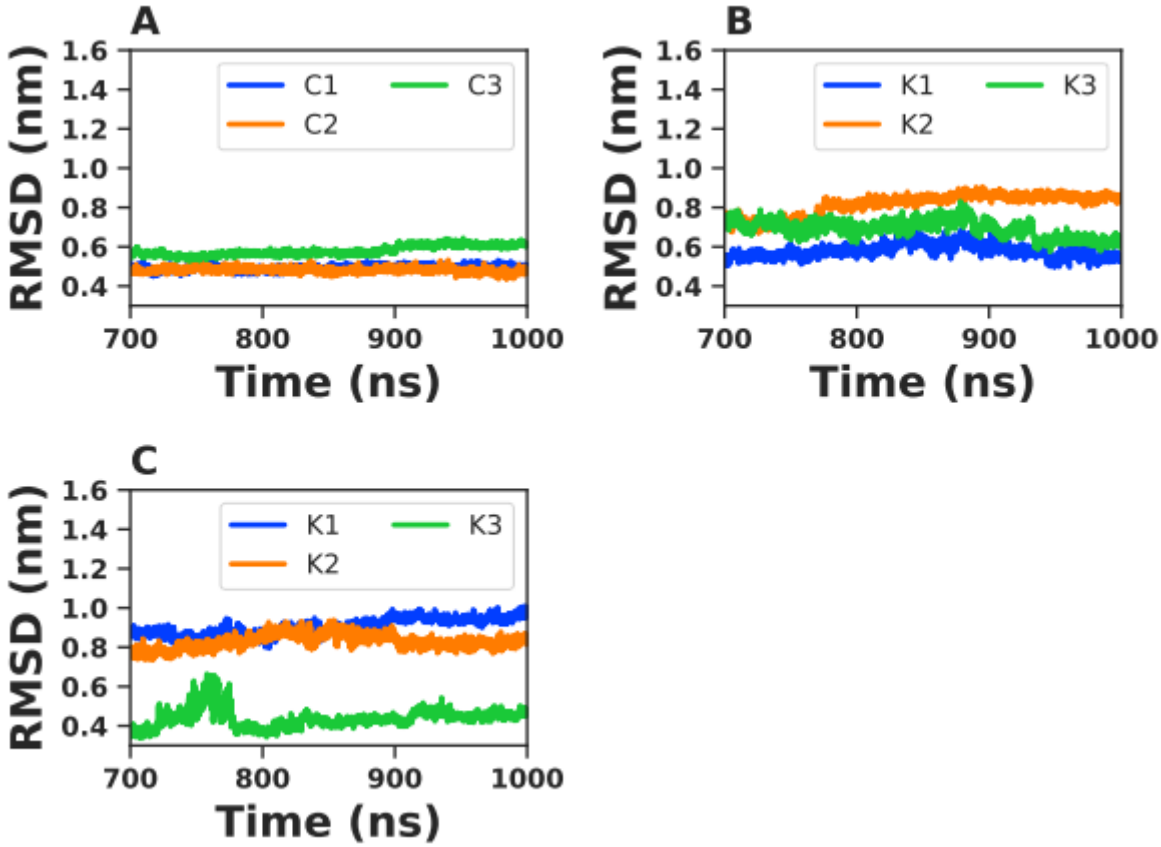


Figure 4.6. Root-mean squared deviation (RMSD) of the Cα atoms of the Aβ protofibrils in all the systems considered. The values of the KP – 1 and KP – 2 systems were higher than the K systems, indicating a greater extent of destabilization of the protofibrils.

The destabilization of the protofibrils in the presence of inhibitors was indicated by an increase in the RMSD value of the protofibrils. In the presence of the KLVFF and $KLVFFP_5$ ligands, the RMSD values of the protofibrils had increased. The average RMSD values of the Aβ protofibrils in the KLVFF-Aβ systems had increased to 0.58 ± 0.04nm, 0.82 ± 0.06nm, and 0.69 ± 0.04 respectively, indicating the destabilizing effect of the KLVFF peptide. In the KP – 1 and KP – 2 systems, the average RMSD values of the Aβ protofibrils further increased to 0.90 ± 0.04nm and 0.83 ± 0.04nm. The increased values of the RMSD indicate that the $KLVFFP_5$ peptide could destabilize the Aβ protofibrils to a greater extent than the KLVFF peptide in the KP – 1 and KP – 2 systems. The RMSD values as a function of time are shown in Figure 6.

The average RMSD values of the three extended β-strand regions β - 1, β - 2 and β - 3 were evaluated in order to determine which region was destabilized by the ligands to the highest extent. The ligands in all the systems made more contacts with protofibril residues in the β - 1 region, and subsequently, this region was destabilized to a greater extent. The KLVFF – Aβ systems had higher average RMSD values in the β-1 region than the controls, and these values

were even higher for the KP – 1 and KP – 2 systems. The RMSD values of the residues in the β-2 regions of the systems were comparable. The increased RMSD values show that both KLVFF and $KLVFFP_5$ peptides could destabilize the Aβ protofibrils, with the latter destabilizing the protofibrils to a greater extent in the KP – 1 and KP – 2 systems. The $KLVFFP_5$ peptide could destabilize the protofibrils in the β - 1 region (in the KP – 1 and KP – 2 systems), the β - 2 region (in the KP – 2 system), and the β - 3 region (in the KP – 1 system) to a greater extent than the KLVFF peptide. These values are shown in Figure 7.

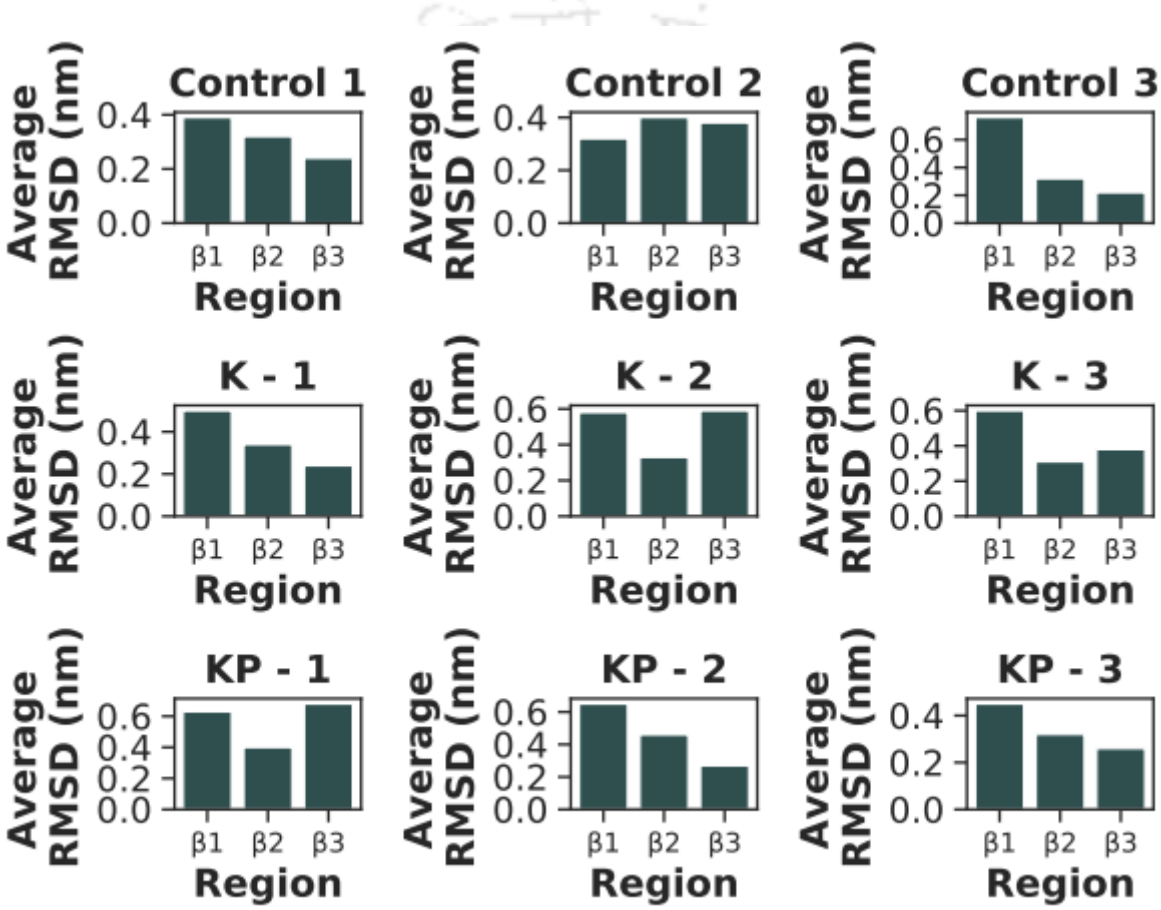


Figure 4.7. The average RMSD values in the last 300ns of the different extended β-sheet regions. The β-1 region was destabilized the most by the $KLVFFP_5$ peptide.

The destabilization of the protofibrils was further evaluated by the effect of the ligands on the secondary structure of the Aβ protofibrils. There was a reduction in the β-sheet content of the protofibrils in the presence of both ligands, indicating their ability to destabilize the protofibrils. In the control Aβ systems, the percentage of residues which formed a β-sheet was 45.91%, 52.55% and 41.87% respectively. In the KLVFF - Aβ systems, these values were 42.20%, 42.38%, and 36.66% respectively, indicating a reduction in beta-sheet content compared to the C – 1 and C – 2 systems. In the KP – 1 and KP – 2 systems there was a further

reduction of the β-sheet content to 28.02% and 35.27% respectively. The drastic reduction in the β-sheet content of the protofibrils in the presence of the KLVFFP$_5$ peptide indicates the improved destabilization ability of the proline enhancement of the KLVFF peptide. The number of protofibril residues which formed random coils increased in the presence of these peptides. The percentage of residues which formed coils in the control Aβ systems was 34.14%, 28.69%, and 36.42% respectively. These values increased to 39.21% and 37.04%, and 38.58% in the KLVFF-Aβ systems and 39.46% and 39.52% in the KP – 1 and KP – 2 systems. These values are shown in Figure 8.

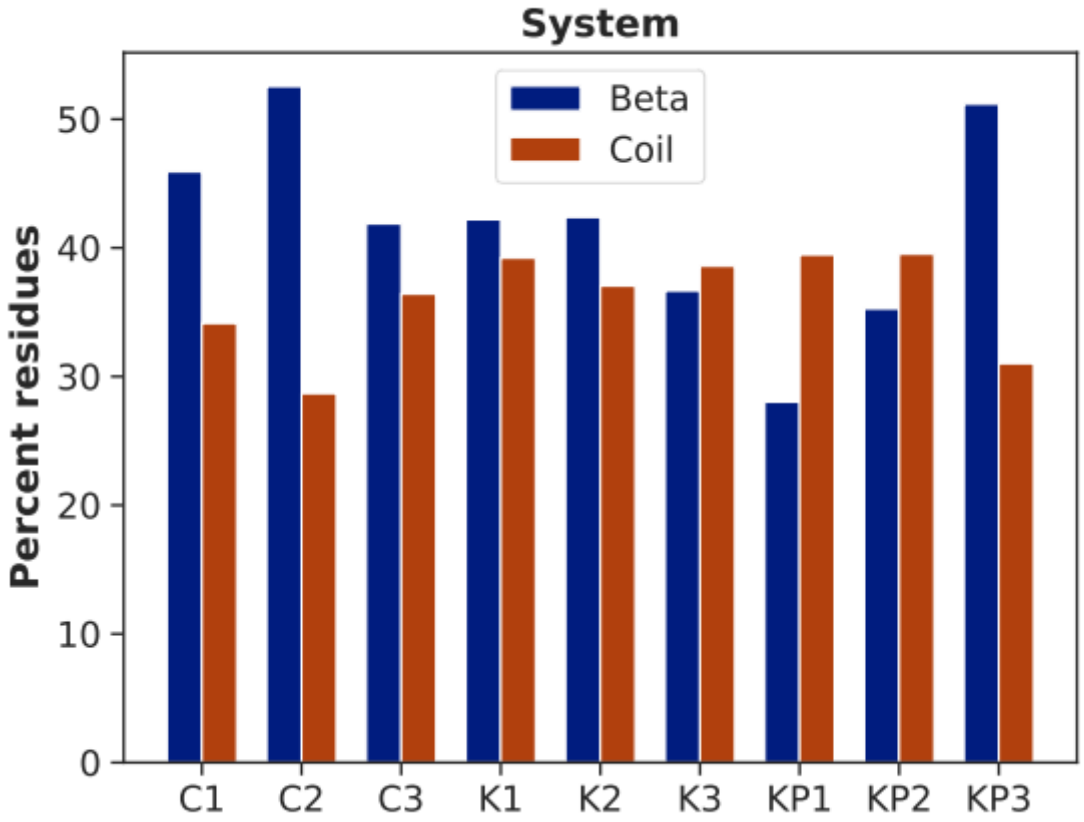


Figure 4.8. The percent of protofibril residues which formed β-sheets and coils in the systems considered. In the presence of the KLVFF peptide, the β-sheet content of the protofibrils reduced, and in the presence of the KLVFFP$_5$ peptide, it reduced even further indicating the increased destabilization of the protofibrils. With a reduction in the beta-sheet content, there was an associated increase of the number of residues which formed coils.

The presence of the KLVFFP$_5$ peptide induced the formation of 3_{10} helices and α-helices in the Aβ protofibrils. In the KP – 2 system, the helices were observed in the residues 25-29 of chain A; and the residues in the two regions 20-27 and 34-37 in chain G. In the KP – 3 system,

the helices were observed in the residues 21-23 of chain A. A plot of the helix-forming residues as a function of the percent duration of the total simulation time is shown in Figure 9. 3_{10} helices were more prominent in chains A of the systems KP – 1 and KP – 2; while in chain G of the KP – 1 system, α-helices were more prominent.

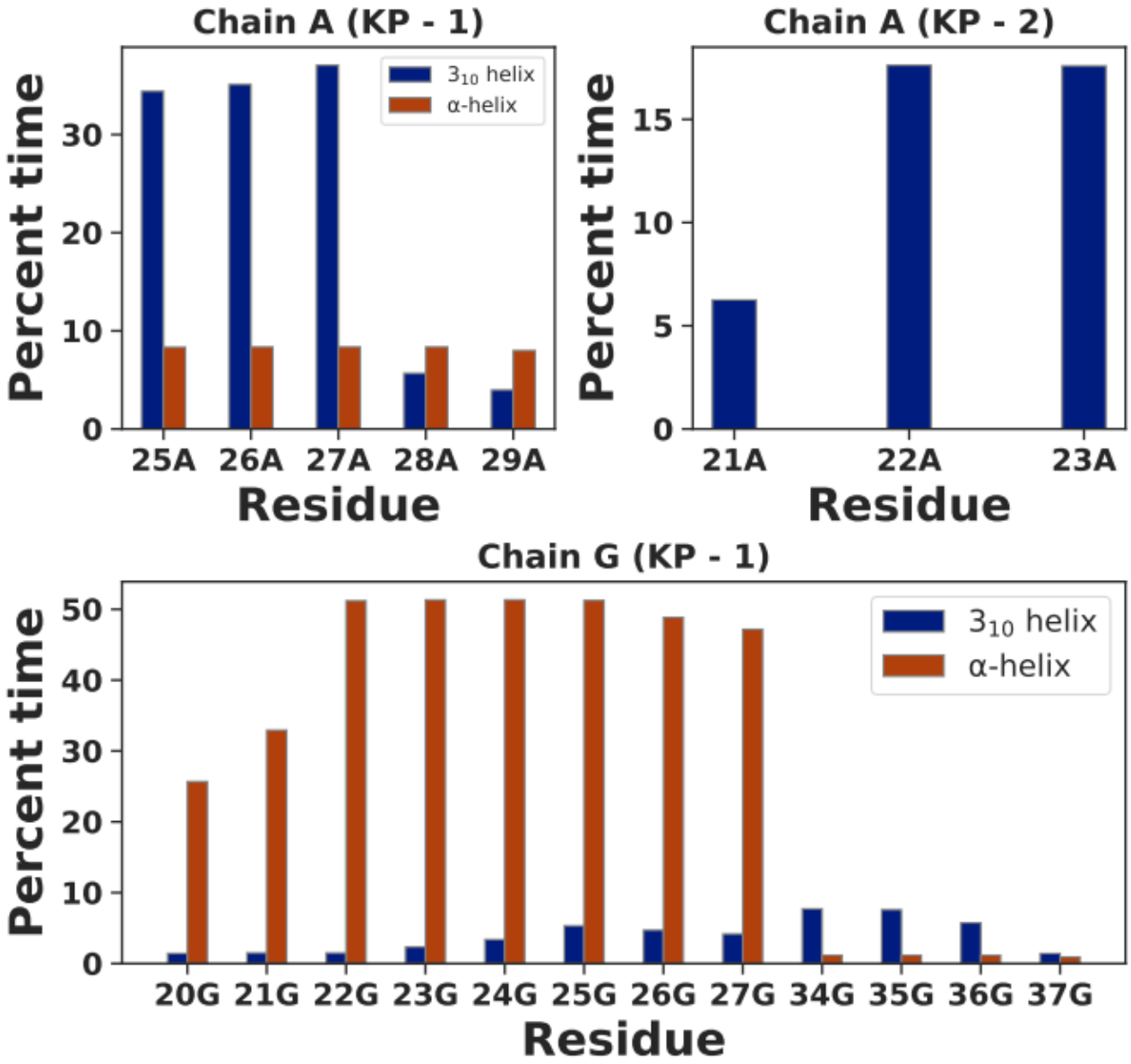


Figure 4.9. The residues which participated in the formation of helices and their associated durations. In chain A of the KP systems, 3_{10} helices were observed. In chain G of the KP – 1 system, α-helices existed for a significant period of time.

A reduction in the number of hydrogen bonds in the presence of the ligands was an indication of the destabilization of the amyloid protofibrils. The $KLVFFP_5$ peptides in the KP – 1 and KP – 2 systems were able to modify the hydrogen bonding network to a greater extent than the KLVFF peptide, as shown in Figure 10, which shows the average number of hydrogen

bonds in the amyloid-β protofibrils in the last 300ns. According to the figure, the average number of hydrogen bonds was the maximum in the systems with the fibrils alone, and the least in the KP – 1 and KP – 2 systems. The hydrogen bonds between neighbouring chains in the systems were further evaluated. The interchain hydrogen bonds which were disrupted by the KLVFF and $KLVFFP_5$ peptides were those between chains D-E, E-F and F-G. Figure 11 shows the average number of inter-chain hydrogen bonds in all the systems considered. The average number of interchain hydrogen bonds between the chains E-F and F-G in the K – 1 and K – 2 systems was less than the controls. In the KP – 1 and KP – 2 systems, the number of hydrogen bonds reduced even further. The average number of interchain hydrogen bonds between chains A-B, B-C, and C-D were comparable in all the systems considered. Since hydrogen bonds play an important role in the stabilization of the amyloid fibrils, a disruption of these indicates the destabilization of the fibrils in the presence of the ligands. These results indicate that the $KLVFFP_5$ peptide could disrupt the hydrogen bonding network of the amyloid protofibrils to a greater extent than the KLVFF peptide.

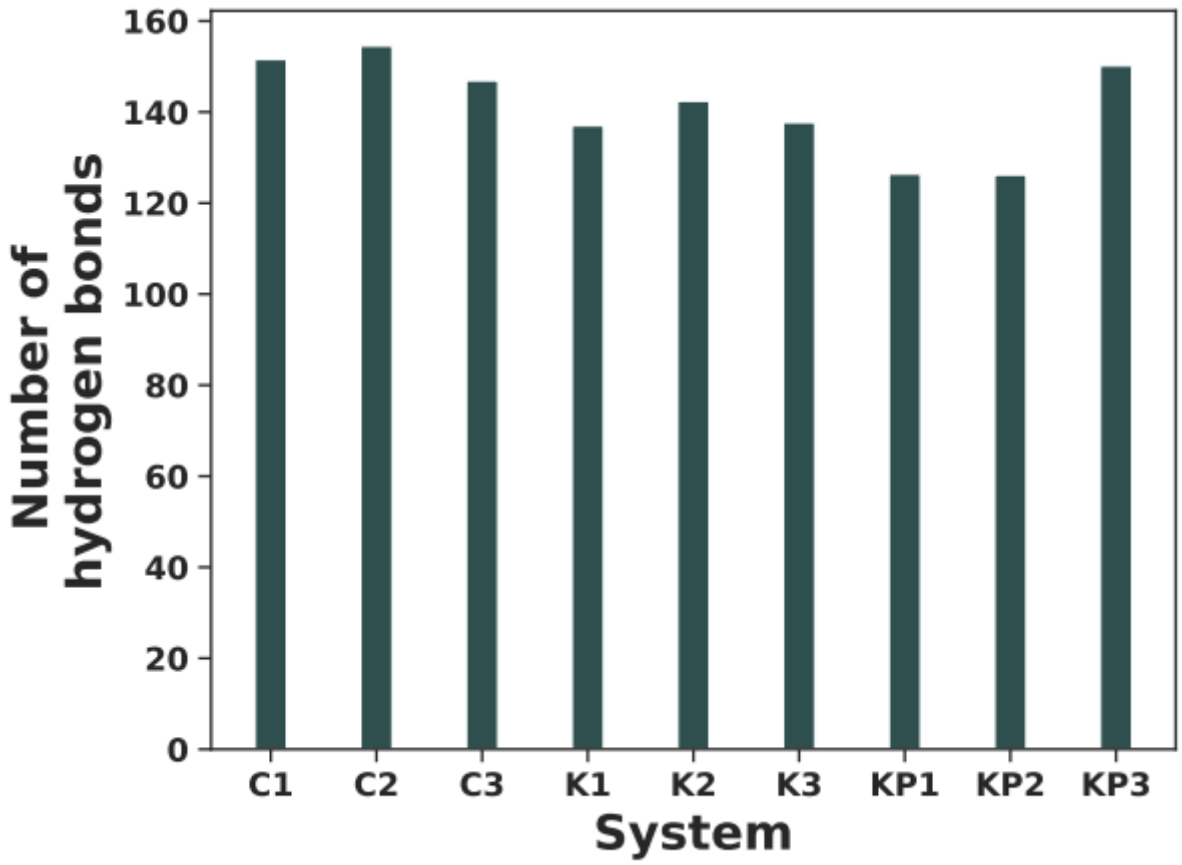


Figure 4.10. The average number of hydrogen bonds in the protofibrils in all the systems. In the KP – 1 and KP – 2 systems the reduction of hydrogen bonds in the protofibrils was more drastic than in the KLVFF systems.

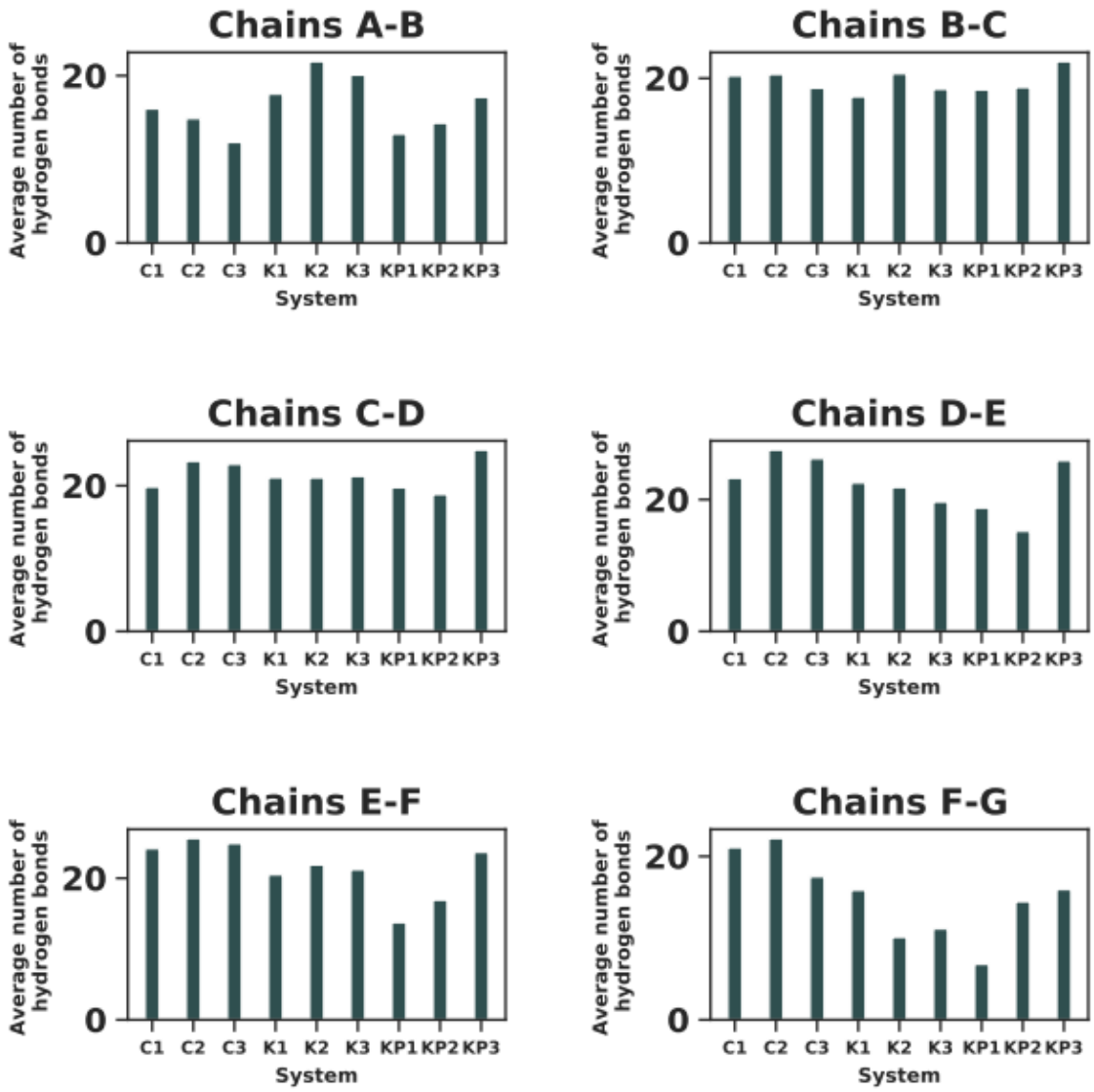


Figure 4.11. The average number of interchain hydrogen bonds. The interchain hydrogen bonds D-E, E-F, and FG were disrupted in the protofibrils. The $KLVFFP_5$ peptide disrupted the hydrogen bonds between these chains to a greater extent than the KLVFF peptide in the KP – 1 and KP – 2 systems.

Salt bridges are important for the stability of the amyloid fibrils [92]. In the solid-state NMR studies from which the structure of the amyloid protofibrils used in the present study was obtained, it was observed that the stabilization of the salt bridge between Lys 28 and Ala 42 was responsible for the S-shaped triple β-sheet motif [51]. This salt bridge is thought to function as a self-recognition molecular switch in the oligomerization of the fibrils [51]. The $KLVFFP_5$ peptide was able to weaken these salt-bridges to a greater extent than the KLVFF peptide in the KP – 1 and KP – 2 systems resulting in the destabilization of the protofibril structure. The

average distance between the salt-bridge forming residues and the fraction of the time in which they were within 4Å of each other was measured to characterize the salt-bridges. The salt bridges in the K – 1 system were well preserved. In the K – 2 system, the salt bridges between the chains C – D, D – E, and F – G were disrupted. In the KP – 1 system, the salt bridges which were disrupted were those between chains A-B, and D-E. The salt bridges in the KP – 2 system were disrupted to the greatest extent compared to the other systems. The average distance between the Ala 42- Lys 28 residues was greater than 4Å between all the chains. The 42E – 28F salt bridge was completely lost. The 42B – 28C and 42F – 28G salt bridge-forming residues were in contact with each other for a negligible fraction of time. Disruption of the other salt bridges was also observed. Figures 12 and 13 show the average distance between the salt bridge-forming residues and the fraction of time in which they were within 4Å of each other.

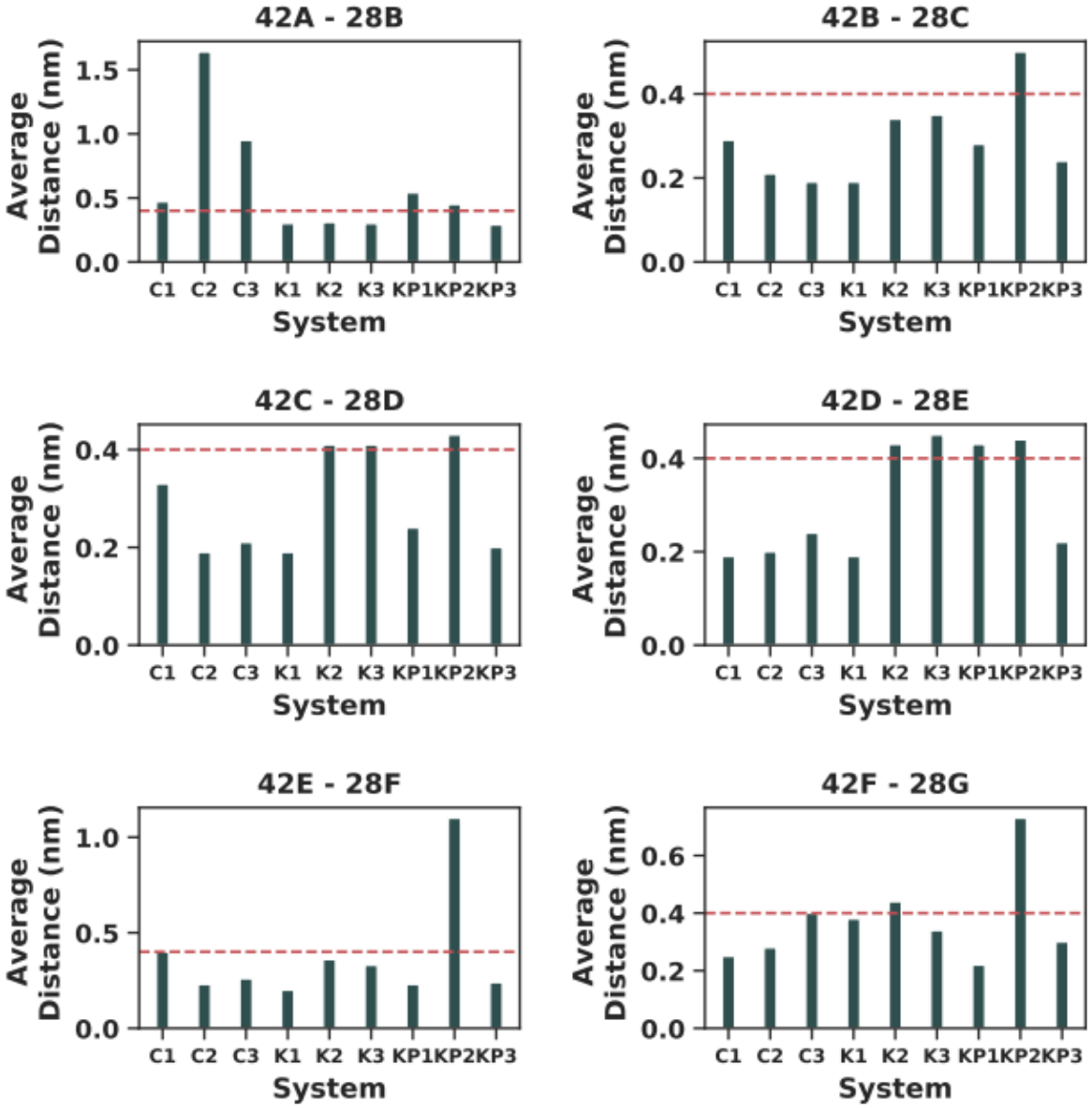


Figure 4.12. The average distance between the salt bridge forming pairs Ala 42 and Lys 28 in all the systems considered. The red dashed line indicates the cut-off distance for the existence of a salt-bridge. The KLVFFP$_5$ peptide could destabilize all the salt bridges in the KP – 1 and KP – 2 systems.

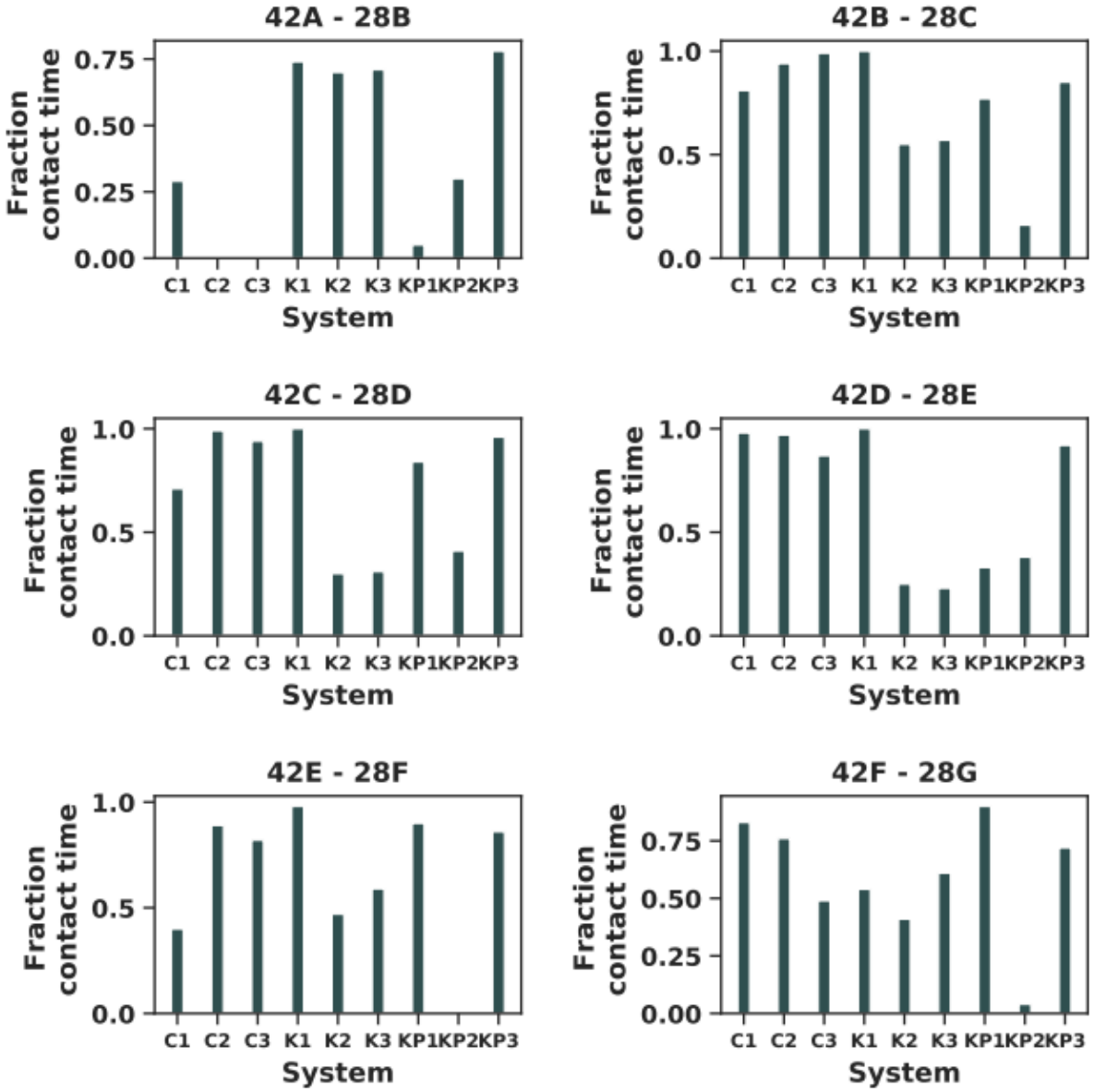


Figure 4.13. The fraction of time in which the salt-bridge forming residue pairs Ala 42 and Lys 28 were within the cut-off distance of 4Å of each other. In the KP – 1 and KP – 2 systems, these pairs were in contact for a significantly reduced fraction of time, indicating the weakening of the salt-bridges.

The tight interatomic packing of the protofibrils was made looser in the presence of the KLVFF and $KLVFFP_5$ peptides. The average values of the radius of gyration and the solvent-accessible surface area increased in the presence of the peptides. This indicates that the protofibrils were less compact and more exposed to the solvent in the presence of the $KLVFFP_5$ peptide. These values were the highest in the presence of the $KLVFFP_5$ peptide as shown in Table 3.

Table 4.3. Radius of gyration and solvent-accessible surface areas (SASA) of the protofibrils.

System	Radius of gyration (nm)	SASA (nm^2)
C – 1	1.66	100.78
C – 2	1.66	103.23
C – 3	1.67	104.32
K – 1	1.70	111.74
K – 2	1.69	109.14
K – 3	1.68	111.71
KP – 1	1.72	115.98
KP – 2	1.73	115.55
KP – 3	1.69	108.68

By disrupting interactions which are critical for the stability of the profibrils, the KLVFF and $KLVFFP_5$ ligands were able to destabilize the protofibrils. For this particular model of the protofibrils, Grasso *et al.* identified three mechanisms in which ligands can destabilize the protofibrils: inter-chain destabilization, pocket distortion and pocket stabilization [91]. In the present study, we found that the mechanism of destabilization was by the pocket distortion mechanism as defined by Grasso *et al.* [91]. The extent of destabilization was measured by the root mean squared deviation (RMSD), the extent of the β – sheet structure, the hydrogen bonding network and stability of the salt bridges. The RMSD values of the $KLVFFP_5$ – Aβ systems were higher than that of the KLVFF – Aβ systems indicating a greater degree of destabilization. There were fewer residues forming extending β-sheets in the presence of the $KLVFFP_5$ peptide than in the presence of the KLVFF peptide. By forming contacts with the $KLVFFP_5$ peptide the protofibril residues lost their native contacts, leading to the loss of their structure. The $KLVFFP_5$ peptide also had a greater ability to disrupt the hydrogen bonding network and the salt bridges of the protofibrils. The tight interatomic packing of the protofibrils was made looser by the $KLVFFP_5$ peptide. Thus the proline modification of the KLVFF peptide improved its ability to bind to the protofibrils and disrupt their structure.

CONCLUSIONS

In the present study the interaction of a designed peptide $KLVFFP_5$ with the amyloid beta protofibrils was investigated. This peptide was designed to exploit the properties of the self-recognition sequence of the amyloid-b peptide KLVFF and the β-sheet breaker amino acid proline. The binding modes of the $KLVFFP_5$ and the destabilization of the amyloid protofibrils were characterized. The mechanism of binding was by the formation of hydrophobic contacts

and hydrogen bonds with the protofibril residues. The $KLVFFP_5$ peptide could destabilize the amyloid protofibrils to a greater extent than the KLVFF peptide. In the presence of the $KLVFFP_5$ peptide the protofibrils lost their beta-sheet structure leading to the formation of coils and helices. The hydrogen bonding network of the protofibrils and the salt bridges are critical for protofibril stability. The $KLVFFP_5$ peptide disrupted the hydrogen bonding network and the salt bridges in the protofibrils to a greater extent than the KLVFF peptide. The tight interatomic packing of the protofibrils was made looser by the $KLVFFP_5$ peptide. Our results indicate that the $KLVFFP_5$ peptide is an effective β-sheet disruptor which can be considered in the therapy of Alzheimer's disease.

REFERENCES

1. Querfurth HW, LaFerla FM (2010) Alzheimer's disease. N Engl J Med 362:329–344. https://doi.org/10.1056/NEJMra0909142
2. Ross CA, Poirier MA (2004) Protein aggregation and neurodegenerative disease. Nat Med 10 Suppl: S10-7. https://doi.org/10.1038/nm1066
3. Prusiner SB (2001) Shattuck lecture--neurodegenerative diseases and prions. N Engl J Med 344:1516–1526. https://doi.org/10.1056/NEJM200105173442006
4. Selkoe DJ (2001) Alzheimer's disease: genes, proteins, and therapy. Physiol Rev 81:741–766. https://doi.org/10.1152/physrev.2001.81.2.741
5. Cummings J, Lee G, Ritter A, *et al.* (2019) Alzheimer's disease drug development pipeline: 2019. Alzheimer's Dement (New York, N Y) 5:272–293. https://doi.org/10.1016/j.trci.2019.05.008
6. Man BY-W, Chan H-M, Leung C-H, *et al.* (2011) Group 9 metal-based inhibitors of β-amyloid (1–40) fibrillation as potential therapeutic agents for Alzheimer's disease. Chem Sci 2:917–921. https://doi.org/10.1039/C0SC00636J
7. Raman B, Ban T, Yamaguchi K-I, *et al*. (2005) Metal ion-dependent effects of clioquinol on the fibril growth of an amyloid-β peptide. J Biol Chem 280:16157–16162. https://doi.org/10.1074/jbc.M500309200
8. Barnham KJ, Kenche VB, Ciccotosto GD, *et al.* (2008) Platinum-based inhibitors of amyloid-beta as therapeutic agents for Alzheimer's disease. Proc Natl Acad Sci U S A 105:6813–6818. https://doi.org/10.1073/pnas.0800712105
9. Iscen A, Brue CR, Roberts KF, *et al*. (2019) Inhibition of Amyloid-β Aggregation by Cobalt (III) Schiff Base Complexes: A Computational and Experimental Approach. J Am Chem Soc 141:16685–16695. https://doi.org/10.1021/jacs.9b06388
10. Lannfelt L, Blennow K, Zetterberg H, *et al.* (2008) Safety, efficacy, and biomarker findings of PBT2 in targeting Abeta as a modifying therapy for Alzheimer's disease: a phase IIa, double-blind, randomised, placebo-controlled trial. Lancet Neurol 7:779–786. https://doi.org/10.1016/S1474-4422(08)70167-4
11. Doens D, Valdés-Tresanco ME, Vasquez V, *et al.* (2019) Hexahydropyrrolo[2,3-b] indole Compounds as Potential Therapeutics for Alzheimer's Disease. ACS Chem Neurosci 10:4250–4263. https://doi.org/10.1021/acschemneuro.9b00297

12. Shi S, Wang H, Wang J, *et al.* (2020) Semi-synthesis and biological evaluation of flavone hybrids as multifunctional agents for the potential treatment of Alzheimer's disease. Bioorg Chem 100:103917. https://doi.org/10.1016/j.bioorg.2020.103917
13. Ehrnhoefer DE, Bieschke J, Boeddrich A, *et al*. (2008) EGCG redirects amyloidogenic polypeptides into unstructured, off-pathway oligomers. Nat Struct Mol Biol 15:558–566. https://doi.org/10.1038/nsmb.1437
14. Ladiwala ARA, Bhattacharya M, Perchiacca JM, *et al*. (2012) Rational design of potent domain antibody inhibitors of amyloid fibril assembly. Proc Natl Acad Sci U S A 109:19965–19970. https://doi.org/10.1073/pnas.1208797109
15. Salloway S, Sperling R, Fox NC, *et al.* (2014) Two phase 3 trials of bapineuzumab in mild-to-moderate Alzheimer's disease. N Engl J Med 370:322–333. https://doi.org/10.1056/NEJMoa1304839
16. Liu E, Schmidt ME, Margolin R, *et al*. (2015) Amyloid-β 11C-PiB-PET imaging results from 2 randomized bapineuzumab phase 3 AD trials. Neurology 85:692–700. https://doi.org/10.1212/WNL.0000000000001877
17. Doody RS, Thomas RG, Farlow M, *et al*. (2014) Phase 3 trials of solanezumab for mild-to-moderate Alzheimer's disease. N Engl J Med 370:311–321. https://doi.org/10.1056/NEJMoa1312889
18. Panza F, Lozupone M, Logroscino G, Imbimbo BP (2019) A critical appraisal of amyloid-β-targeting therapies for Alzheimer disease. Nat Rev Neurol 15:73–88. https://doi.org/10.1038/s41582-018-0116-6
19. Goyal D, Shuaib S, Mann S, Goyal B (2017) Rationally Designed Peptides and Peptidomimetics as Inhibitors of Amyloid-β (Aβ) Aggregation: Potential Therapeutics of Alzheimer's Disease. ACS Comb Sci 19:55–80. https://doi.org/10.1021/acscombsci.6b00116
20. Craik DJ, Fairlie DP, Liras S, Price D (2013) The future of peptide-based drugs. Chem Biol Drug Des 81:136–147. https://doi.org/10.1111/cbdd.12055
21. Matsunaga Y, Fujii A, Awasthi A, *et al*. (2004) Eight-residue Abeta peptides inhibit the aggregation and enzymatic activity of Aβ42. Regul Pept 120:227–236. https://doi.org/10.1016/j.regpep.2004.03.013
22. Tjernberg LO, Näslund J, Lindqvist F, *et al.* (1996) Arrest of beta-amyloid fibril formation by a pentapeptide ligand. J Biol Chem 271:8545–8548. https://doi.org/10.1074/jbc.271.15.8545
23. Soto C, Sigurdsson EM, Morelli L, *et al*. (1998) Beta-sheet breaker peptides inhibit fibrillogenesis in a rat brain model of amyloidosis: implications for Alzheimer's therapy. Nat Med 4:822–826. https://doi.org/10.1038/nm0798-822
24. Lowe TL, Strzelec A, Kiessling LL, Murphy RM (2001) Structure-function relationships for inhibitors of beta-amyloid toxicity containing the recognition sequence KLVFF. Biochemistry 40:7882–7889. https://doi.org/10.1021/bi002734u
25. Kapurniotu A, Buck A, Weber M, *et al*. (2003) Conformational restriction via cyclization in beta-amyloid peptide Abeta (1-28) leads to an inhibitor of Aβ (1-28) amyloidogenesis and cytotoxicity. Chem Biol 10:149–159. https://doi.org/10.1016/s1074-5521(03)00022-x
26. Austen BM, Paleologou KE, Ali SAE, *et al*. (2008) Designing peptide inhibitors for oligomerization and toxicity of Alzheimer's beta-amyloid peptide. Biochemistry 47:1984–1992. https://doi.org/10.1021/bi701415b
27. Findeis MA, Lee JJ, Kelley M, *et al*. (2001) Characterization of cholyl-leu-val-phe-phe-ala-OH as an inhibitor of amyloid beta-peptide polymerization. Amyloid Int J Exp Clin Investig Off J Int Soc Amyloidosis 8:231–241. https://doi.org/10.3109/13506120108993819

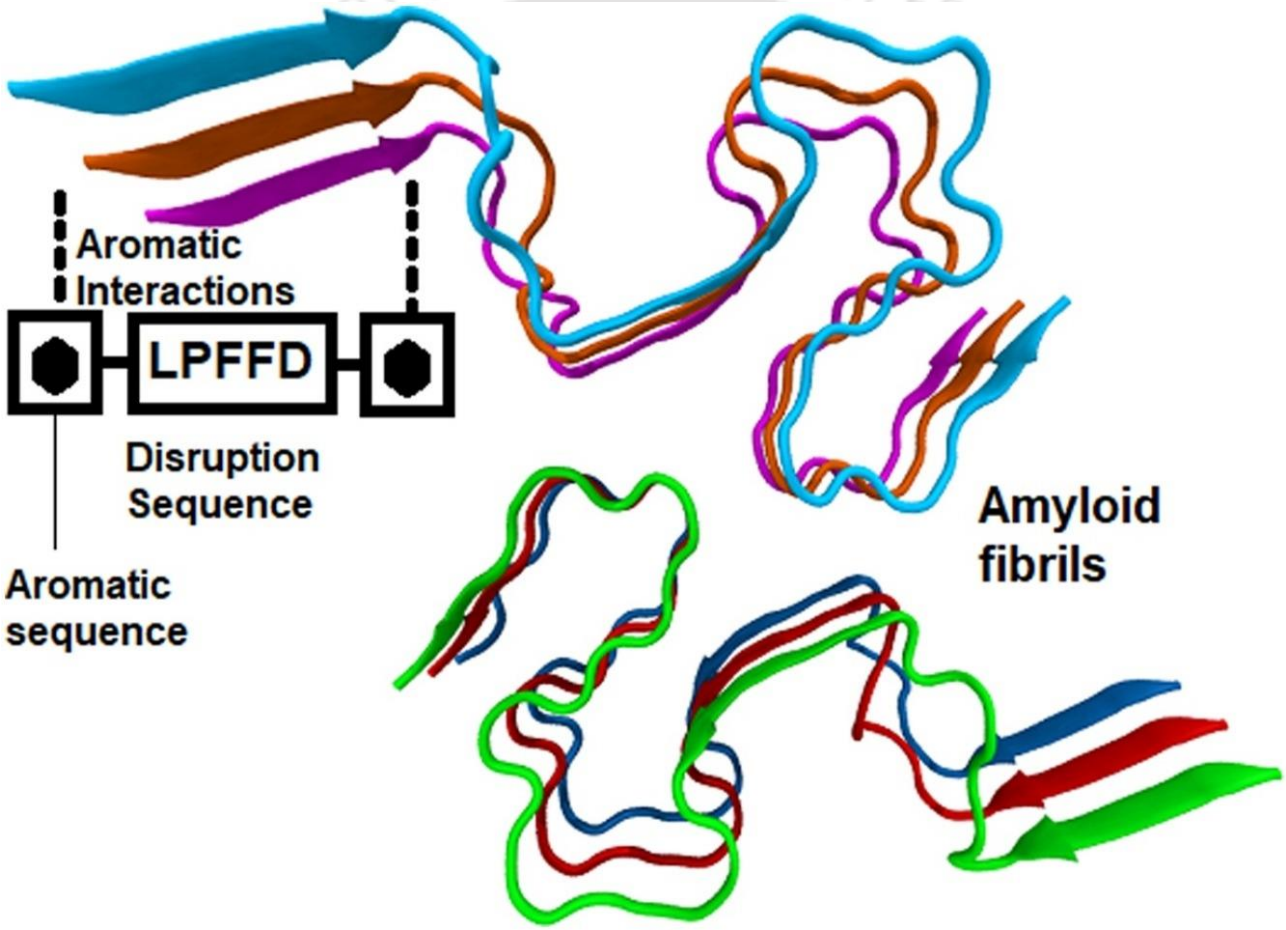
28. Poduslo JF, Curran GL, Kumar A, *et al.* (1999) Beta-sheet breaker peptide inhibitor of Alzheimer's amyloidogenesis with increased blood-brain barrier permeability and resistance to proteolytic degradation in plasma. J Neurobiol 39:371–382
29. Poduslo JF, Curran GL (1996) Polyamine Modification Increases the Permeability of Proteins at the Blood-Nerve and Blood-Brain Barriers. J Neurochem 66:1599–1609. https://doi.org/https://doi.org/10.1046/j.1471-4159.1996.66041599.x
30. Gordon DJ, Sciarretta KL, Meredith SC (2001) Inhibition of beta-amyloid (40) fibrillogenesis and disassembly of beta-amyloid (40) fibrils by short beta-amyloid congeners containing N-methyl amino acids at alternate residues. Biochemistry 40:8237–8245. https://doi.org/10.1021/bi002416v
31. Gordon DJ, Tappe R, Meredith SC (2002) Design and characterization of a membrane permeable N-methyl amino acid-containing peptide that inhibits Abeta1-40 fibrillogenesis. J Pept Res 60:37–55. https://doi.org/10.1034/j.1399-3011.2002.11002.x
32. Cruz M, Tusell JM, Grillo-Bosch D, *et al.* (2004) Inhibition of beta-amyloid toxicity by short peptides containing N-methyl amino acids. J Pept Res 63:324–328. https://doi.org/10.1111/j.1399-3011.2004.00156.x
33. Grillo-Bosch D, Carulla N, Cruz M, *et al*. (2009) Retro-enantio N-methylated peptides as beta-amyloid aggregation inhibitors. ChemMedChem 4:1488–1494. https://doi.org/10.1002/cmdc.200900191
34. Findeis MA (2002) Peptide inhibitors of beta amyloid aggregation. Curr Top Med Chem 2:417–423. https://doi.org/10.2174/1568026024607508
35. Chalifour RJ, McLaughlin RW, Lavoie L, *et al.* (2003) Stereoselective interactions of peptide inhibitors with the beta-amyloid peptide. J Biol Chem 278:34874–34881. https://doi.org/10.1074/jbc.M212694200
36. Jagota S, Rajadas J (2013) Synthesis of d-amino acid peptides and their effect on beta-amyloid aggregation and toxicity in transgenic Caenorhabditis elegans. Med Chem Res 22:3991–4000. https://doi.org/10.1007/s00044-012-0386-2
37. Watanabe K, Nakamura K, Akikusa S, *et al.* (2002) Inhibitors of Fibril Formation and Cytotoxicity of β-Amyloid Peptide Composed of KLVFF Recognition Element and Flexible Hydrophilic Disrupting Element. Biochem Biophys Res Commun 290:121–124. https://doi.org/https://doi.org/10.1006/bbrc.2001.6191
38. Amijee H, Bate C, Williams A, *et al.* (2012) The N-Methylated Peptide SEN304 Powerfully Inhibits Aβ (1–42) Toxicity by Perturbing Oligomer Formation. Biochemistry 51:8338–8352. https://doi.org/10.1021/bi300415v
39. Etienne MA, Aucoin JP, Fu Y, *et al*. (2006) Stoichiometric inhibition of amyloid beta-protein aggregation with peptides containing alternating alpha, alpha-disubstituted amino acids. J Am Chem Soc 128:3522–3523. https://doi.org/10.1021/ja0600678
40. Chafekar SM, Malda H, Merkx M, *et al*. (2007) Branched KLVFF tetramers strongly potentiate inhibition of beta-amyloid aggregation. Chembiochem 8:1857–1864. https://doi.org/10.1002/cbic.200700338
41. Rocha S, Cardoso I, Börner H, *et al.* (2009) Design and biological activity of beta-sheet breaker peptide conjugates. Biochem Biophys Res Commun 380:397–401. https://doi.org/10.1016/j.bbrc.2009.01.090
42. Viet MH, Siposova K, Bednarikova Z, *et al*. (2015) In Silico and in Vitro Study of Binding Affinity of Tripeptides to Amyloid β Fibrils: Implications for Alzheimer's Disease. J Phys Chem B 119:5145–5155. https://doi.org/10.1021/acs.jpcb.5b00006
43. Gladkevich A, Bosker F, Korf J, *et al*. (2007) Proline-rich polypeptides in Alzheimer's disease and neurodegenerative disorders - therapeutic potential or a mirage? Prog

Neuropsychopharmacol Biol Psychiatry 31:1347–1355. https://doi.org/10.1016/j.pnpbp.2007.06.005
44. Bilikiewicz A, Gaus W (2004) Colostrinin (a naturally occurring, proline-rich, polypeptide mixture) in the treatment of Alzheimer's disease. J Alzheimers Dis 6:17–26. https://doi.org/10.3233/jad-2004-6103
45. Leszek J, Inglot AD, Janusz M, *et al*. (1999) Colostrinin: a proline-rich polypeptide (PRP) complex isolated from ovine colostrum for treatment of Alzheimer's disease. A double-blind, placebo-controlled study. Arch Immunol Ther Exp (Warsz) 47:377–385
46. Yenkoyan K, Safaryan K, Chavushyan V, *et al*. (2011) Neuroprotective action of proline-rich polypeptide-1 in β-amyloid induced neurodegeneration in rats. Brain Res Bull 86:262–271. https://doi.org/10.1016/j.brainresbull.2011.08.003
47. Chou PY, Fasman GD (1974) Prediction of protein conformation. Biochemistry 13:222–245. https://doi.org/10.1021/bi00699a002
48. Wood SJ, Wetzel R, Martin JD, Hurle MR (1995) Prolines and amyloidogenicity in fragments of the Alzheimer's peptide beta/A4. Biochemistry 34:724–730. https://doi.org/10.1021/bi00003a003
49. Herning T, Yutani K, Inaka K, *et al*. (1992) Role of proline residues in human lysozyme stability: a scanning calorimetric study combined with X-ray structure analysis of proline mutants. Biochemistry 31:7077–7085. https://doi.org/10.1021/bi00146a008
50. Rauscher S, Baud S, Miao M, *et al*. (2006) Proline and glycine control protein self-organization into elastomeric or amyloid fibrils. Structure 14:1667–1676. https://doi.org/10.1016/j.str.2006.09.008
51. Xiao Y, Ma B, McElheny D, *et al.* (2015) Aβ (1-42) fibril structure illuminates self-recognition and replication of amyloid in Alzheimer's disease. Nat Struct Mol Biol 22:499–505. https://doi.org/10.1038/nsmb.2991
52. Xiang N, Lyu Y, Zhu X, Narsimhan G (2018) Investigation of the interaction of amyloid β peptide (11–42) oligomers with a 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) membrane using molecular dynamics simulation. Phys Chem Chem Phys 20:6817–6829. https://doi.org/10.1039/C7CP07148E
53. Jahanbin F, Bozorgmehr MR, Morsali A, Beyramabadi SA (2019) The effect of different alcohols on the Asp23-Lys28 and Asp23-Ala42 salt bridges of the most effective peptide in Alzheimer's disease: Molecular dynamics viewpoints. J Mol Graph Model 86:199–208. https://doi.org/10.1016/j.jmgm.2018.10.022
54. Grasso G, Rebella M, Muscat S, *et al*. (2018) Conformational Dynamics and Stability of U-Shaped and S-Shaped Amyloid β Assemblies. Int. J. Mol. Sci. 19
55. Jakubowski JM, Orr AA, Le DA, Tamamis P (2020) Interactions between Curcumin Derivatives and Amyloid-β Fibrils: Insights from Molecular Dynamics Simulations. J Chem Inf Model 60:289–305. https://doi.org/10.1021/acs.jcim.9b00561
56. Zhang M, Zheng J, Nussinov R, Ma B (2018) Molecular Recognition between Aβ-Specific Single-Domain Antibody and Aβ Misfolded Aggregates. Antibodies 7
57. Thai NQ, Nguyen HL, Linh HQ, Li MS (2017) Protocol for fast screening of multi-target drug candidates: Application to Alzheimer's disease. J Mol Graph Model 77:121–129. https://doi.org/https://doi.org/10.1016/j.jmgm.2017.08.002
58. Gautieri A, Beeg M, Gobbi M, *et al*. (2019) The Anti-Amyloidogenic Action of Doxycycline: A Molecular Dynamics Study on the Interaction with Aβ42. Int. J. Mol. Sci. 20
59. Grasso G, Rebella M, Morbiducci U, *et al*. (2019) The Role of Structural Polymorphism in Driving the Mechanical Performance of the Alzheimer's Beta Amyloid Fibrils. Front Bioeng Biotechnol 7:83. https://doi.org/10.3389/fbioe.2019.00083

60. Villalobos Acosta DMÁ, Chimal Vega B, Correa Basurto J, *et al*. (2018) Recent Advances by In Silico and In Vitro Studies of Amyloid-β 1-42 Fibril Depicted a S-Shape Conformation. Int J Mol Sci 19:2415. https://doi.org/10.3390/ijms19082415
61. Hou S, Gu R-X, Wei D-Q (2017) Inhibition of β-Amyloid Channels with a Drug Candidate wgx-50 Revealed by Molecular Dynamics Simulations. J Chem Inf Model 57:2811–2821. https://doi.org/10.1021/acs.jcim.7b00452
62. Fan H-M, Gu R-X, Wang Y-J, *et al*. (2015) Destabilization of Alzheimer's Aβ42 Protofibrils with a Novel Drug Candidate wgx-50 by Molecular Dynamics Simulations. J Phys Chem B 119:11196–11202. https://doi.org/10.1021/acs.jpcb.5b03116
63. Battisti A, Palumbo Piccionello A, Sgarbossa A, *et al*. (2017) Curcumin-like compounds designed to modify amyloid beta peptide aggregation patterns. RSC Adv 7:31714–31724. https://doi.org/10.1039/C7RA05300B
64. Xi W, Wang W, Abbott G, Hansmann UHE (2016) Stability of a Recently Found Triple-β-Stranded Aβ1-42 Fibril Motif. J Phys Chem B 120:4548–4557. https://doi.org/10.1021/acs.jpcb.6b01724
65. Bitan G, Kirkitadze MD, Lomakin A, *et al*. (2003) Amyloid beta -protein (Abeta) assembly: Abeta 40 and Abeta 42 oligomerize through distinct pathways. Proc Natl Acad Sci U S A 100:330–335. https://doi.org/10.1073/pnas.222681699
66. Cheon M, Kang M, Chang I (2016) Polymorphism of fibrillar structures depending on the size of assembled Aβ17-42 peptides. Sci Rep 6:38196. https://doi.org/10.1038/srep38196
67. Kahler A, Sticht H, Horn AHC (2013) Conformational Stability of Fibrillar Amyloid-Beta Oligomers via Protofilament Pair Formation – A Systematic Computational Study. PLoS One 8:e70521
68. Kozakov D, Brenke R, Comeau SR, Vajda S (2006) PIPER: an FFT-based protein docking program with pairwise potentials. Proteins 65:392–406. https://doi.org/10.1002/prot.21117
69. Kozakov D, Beglov D, Bohnuud T, *et al*. (2013) How good is automated protein docking? Proteins 81:2159–2166. https://doi.org/10.1002/prot.24403
70. Kozakov D, Hall DR, Xia B, *et al*. (2017) The ClusPro web server for protein-protein docking. Nat Protoc 12:255–278. https://doi.org/10.1038/nprot.2016.169
71. Comeau SR, Gatchell DW, Vajda S, Camacho CJ (2004) ClusPro: an automated docking and discrimination method for the prediction of protein complexes. Bioinformatics 20:45–50. https://doi.org/10.1093/bioinformatics/btg371
72. Comeau SR, Gatchell DW, Vajda S, Camacho CJ (2004) ClusPro: a fully automated algorithm for protein-protein docking. Nucleic Acids Res 32:W96-9. https://doi.org/10.1093/nar/gkh354
73. Nosé S (1984) A unified formulation of the constant temperature molecular dynamics methods. J Chem Phys 81:511–519. https://doi.org/10.1063/1.447334
74. Hoover WG (1985) Canonical dynamics: Equilibrium phase-space distributions. Phys Rev A, Gen Phys 31:1695–1697. https://doi.org/10.1103/physreva.31.1695
75. Parrinello M, Rahman A (1981) Polymorphic transitions in single crystals: A new molecular dynamics method. J Appl Phys 52:7182–7190. https://doi.org/10.1063/1.328693
76. Nosé S, Klein ML (1983) Constant pressure molecular dynamics for molecular systems. Mol Phys 50:1055–1076. https://doi.org/10.1080/00268978300102851
77. Abraham MJ, Murtola T, Schulz R, *et al*. (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX 1–2:19–25. https://doi.org/https://doi.org/10.1016/j.softx.2015.06.001

78. Hess B (2008) P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. J Chem Theory Comput 4:116–122. https://doi.org/10.1021/ct700200b
79. Essmann U, Perera L, Berkowitz ML, *et al*. (1995) A smooth particle mesh Ewald method. J Chem Phys 103:8577–8593. https://doi.org/10.1063/1.470117
80. Touw WG, Baakman C, Black J, *et al.* (2015) A series of PDB-related databanks for everyday needs. Nucleic Acids Res 43: D364–D368. https://doi.org/10.1093/nar/gku1028
81. Kabsch W, Sander C (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 22:2577–2637. https://doi.org/10.1002/bip.360221211
82. van der Spoel D, van Maaren PJ, Larsson P, Tîmneanu N (2006) Thermodynamics of hydrogen bonding in hydrophilic and hydrophobic media. J Phys Chem B 110:4393–4398. https://doi.org/10.1021/jp0572535
83. Eisenhaber F, Lijnzaad P, Argos P, *et al*. (1995) The double cubic lattice method: Efficient approaches to numerical integration of surface area and volume and to dot surface contouring of molecular assemblies. J Comput Chem 16:273–284. https://doi.org/https://doi.org/10.1002/jcc.540160303
84. Srinivasan J, Miller J, Kollman PA, Case DA (1998) Continuum solvent studies of the stability of RNA hairpin loops and helices. J Biomol Struct Dyn 16:671–682. https://doi.org/10.1080/07391102.1998.10508279
85. Kollman PA, Massova I, Reyes C, *et al*. (2000) Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. Acc Chem Res 33:889–897. https://doi.org/10.1021/ar000033j
86. Kumari R, Kumar R, Lynn A (2014) g_mmpbsa--a GROMACS tool for high-throughput MM-PBSA calculations. J Chem Inf Model 54:1951–1962. https://doi.org/10.1021/ci500020m
87. Baker NA, Sept D, Joseph S, *et al*. (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. Proc Natl Acad Sci U S A 98:10037–10041. https://doi.org/10.1073/pnas.181342398
88. Tian W, Chen C, Lei X, *et al*. (2018) CASTp 3.0: computed atlas of surface topography of proteins. Nucleic Acids Res 46: W363–W367. https://doi.org/10.1093/nar/gky473
89. Kozlovskii I, Popov P (2020) Spatiotemporal identification of druggable binding sites using deep learning. Commun Biol 3:618. https://doi.org/10.1038/s42003-020-01350-0
90. Singh T, Biswas D, Jayaram B (2011) AADS--an automated active site identification, docking, and scoring protocol for protein targets based on physicochemical descriptors. J Chem Inf Model 51:2515–2527. https://doi.org/10.1021/ci200193z
91. Muscat S, Pallante L, Stojceski F, *et al*. (2020) The Impact of Natural Compounds on S-Shaped Aβ42 Fibril: From Molecular Docking to Biophysical Characterization. Int J Mol Sci 21:2017. https://doi.org/10.3390/ijms21062017
92. Kumar S, Tsai CJ, Ma B, Nussinov R (2000) Contribution of salt bridges toward protein thermostability. J Biomol Struct Dyn 17 Suppl 1:79–85. https://doi.org/10.1080/07391102.2000.10506606

Chapter 5. Enhancing the binding of the β-sheet breaker peptide LPFFD to the amyloid-β fibrils by aromatic modifications: A Molecular Dynamics Simulation study



5.1. Introduction

Alzheimer's disease is a fatal neurodegenerative disease and a primary source of dementia which affects millions of people worldwide who suffer from memory loss and cognitive deterioration [1][2]. Alzheimer's disease belongs to a class of neurodegenerative diseases caused due the abnormal processing of neuronal proteins [3]. Other examples of such diseases are Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, frontotemporal dementia, prion diseases and spinocerebellar ataxias [3]. Neuritic plaques, which consist of extracellular aggregates of amyloid fibrils, and intracellular aggregates (neurofibrillary tangles) formed by tau proteins are found in the brains of patients with Alzheimer's disease [2], [4].

The use of small molecule drugs to destabilize and prevent the aggregation of the misfolded amyloid-beta proteins is one of the therapeutic strategies for Alzheimer's disease [2]. Drug discovery and development for Alzheimer's disease is difficult and new drugs have not been approved since 2003 - no disease-modifying treatments, which are approved exist [5]. Small molecules-based drugs have attracted attention in treating various diseases, as they may be able to cross the blood-brain barrier.

Several efforts had been made over the past two decades to identify different classes of small molecules which could inhibit the aggregation of amyloid-beta fibrils into neurotoxic species and to destabilize the amyloid-beta fibrils. One class of such molecules includes the chelates and metal coordination complexes, which have demonstrated the inhibition of amyloid-beta aggregation and reduction in the toxicity of the amyloid fibrils. Man and group had shown the inhibition of fibrillation by the amyloid-beta peptide by iridium (III) and rhodium (III) metal complexes [6]. The clioquinol – zinc ion complex disrupted the amyloid-beta aggregates and could prevent further aggregation in a study by Raman *et al.* [7]. Similar observations were made by Barnham *et al.* who used platinum phenanthroline derivatives to inhibit amyloid-beta aggregation [8] and Iscen *et al.* who used Cobalt (III) Schiff bases [9]. While metal chelators had shown therapeutic potential in experimental studies, in clinical trials the metal chelators PBT_1 and PBT_2 had not shown promising results [10].

Several synthetic compounds were also designed to interact with the amyloid-beta fibrils. Doens and group identified four new hexahydropyrroloindoles (HPI) synthetic compounds which could disrupt the fibrils and prevent further aggregation [11]. Flavone hybrids were designed by Shi and group [12]; and Kaur and group designed and synthesized triazole-based

derivates which could inhibit the aggregation of the amyloid-beta fibrils [13]. The green tea polyphenol epigallocatechin-3-gallate (EGCG) was able to alter the structures of natively unfolded amyloid peptides into off-pathway and non-toxic structures in a study by Ernhoefer *et al.* [14]. Several antibodies were also studied for the treatment of Alzheimer's disease, however, due to their large size they were not able to cross the blood-brain barrier [15]. In clinical trials the monoclonal antibodies Bapineuzumab and Solanezumab were not successful [16], [17], and [18]. A monolonal antibody Aducanumab showed promising results in patients with mild Alzheimer's disease [19].

Conventional small molecule drugs are limited by their low selectivity indicated by side-effects in humans and a low affinity to the amyloid fibrils [20], [21]. These drawbacks can be overcome by peptide-based drugs which may serve as an alternative option to these chemical compounds. Protein therapeutics have a greater specificity for their targets because of their ability to have more interactions with them [21]. There are more than 100 peptide-based drugs in the ethical pharmaceutical market with a market share of about 10% [21]. More commonly, these peptides have a sequence of about 8 to 10 amino acids [21]. Peptide-based drugs with molecular weight greater than 500 Da are usually delivered by injection [21]. Peptide-based drugs are highly potent and selective with a broad range of targets. They are less toxic than small molecule drugs and do not accumulate in tissues. They are chemically and biologically diverse and are discoverable at peptide and/or nucleic acid levels [21]. However, peptide-based drugs have several disadvantages: poor metabolic stability, poor membrane permeability, poor oral availability, high cost, and poor solubility in some cases [21]. If the challenges of using peptide-based drugs are overcome, they can potentially be used to treat diseases effectively and reliably [21]. A comprehensive review of various peptide-based drugs that have been studied so far for the treatment of Alzheimer's disease has been provided by Goyal *et al.* [20].

Peptides obtained from the central hydrophobic core and the C-terminal of the $A\beta_{1-42}$ peptide can inhibit the aggregation of amyloid fibrils [20]. Tjernberg *et al.* systematically synthesized short peptides, which had sequences corresponding to the Aβ peptide in order to determine the minimum length of a binding sequence which was critical for binding [22]. They found that the residues in the positions 16-20 which form the central hydrophobic core, KLVFF, were the most important in the aggregation process and could inhibit the aggregation of fibrils [22]. This sequence is the self-recognition sequence and is the site in the fibrils where nucleation occurs. Based on this sequence, Soto *et al.* designed a pentapeptide LPFFD, which could destabilize Aβ fibrils and prevent their aggregation [23]. When this was injected into the brains of a rat, there was a reduction in Aβ protein deposition and a blockage of the formation of amyloid fibrils

(Soto *et al*., 1998). Similar neuroprotective effects were observed in rats when the LPFFD peptide was modified with an acetyl group (acetyl-LPFFD-amide) [24]. Several modifications to the LPFFD peptide were investigated, such as iAβ11 (RDLPFFPVRID) [25], which could inhibit amyloidogenesis and partially disassemble preformed Aβ fibrils in vitro; Ac-LPFFD-amide [26], LPYFD-amide [27], [28], Ac-LPFFN-NH2 [29], and LPfFFD-PEG [30], cholyl-(LVFFA)−OH [31], iAβ11 [32], and trehalose conjugated LPFFD [33].

Viet *et al*. considered the binding of all 8000 possible combinations of amino acid trimers in order to find the combination which could bind the best to the amyloid-β fibrils [34]. They found that combinations comprising of aromatic amino acids and proline could bind better to the fibrils than other amino acids [34]. They also found a significant correlation between binding affinities of aromatic trimers and their ability to depolymerize Aβ [34]. Their in silico and in vitro study showed that the presence of tryptophan and proline always enhanced the binding of trimers to the amyloid fibrils [34].

There is a significant occurrence of aromatic residues in many amyloid-related proteins [35], [36] which play a role in the self-assembly of proteins by aromatic stacking [37][38][39][40]. It has been shown that by blocking the interactions of these aromatic residues, the self-assembly of fibrils can be inhibited. For example, it was shown that Congo red inhibits insulin fibril formation by interacting with the aromatic phenylalanine residues in the insulin dimer [41]. Polyphenols inhibit fibril assembly by virtue of their phenol rings which stack with aromatic amyloidogenic residues differently compared to benzene rings [42]. This phenol ring must be a part of molecule which can bind to the fibrils well by other interactions such as hydrogen bonds [43]. A study by Xie *et al*. had revealed that the hydrophobic and aromatic stacking interactions between phenylalanine residues in the hydrophobic core of the fibrils played an essential role in the formation of β-barrels, and bilayer β-sheets [44]. The substitution of the phenylalanine residues by tryptophan reduced the β-sheet content and increased the population of disordered aggregates [44]. An orally absorbed dipeptide Tyr-Pro improved cognitive impairment in Aβ25-35-induced mice [45].

In our present study we investigated the effect of modifying the LPFFD peptide by aromatic amino acids. The model of the amyloid fibrils used in the present study is an S-shaped disease-relevant form of the fibrils. This model is known to be a difficult target for drug candidates because of the absence of suitable binding sites for conventional molecules, particularly in the hydrophobic core region, as these side chains remained buried in solution. A study by Marondedze *et al.* [46] found that the binding sites in the model used in the present study was undruggable due to the absence of hydrophobic surfaces for the binding of the ligand.

In this context, the design of ligands, which can serve as either drug candidates or as probes for the detection of the fibrils is important for this particular model of the fibrils. The presence of aromatic resides in the N-terminal region were a motivation for the investigation of the binding affinity of the LPFFD peptide when it was modified by the addition of aromatic amino acids. Such a modification allowed the increased prevalence of aromatic interactions, which could lead to better binding. We considered the aromatic amino acids tryptophan, phenylalanine, tyrosine and histidine. The aromatic amino acids have different properties: phenylalanine is hydrophobic, histidine is polar, and tryptophan & tyrosine are amphipathic. Tryptophan is a bulky amino acid, while proline has the smallest size compared to the aromatic amino acids. Based on these differences, their interactions with the amyloid β fibrils may also be expected to be different. We characterized these interactions with the fibrils and the binding mechanism of these different ligands. We compared the frequency of sustained contacts between the aromatic, hydrophobic, and charged residues of the ligands and the fibrils and their associated binding affinities. Our results indicate that tryptophan enhanced the binding of the LPFFD peptide more than the other aromatic amino acids.

5.2. Materials and Methods

The crystal structure for the $A\beta_{1-42}$ fibrils was obtained from the Protein Data Bank (PDB ID: 2NAO) [47]. This model has a dimeric double horseshoe structure. This structure comprises of 6 chains of the fibrils composed of two molecules per fibril layer. Residues 15 - 42 form a double-horseshoe-like-cross-β-sheet entity and have maximally buried hydrophobic side chains. Residues 1-14 are partially ordered and in a β-strand conformation. This structure is shown in Fig. 5.1. For convenience, the six chains of the fibrils will be referred to as chains A, B, C, D, E, and F. This structure has been characterized and studied in many previous studies [48], [49], [50], [51], and [52]. Fig. 5.1 shows the model of the fibrils that was used in this study. The six chains are labelled A, B, C, D, E, and F for convenience. The N-termini of these chains are labelled by the corresponding alphabet letter. In this structure, the side chains of residues comprising the central hydrophobic core and the hydrophobic C-terminal residues, namely the residues occupying positions 16-42, are buried and not accessible to the solvent. Thus, they are packed tightly.

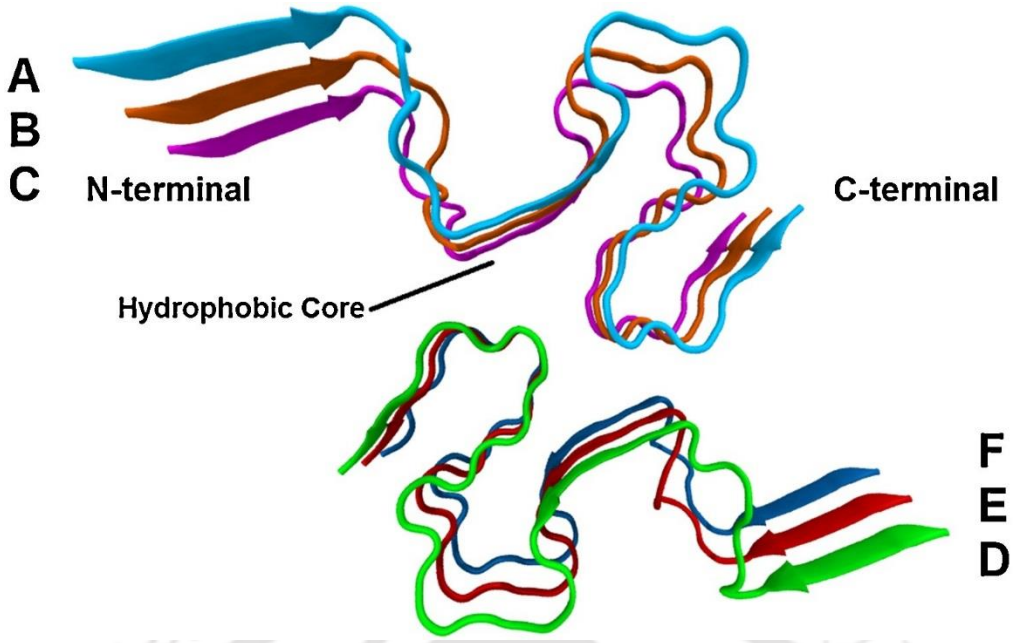


Fig. 5.1. The model of the $A\beta_{1-42}$ fibrils used in this study. This model comprises of 6 chains of the fibrils. The different chains are represented by a specific color: chain A is cyan, chain B is orange, chain C is magenta, chain F is blue, chain E is red and chain D is green. The N-termini of the chains are denoted by the alphabet corresponding to the chain.

The structure of the monomer $A\beta_{17-21}$ (LVFFA) was extracted from the structure of $A\beta_{10-35}$ peptide (PDB code 1hz3). The peptide LPFFD was constructed by the mutations V18 P and A21D on the derived peptide. The mutations were done using UCSF Chimera 1.13.1 [53]. Modifications to these peptides were made by adding aromatic residues to the termini using Pymol 2.3.1. The residues which were added to the termini were tryptophan, phenylalanine, tyrosine, histidine and proline. For convenience, the ligand chains are labelled as chain G.

The ligands were then docked to the receptor amyloid-β fibrils by using the automated protein docking server ClusPro [54], [55], [56], [57], and [58]. The ClusPro algorithm uses a rigid body method for docking. ClusPro yields acceptable or medium accuracy models in many tests and can therefore identify best near native conformations [55]. The recommended default procedure was followed in ClusPro. Based on favorable desolvation energies and electrostatics, the top 2000 models were clustered and ranked by the ClusPro algorithm [57].

We performed three control simulations of the $A\beta_{1-42}$ fibrils in water, and two control simulations of the system of the LPFFD peptide and the $A\beta_{1-42}$ fibrils. We classified our simulations involving the modified ligands into two sets: Set 1 and Set 2. In Set 1 of the simulations, the LPFFD peptide was modified by adding an amino acid monomer to its N- and C-term7ini. In Set 2 of the simulations, an amino acid trimer was added to the termini. Table

5.1 describes the nomenclature that we use to describe every system. Two replicas of each system are considered.

Table 5.1. Nomenclature and composition of all the systems considered. Two replicas of each system are considered.

Nomenclature	Replica	Composition	Set
C1	First	$A\beta_{1-42}$	Control
C2	Second	$A\beta_{1-42}$	Control
C3	Third	$A\beta_{1-42}$	Control
L1	First	LPFFD - $A\beta_{1-42}$	Control
L2	Second	LPFFD - $A\beta_{1-42}$	Control
W1	First	WLPFFDW - $A\beta_{1-42}$	1
W2	Second	WLPFFDW - $A\beta_{1-42}$	1
F1	First	FLPFFDF - $A\beta_{1-42}$	1
F2	Second	FLPFFDF - $A\beta_{1-42}$	1
Y1	First	YLPFFDY - $A\beta_{1-42}$	1
Y2	Second	YLPFFDY - $A\beta_{1-42}$	1
H1	First	HLPFFDH - $A\beta_{1-42}$	1
H2	Second	HLPFFDH - $A\beta_{1-42}$	1
WWW1	First	WWWLPFFDWWW - $A\beta_{1-42}$	2
WWW2	Second	WWWLPFFDWWW - $A\beta_{1-42}$	2
FFF1	First	FFFLPFFDFFF - $A\beta_{1-42}$	2
FFF2	Second	FFFLPFFDFFF - $A\beta_{1-42}$	2
YYY1	First	YYYLPFFDYYY - $A\beta_{1-42}$	2
YYY2	Second	YYYLPFFDYYY - $A\beta_{1-42}$	2
HHH1	First	HHHLPFFDHHH - $A\beta_{1-42}$	2
HHH2	Second	HHHLPFFDHHH - $A\beta_{1-42}$	2

*Molecular Dynamics Simulations***:** Each system was placed in a cubic box of size 10.57 nm of TIP3P water. The simulations employed the Amber99SB-ILDN all-atom force field. Sodium ions were added to each system to make the system neutral in charge. Following energy minimization using the steepest descents algorithm, the 25 systems were equilibrated in two steps with position restraints applied to heavy atoms throughout. An equilibration was carried out under NVT conditions at 300 K using coupling to the Nosé Hoover thermostat [59], [60]. Following this, an equilibration was carried out under NPT conditions with the system coupled

to a Parinello-Rahman barostat [61], [62] and a thermostat at 300 K. All the equilibrations were done for 200 ps. Finally, the MD simulation with position restraints removed was run for 300 ns using GROMACS 5.1.1 [63]. The periodic boundary conditions were applied in all the directions. The constraint algorithm used for bond lengths was P-LINCS [64]. The neighbour search cut-off was at approximately 1 nm. The fast-smooth particle mesh Ewald summation method was used to calculate the long-range electrostatic interactions and the Fourier grid spacing was 0.16 nm [65].

The minimum distance of a protein system to its periodic image was measured to check for artefacts. The secondary structure was calculated using the do_dssp [66], [67] tool in GROMACS. Hydrogen bonds were calculated using the hbond [68] tool in GROMACS. The root squared mean deviation (RMSD) and root mean squared fluctuation (rmsf) were calculated using the GROMACS tools rms and rmsf. Python 3.7.3 was used to calculate averages and for plotting results. Contact surface areas were calculated using VMD [69] based on the following formula [70]:

$$Contact\ Surface\ Area = \frac{SASA_{protein} + SASA_{ligand} - SASA_{complex}}{2} \qquad \text{(Equation 5.1)}$$

where SASA represents the solvent accessible surface area.

The MM/PBSA method was used to calculate the binding free energies of the amyloid protofibrils with the modified ligands. In MM/PBSA, the binding energy is calculated as follows [71], [72]:

$$\Delta G^{bind} = \Delta EMM + \Delta G^{psolv} + \Delta G^{npsolv} - T\Delta S \qquad \text{(Equation 5.2)}$$

Where ΔEMM is the molecular mechanics contribution to the binding free energy in vacuum, ΔG^{psolv} is the polar contribution to the solvation energy (calculated by solving the Poisson-Boltzmann equation) and ΔG^{npsolv} is the non-polar contribution to the solvation energy (calculated by using the solvent-accessible surface area (SASA) model.) Entropic terms were not included in calculating the binding energies of the ligands.

The g_mmpbsa tool [73], [74] was used to calculate the binding free energies using the single trajectory protocol. Snapshots from the last 10 ns of the production run extracted every 100 ps were used for the analysis. Energetic components are expressed as an average with standard

deviations. The contributions to the binding energy made by individual residues were also calculated using 1000 bootstraps in the g_mmpbsa program. The solute and solvent dielectric constants were taken as 4 and 80 respectively. The non-linear Poisson-Boltzmann equation was solved to calculate the polar solvation energy.

Representative snapshots from the most occupied conformation cluster for each system were derived from the trajectory using the method of Daura *et al*. [75] using a cut-off of 1.5 Å using the GROMACS modules gmx cluster with a 1.5 Å cut-off on the active site residues of the fibrils. Sustained contacts between ligand and fibril residues were calculated using the vmd script contactFreq.tcl.

5.3. Results and Discussion

5.3.1. Docking results

In our present study, we used the best ranked structure obtained from the docking algorithm for all the systems considered. The algorithm produced structures in which the ligand was in similar region in all systems, enabling the easy comparison of the different systems. In these structures, the ligands were in the vicinity of the region spanned by the N – terminal residues 1 – 14 of chains D, E, and F of the fibrils. The structure obtained from the docking algorithm for the LPFFD – Aβ system is shown in Fig. 5.2. There were very few docking poses predicted by ClusPro, which were similar in nature due to the symmetric nature of the fibril structure. These docking poses were either in the N-terminal region of chains D, E and F used in the present study or in the similar N-terminal region of chains A, B, and C. In the case of the control LPFFD peptide, only one docking pose was predicted by ClusPro which was used as the basis for the modified ligands. In the case of the modified ligands, the highest ranked model also corresponded to this same region. In the absence of experimentally verified structures, we feel that the results produced by ClusPro are reasonable to consider for the present study. The MD simulations, which were extended up to 300 ns allowed for the refinement of the docked poses.

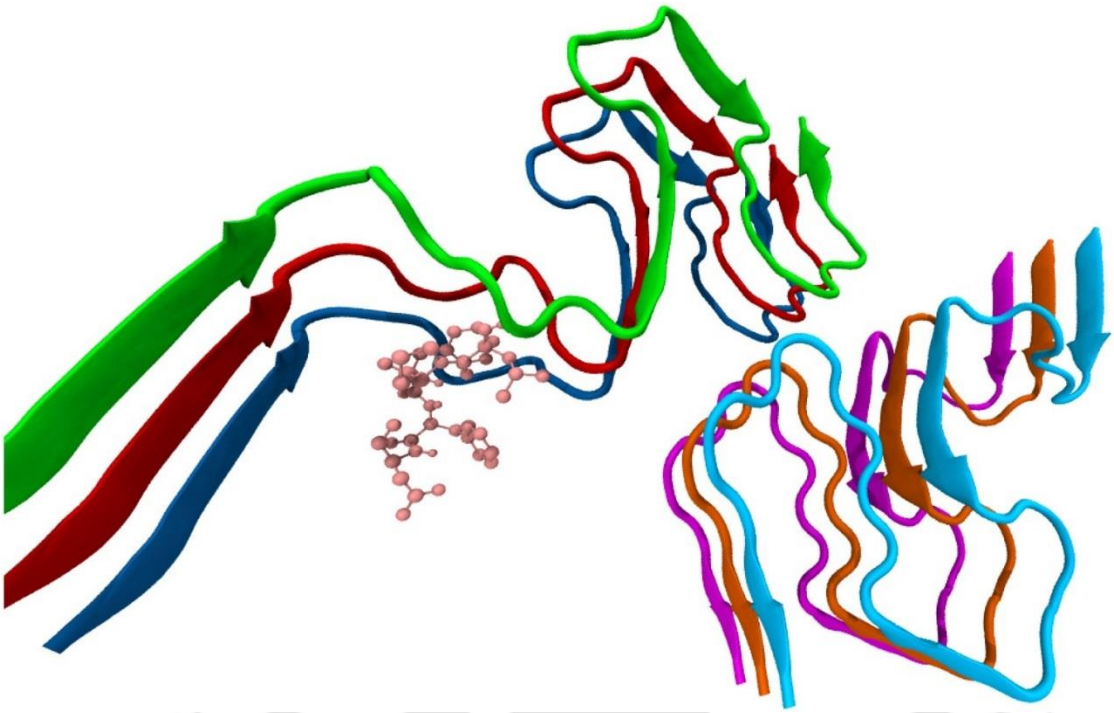


Fig. 5.2. The docking result of the LPFFD – Aβ system that was used for the simulations. LPFFD docked to a region corresponding to the N-terminal residues of chains D, E and F. chain F is blue, chain E is red and chain D is green.

The binding region in the fibrils, which was spanned by the first 14 residues, and in some cases up to the 16th residue of the fibrils target, contains charged, aromatic, and hydrophobic amino acids. It consists of the negatively charged residues Asp 1, Glu 3, Asp 7, and Glu 11 which are on the surface of the molecule [76]. There is a cluster of positively charged residues formed by His 13, His 14 and Lys 16, which are hypothesized to participate in the activation of microglia, which are involved in the inflammatory response to Aβ observed in Alzheimer's disease [76]. The positively charged residue Arg 5 is in the neighbourhood of negatively charged residues. The aromatic fibril residues are Phe 4, His 6, Tyr 10, His 13 and His 14. The residues within 4 Å of the LPFFD ligand in the initial structure were Tyr 10D, Val 12D, His 13D, His 14D, Phe 4E, His 6E, Gly 9E, Tyr 10E, Glu 11E, Val 12E, His 13E, His 14E, Phe 4 F, His 6 F, Ser 8 F, Gly 9 F, Tyr 10 F, Val 12 F, and His 13 F.

The model of the fibrils used in this present study is a disease-relevant form of the Aβ aggregates, and hence an important and relevant model to consider in the design of novel peptides which are candidates for drug development. The docking algorithm did not produce structures in which the ligand could interact with the residues in the hydrophobic core of the fibrils, as these were tightly packed and inaccessible to the ligands, as also observed in a previous study by Jakubowski *et al.* [77] who studied the interaction of curcumin with Aβ fibrils. There are three

distinct regions of the fibrils, which have been used to target by drug candidates for the fibrils: the N-terminal region, the central hydrophobic core and the C-terminal region [78]. Oligomer-defining interactions involving the Aβ peptide N-terminus may be important in production of the neurotoxic forms and thus should not be neglected [79]. It has been suggested that blocking the N-terminus interactions of the Aβ oligomers can block further aggregation [80]. Aducanumab, a drug candidate, which was in Phase 3 trials, was shown to bind to residues 3 to 6 of the fibrils [80]. Further, glutamate or N-methyl-d-aspartate receptors bind to the N-terminus residues of the fibrils [80]. Residues in this region have shown to play a role in the transition of oligomers to mature fibrils and the stabilization [79] and the residue tyrosine in the 10^{th} position is important in this process [81]. It has been hypothesized that the imidazole sidechain of histidine 14 which modulates the interaction with the negatively charged head group of phosphatidylserine of cell membranes and strategies inhibiting this interaction may have therapeutic potential for Alzheimer's disease [82]. Another study showed that His6, His13, and His14 residues in Aβ 1-40 peptide significantly affect oligomeric equilibria [83]. This region has also been shown to play a role in the formation of an interface in the zinc-mediated oligomerization of the fibrils [84]. Hence, the consideration of the residues in the N-terminus is important in the design of novel drug candidates. A molecular dynamics study of the interaction of the LPFFD peptide with the Aβ monomers [85] showed that the LPFFD peptide could interact with the N-terminus residues Phe 4, His13, Asp7 and Val 12. In another study [86] it was able to interact with the residues Tyr 10 and His 14. In our simulations, the ligands considered in our study were also able to bind to C-terminal residues of chains A, B, and C which is another important region to block aggregation [87], and is discussed in later sections. A representation of the fibril sequence is shown in Fig. 5.3.

{ASP 1}{ALA 2}{GLU 3}{PHE 4}{ARG 5}{HIS 6}{ASP 7}{SER 8}{GLY 9}{TYR 10}{GLU 11}{VAL 12}{HIS 13}{HIS 14}{GLN 15}{LYS 16}{LEU 17}{VAL 18}{PHE 19}{PHE 20}{ALA 21}{GLU 22}{ASP 23}{VAL 24}{GLY 25}{SER 26}{ASN 27}{LYS 28}{GLY 29}{ALA 30}{ILE 31}{ILE 32}{GLY 33}{LEU 34}{MET 35}{VAL 36}{GLY 37}{GLY 38}{VAL 39}{VAL 40}{ILE 41}{ALA 42}

Figure 5.3. Sequence representation of the amyloid-beta peptide. The residues highlighted in blue are positively charged, the residues in red are negative and the green residues are the aromatic residues.

5.3.2. Stability of the fibrils

The stability of the control systems was assessed by the root mean squared deviation (RMSD) of the C-α atoms with respect to the equilibriated structures. The RMSD of C-α atoms in all the control simulations reached equilibrium in the course of the MD simulations, indicating the stability of the fibrils. The increase in the RMSD and its fluctuations were from the contributions of the flexible domains of the fibrils. A plot of the RMSF shows the flexible residues in the N-terminal of the fibrils. This region tended to move closer to the hydrophobic core of the fibrils in order to maximize contact in the course of the simulations, as it did not have any other chains adjacent to it. The flexibility of the control systems was further assessed by the root mean squared fluctuations (RMSF) of the residues of the individual chains of the control systems, which is a measure of the mobility of atoms in these residues. These values are shown in Fig. 5.4 and 5.5. The 3D structure resolved by solid-state NMR spectroscopy and mass-per-length measurements from EM [47] for this system of the amyloid fibrils suggested that residues 1 - 14 of the N – terminal are dynamic and disordered and not rigid. This is reflected in the RMSF values of this segment in the control systems as can be seen in Fig. 5.5. The residues in the 15 – 42 regions were stable, as they contain maximally buried side-chain atoms [47], and did not have any significant fluctuations. The residues 16 - 21 of the central hydrophobic were the most stable as the side chains of these hydrophobic residues were buried in the solution. Hence, the flexibility of the N – terminal 1 – 14 residues contributed the most to the RMSD and RMSF values in the control systems. The RMSD of C-α atoms with respect to the equilibriated structures of all the ligand-fibril complexes was calculated. These values are shown in Fig. 5.6 and 5.7. The RMSD of C -α atoms in all the simulations reached equilibrium in the course of the MD simulations, indicating the stability of the fibrils.

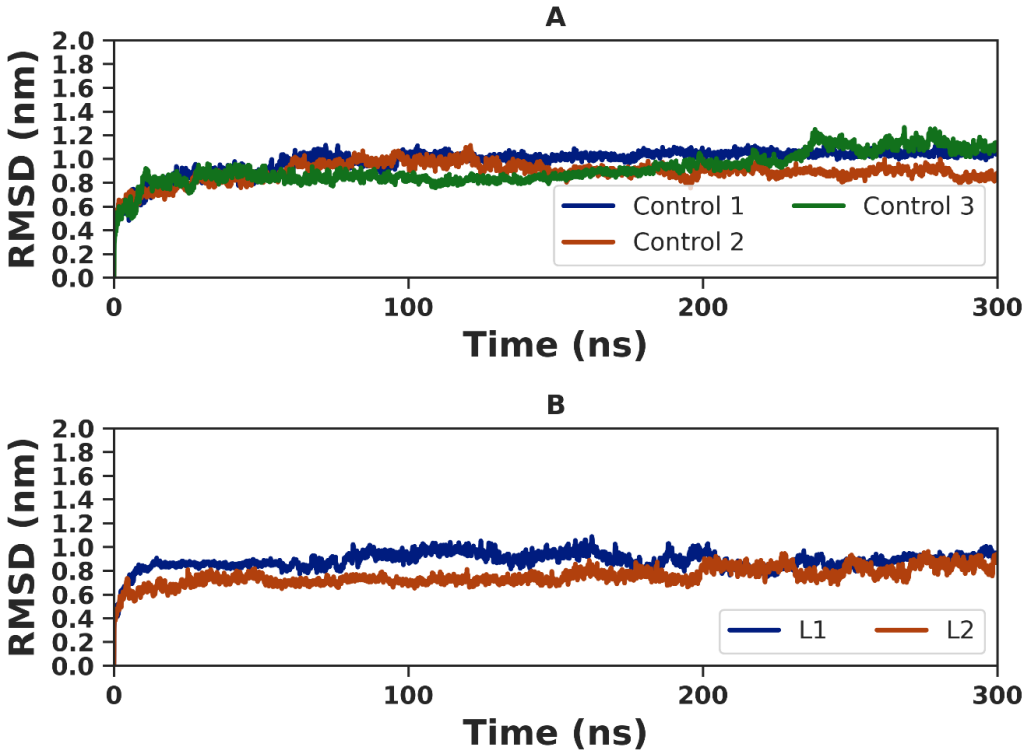


Figure 5.4. RMSD values in the **A.** control systems and **B.** in the LPFFD-amyloid-beta systems.

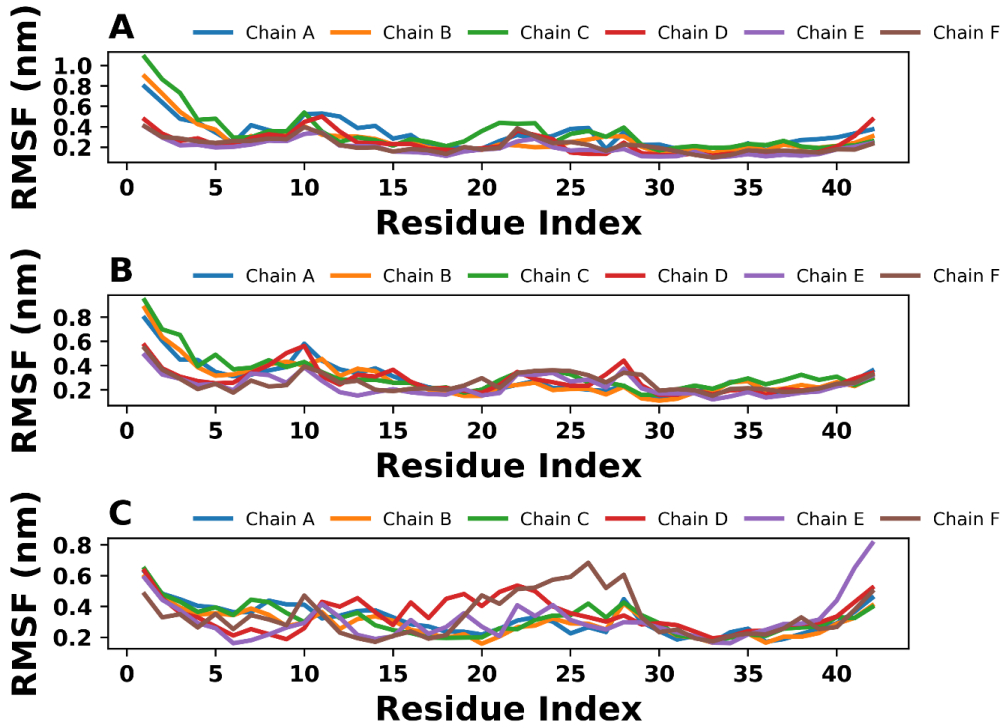


Figure 5.5. RMSF values of the control systems.

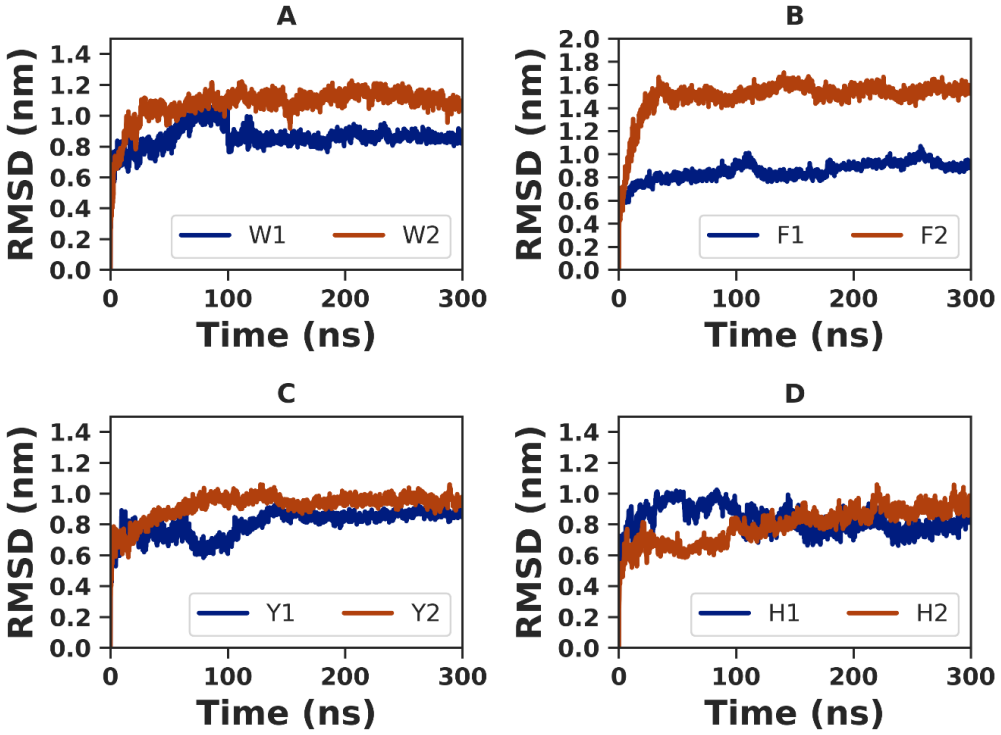


Figure 5.6. The RMSD values in set 1 of the simulations

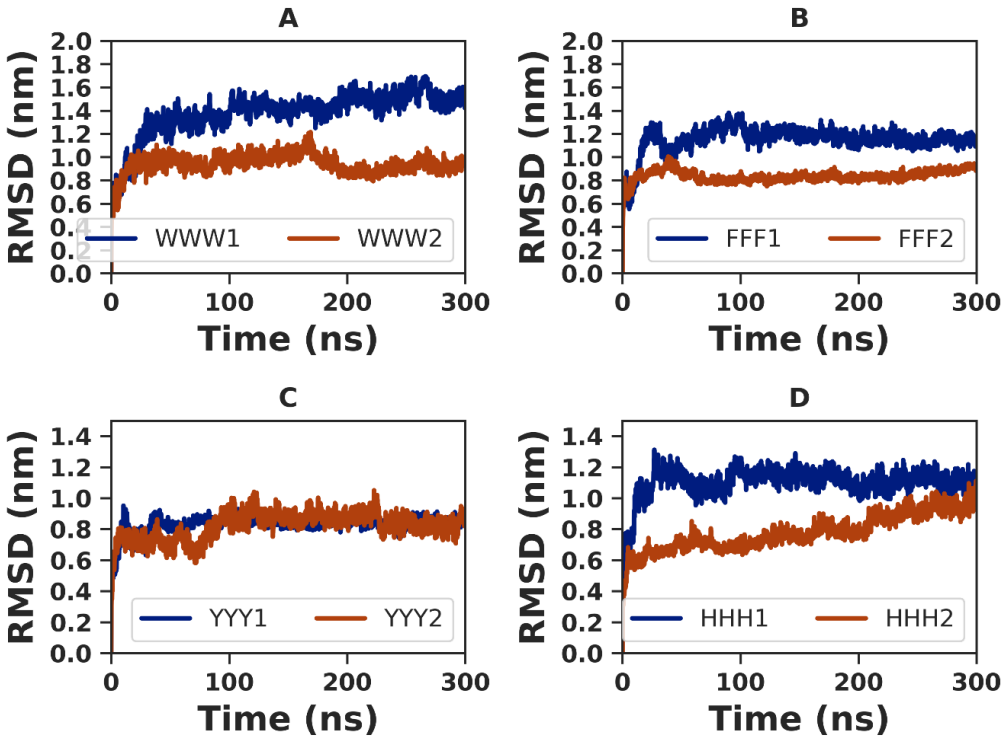


Figure 5.7. RMSD values of the systems in set 2 of the simulations.

5.3.3. Sustained contacts

We quantified the duration of important contacts between ligand and fibril residues by using the definition of Ghosh *et al.* [88] in which sustained contacts are defined as those contacts between the ligand and fibril residues which lasted for more than 40% of the total simulation time. We considered four types of interactions which made an important contribution to the binding of the ligand residues to the fibril residues: contacts between aromatic-aromatic residues (which accounted for aromatic interactions), hydrophobic-hydrophobic residues (which accounted for hydrophobic interactions), salt-bridge forming contacts, and contacts between residues which were capable of forming cation-pi interactions. We also measured the hydrogen-bonding ability of the various ligands to the fibrils.

Aromatic interactions play an important role in biological molecular recognition [89] and are important for ligand binding. A cut-off of 4 Å was used to account for hydrophobic and electrostatic interactions and 6.5 Å for cation-pi interactions.

5.3.3.1. The control LPFFD systems

There were two replicates of the LPFFD systems. In both these replicates, the LPFFD ligands were not able to bind to the fibrils for more than 100 ns. These form the basis for the comparison of the various modifications of the LPFFD peptide considered in this study.

5.3.3.2. Sustained contacts in the systems in Set 1

We next considered the sustained contacts between the ligands and fibrils in Set 1 of the simulations. Fig. 5.8 shows a plot of the sustained contacts in all the systems. The W1 ligand was extended as a coil while the W2 ligand tended to adopt a 3_{10} helical structure. Due to this difference in the secondary structure adopted by the ligands, the W1 ligand was able to make more sustained contacts with the amyloid fibril residues, particularly with those in the C-terminal region of chains A, B and C. The F1 and F2 ligands made a similar number of sustained contacts with the fibrils and had similar binding mechanisms. The Y1 ligand made more sustained contacts with the fibrils than the Y2 ligand as the latter did not make any contacts with the residues in the C-terminal region of chains A, B and C. Likewise, the H1 ligand made

significantly more sustained contacts with the fibril residues than the H2 ligand on account of making sustained contacts with the C-terminal residues of chains A, B, and C.

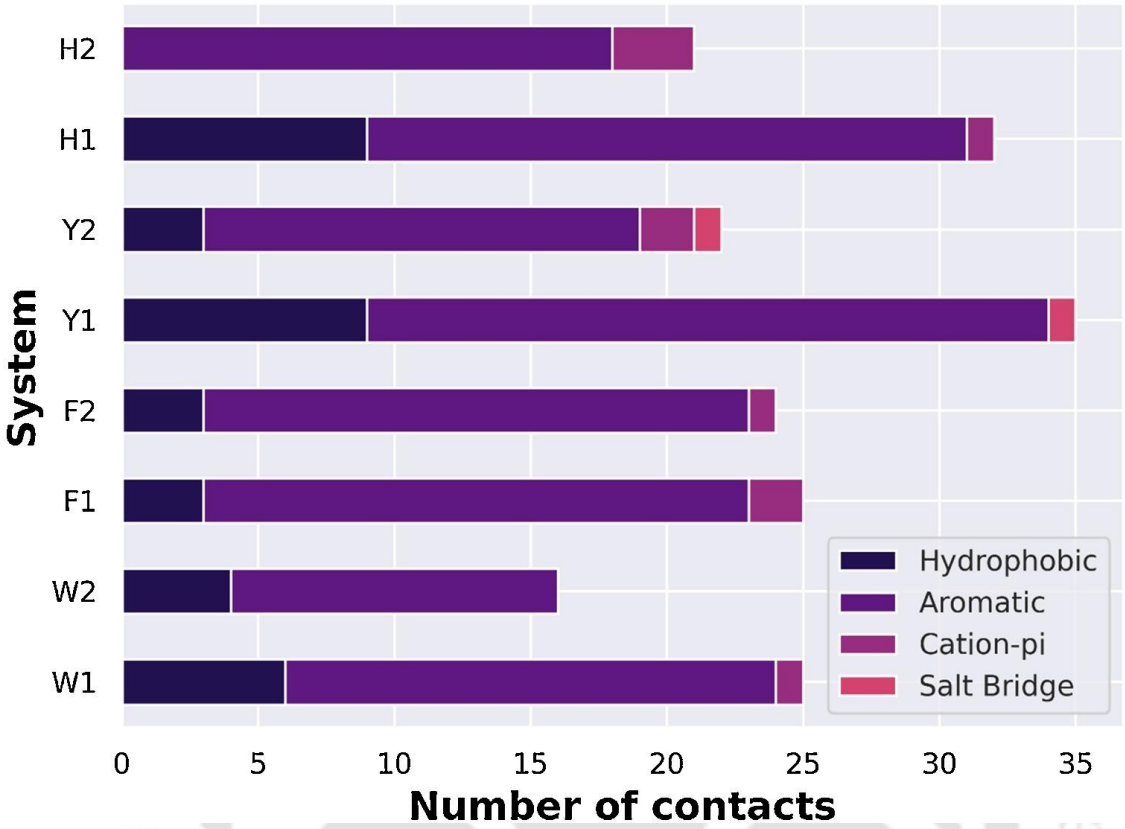


Fig. 5.8. A plot comparing the number of sustained contacts in each of the systems of set 1

The Y1 ligand made the maximum number of sustained contacts, while the H1 ligand made the second highest number of sustained contacts. An important observation was that the number of sustained contacts was higher when the ligands could establish new contacts with the C-terminal region residues of chains A, B and C in addition to the N-terminal region residues of chains D, E and F where they were initially docked. The Y1 and H1 ligands made the maximum number of sustained contacts on account of the highest number of hydrophobic and aromatic contacts formed with the fibril residues.

In the various ligands considered, an aromatic amino acid was added to the N- and C-termini of the LPFFD peptide. In order to investigate which of these aromatic residues was able to form more contacts with the fibrils, we counted the number of sustained contacts formed by these aromatic residues at both ends of the ligand. Fig. 5.9 shows the number of contacts made by the N-terminal aromatic residues of the modified ligands. There were no sustained contacts formed by the aromatic residues in the C-terminal of the modified ligands, implying that the

binding was driven by the aromatic amino acids at the N-termini. A significant number of sustained contacts were made by these N-terminal aromatic amino acids with respect to the total number of sustained contacts, particularly in the cases of the W1 and Y2 ligands.

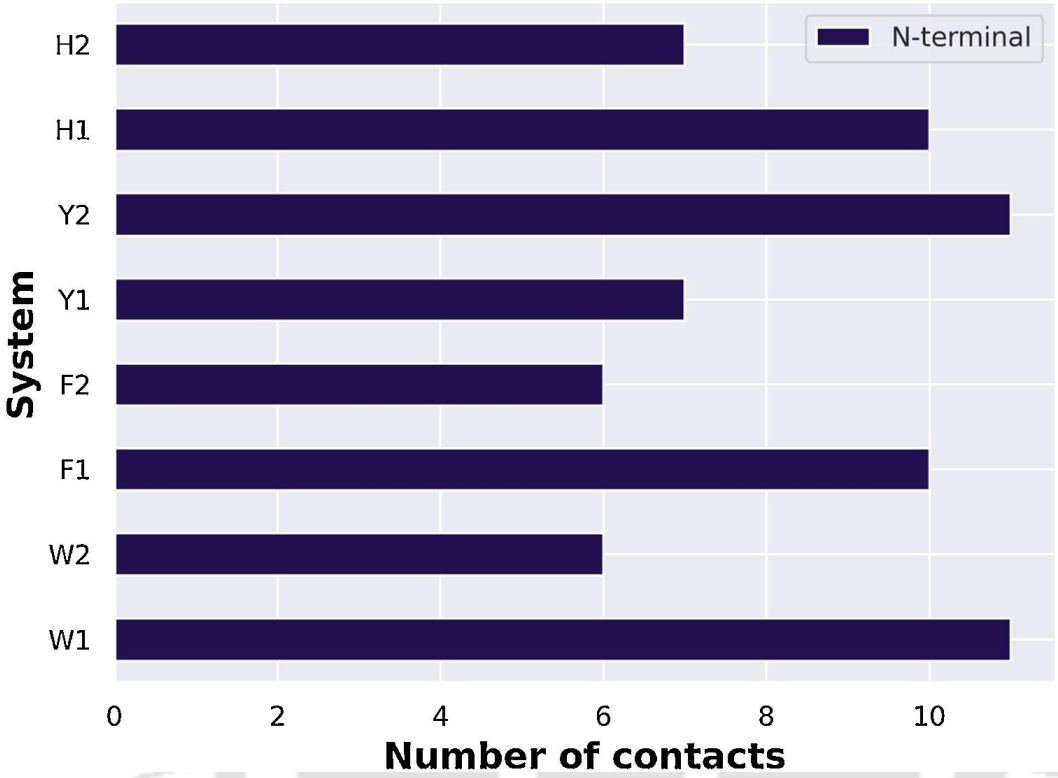


Fig. 5.9. Number of contacts made by the N-terminal aromatic amino acid with the fibril residues in the modified LPFFD ligand systems. There were no contacts made by the C-terminal amino acids with the fibril residues. The N-terminal Trp and Tyr residues in the W1 and Y2 ligands made the highest number of contacts with fibril residues.

5.3.3.3. Sustained contacts in the systems in Set 2

We next consider the sustained contacts in the systems in Set 2 of the simulations in which the LPFFD ligand was modified by adding an aromatic trimer to its termini. Fig. 5.10 shows the number of sustained contacts in all the systems in set 2.

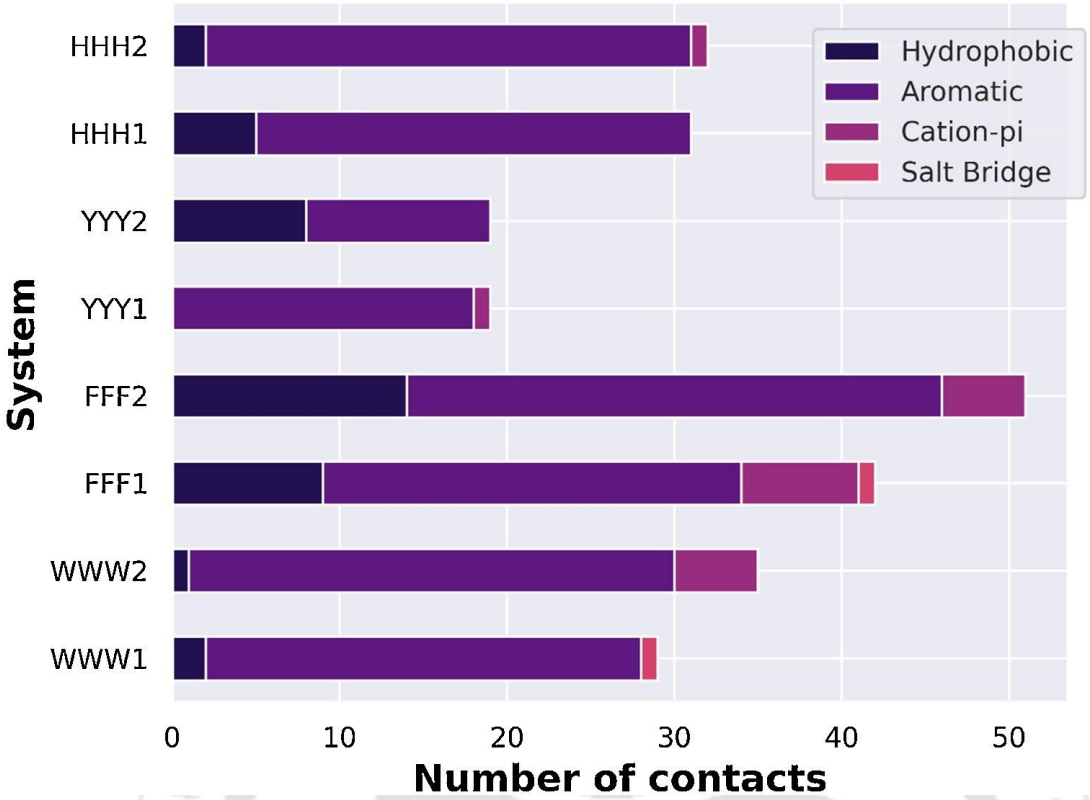


Fig. 5.10. A plot comparing the number of sustained contacts in each of the systems of set 2. The contacts are classified as aromatic contacts, hydrophobic contacts, cation-pi and salt bridge contacts. The FFF ligands made the highest number of contacts with the fibril residues.

The WWW2 ligand made more sustained aromatic and cation-pi contacts with the fibrils than the WWW1 ligand. The ligand in the both the FFF1 and FFF2 systems made significantly more sustained aromatic contacts with the fibril residues than the ligands in the other systems. The FFF2 system had the greatest number of sustained contacts between the ligand and the fibrils in set 2 of the simulations. The dual nature of phenylalanine as an aromatic and hydrophobic residue enabled it to make several contacts with fibril residues. Phenylalanine is an important residue in the sequence of the $A\beta_{1-42}$ peptide for the aggregation of the Aβ peptides as the side chains of this residue tends to interact with the side chain of other Phe residues [90]. It occurs at positions 4, 19 and 20 in the fibrils. The Phe-modifications of the LPFFD peptide enhanced its binding to the fibrils.

The YYY1 and YYY2 ligands made the same number of sustained contacts with the fibril residues, which was the least in set 2 of the simulations. The HHH1 and HHH2 ligands made a similar number of sustained contacts with the amyloid fibrils. The HHH2 ligand adopted an alpha-helical and 310 helix structure in the course of the simulations. Except in the case of the

YYY1 and YYY2 systems, the sustained aromatic contacts made a significant contribution to the total number of contacts.

A comparison between the numbers of sustained contacts formed by the aromatic trimers at the N- and C-termini of the ligands revealed that in set 2 of simulations, the C-terminal aromatic trimers made a contribution to the total number of sustained contacts in all the systems except the YYY2 system. This is shown in Fig. 5.11. In the case of the WWW2 and HHH1 systems, these contributions were more prominent than the N-terminal trimers.

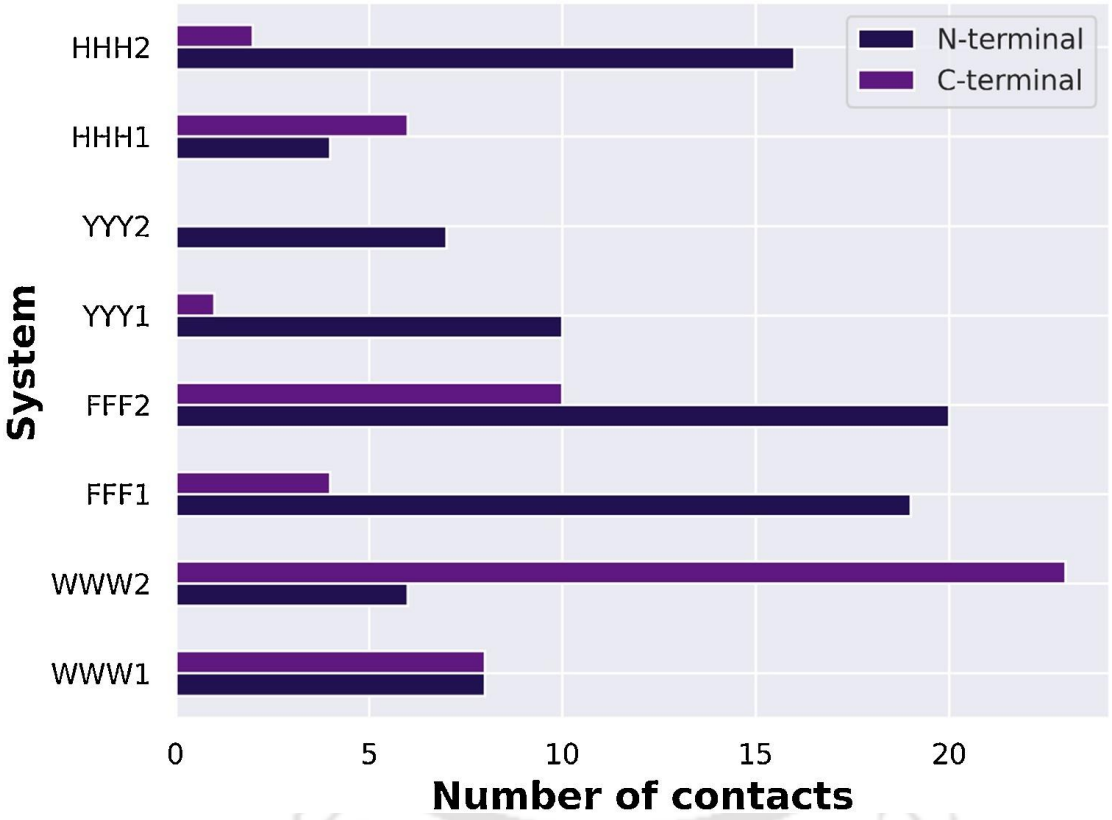


Fig. 5.11. A comparison of the number of sustained contacts made by the N-terminal and C-terminal aromatic trimers in the various ligands in set 2 of the simulations. The C-terminal aromatic amino acids made more sustained contacts in the WWW2 and HHH1 systems than the corresponding N-terminal aromatic amino acids of the ligands.

5.3.4. Hydrogen Bonding

Hydrogen bonds between receptor and drug molecules are an important criterion for binding affinities [91]. We found that the hydrogen bonding abilities of the ligands in set 1 of the simulations were similar, except in the case of the Y1 ligand which made significantly more

hydrogen bonds. In set 2 of the simulations, the ligands with the best hydrogen bonding abilities were the WWW1 and YYY1 ligands. The average number of hydrogen bonds between the ligands and fibrils are shown in Table 5.2

Table 5.2. Average number of hydrogen bonds between the ligands and the fibrils.

System	Average Number of Hydrogen Bonds
W1	2.38
W2	1.93
F1	2.92
F2	2.54
Y1	6.51
Y2	2.73
H1	2.69
H2	3.24
WWW1	8.75
WWW2	6.13
FFF1	4.99
FFF2	4.76
YYY1	7.41
YYY2	1.35
HHH1	2.00
HHH2	2.34

5.3.5. Binding energies in Set 1 of the simulations

In set 1 of the simulations, the Y1 ligand had the highest binding affinity to the fibrils. The W1, W2, F1 and F2 ligands had similar binding affinities while the H1 ligand had the lowest binding affinity. Fig. 5.12 shows a comparison of the binding energies. The binding energies are summarized in Table 5.3 along with the associated contact surface areas in all the systems in this set.

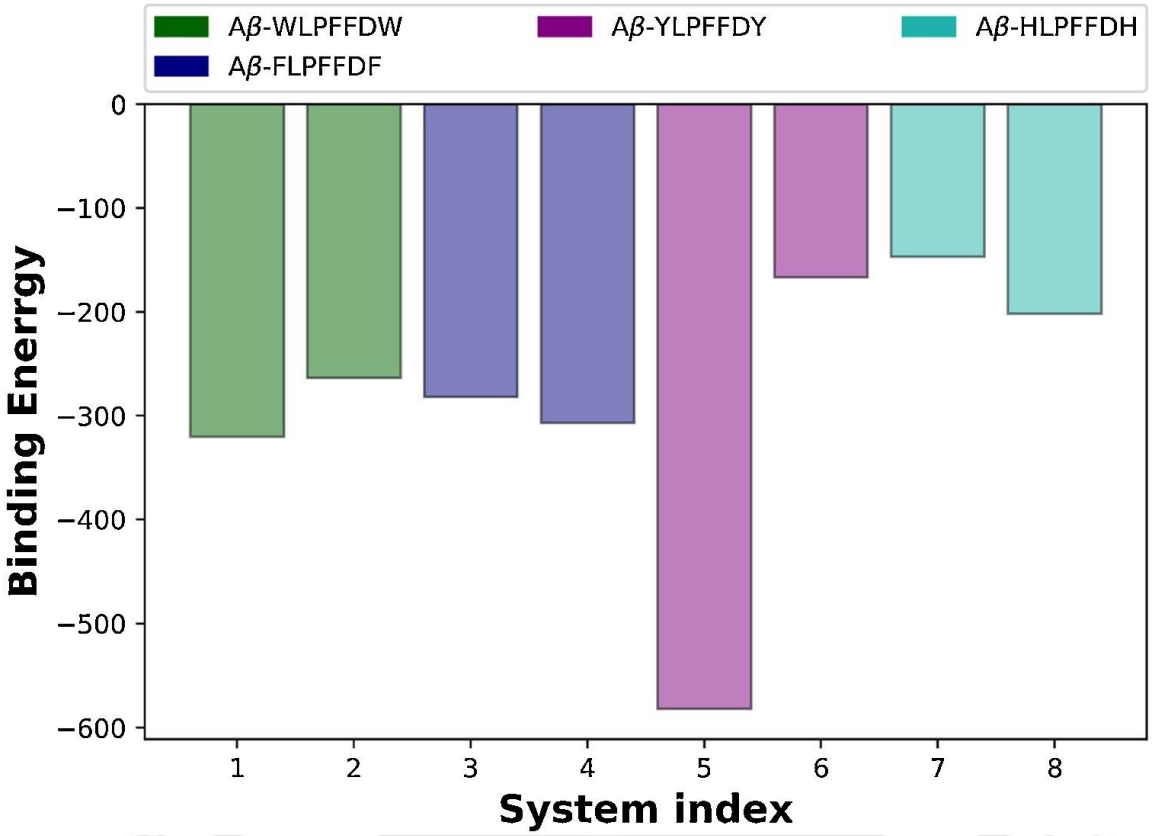


Fig. 5.12. Comparison of the binding energies of all the ligands in set 1. The Y1 ligand had the highest binding affinity.

Table 5.3. Summary of the binding energies and contact surface areas in set 1 of the simulations. Energies are expressed in kJ $mol^{-1.}$

System	van der Waal energy	Electrostatic energy	Polar solvation energy	SASA energy	Binding energy	Contact Surface Area (Å^2)
W1	-323.477 ± 1.229	358.619 ± 8.207	-322.399 ± 4.010	-32.817 ± 0.099	-320.191 ± 4.894	776.20
W2	-187.899 ± 1.999	319.598 ± 6.671	-374.071 ± 4.779	-21.118 ± 0.156	-263.439 ± 3.388	580.52
F1	-206.296 ± 1.357	236.838 ± 8.011	-289.626 ± 4.132	-23.123 ± 0.117	-281.971 ± 4.898	577.87
F2	-274.298 ± 1.139	245.310 ± 4.142	-248.908 ± 2.420	-28.933 ± 0.074	-306.812 ± 2.954	667.98
Y1	-347.791 ± 1.594	-84.113 ± 5.718	-113.849 ± 4.752	-36.196 ± 0.117	-581.834 ± 5.382	1007.81
Y2	-223.749 ± 1.719	448.733 ± 10.040	-365.580 ± 6.225	-25.291 ± 0.194	-166.797 ± 5.220	723.78
H1	-278.241 ± 1.289	439.571 ± 8.065	-276.857 ± 6.099	-30.336 ± 0.107	-146.691 ± 4.593	843.29
H2	-188.514 ± 1.172	303.094 ± 7.638	-294.003 ± 4.367	-22.247 ± 0.113	-201.652 ± 4.652	612.78

The van der Waals energetic term was the dominant term in all the systems considered by virtue of being larger than the sum of the corresponding electrostatic interaction energy term and the

polar solvation energy term. Therefore, the association of the ligands with the fibrils by the hydrophobic and aromatic interactions that we observed will depend on the buried contact surface area [92]. According to Martinez and Iverson, there is a relationship between the electrostatic complementarity of the ligands and the receptor, the buried contact surface area between them, and the favorable aromatic and hydrophobic interactions between them [92]. Except in the case of the Y2, H1 and H2 replicates, in all the replicates, a higher van der Waals energy term was associated with a higher contact surface area. In the case of the Y2, H1 and H2 systems, the comparatively lower binding affinity was associated with unfavourable electrostatic interactions.

The Y1 system had the highest van der Waals interaction energy term and buried surface area, and correspondingly the higher binding affinity. The difference in the binding modes of the ligands in the W1 and W2 systems arose due to the different secondary structures adopted by the ligands. The W1 ligand was in an extended conformation while the W2 ligand formed a helix in the course of the simulations. Hence, the buried contact surface area between this ligand and the fibrils was much lower than that of the ligand in the W1 system, which was a random coil. The W1 ligand made additional contacts with the C-terminal residues of chains A and B and was less exposed to the solvent with a higher buried contact surface area. Fig. 5.13, Fig. 5.14 show the difference in orientation of these peptides (Table 5.3).

Fig. 5.13. The binding mode of the W1 ligand. The W1 ligand is shown in blue, the amyloid residues within 4 Å of the ligand are shown in green and the rest of the fibril residues are shown in black.

Fig. 5.14. The binding mode of the W2 ligand. The W2 ligand is shown in blue, the amyloid residues within 4 Å of the ligand are shown in green and the rest of the fibril residues are shown in black. The W2 ligand has adopted a helical structure.

Table 5.4. Summary of the binding energies and contact surface areas in set 2 of the simulations. Energies are expressed in kJ mol^{-1}.

System	van der Waal energy	Electrostatic energy	Polar solvation energy	SASA energy	Binding energy	Contact Surface Area ($Å^2$)
WWW1	-453.742 ± 1.696	58.041 ± 7.263	-110.443 ± 3.255	-49.081 ± 0.120	-555.130 ± 4.304	1226.14
WWW2	-471.418 ± 1.384	397.717 ± 3.827	-297.993 ± 2.071	-42.869 ± 0.106	-414.475 ± 3.080	1116.25
FFF1	-432.136 ± 1.873	329.728 ± 6.009	-128.198 ± 3.298	-47.276 ± 0.134	-277.790 ± 3.642	1130.23
FFF2	-465.181 ± 1.973	318.827 ± 8.130	-202.371 ± 6.343	-50.437 ± 0.172	-399.914 ± 3.837	1177.52
YYY1	-323.641 ± 1.352	356.290 ± 9.653	-256.028 ± 7.894	-36.812 ± 0.138	-260.232 ± 3.891	818.82
YYY2	-225.384 ± 1.580	520.878 ± 5.474	-427.782 ± 5.179	-25.765 ± 0.203	-158.191 ± 4.444	736.84
HHH1	-297.288 ± 2.143	493.720 ± 8.017	-437.936 ± 5.696	-31.946 ± 0.179	-272.863 ± 3.869	709.65
HHH2	-199.079 ± 1.724	382.888 ± 7.880	-377.308 ± 6.247	-21.590 ± 0.154	-214.995 ± 4.537	558.10

5.3.6. Binding energies in Set 2 of the simulations

We consider the binding energies in Set 2 of the simulations in which the LPFFD peptide was modified by adding an aromatic trimer to both the termini, resulting in an eleven residue ligand. In this set, the top 3 binding affinities were of the WWW1, WWW2, and FFF2 systems. In all the systems except the WWW1 system, the electrostatic interactions were overall unfavourable for binding as the sums of the polar solvation energy term and the electrostatic energy term were positive. The van der Waals energy term was the driving force for the binding process, as this term was significantly higher than the sum of the respective electrostatic energy terms and polar solvation energy terms. Fig. 5.15 compares the binding energies of the systems in this set,

and Table 5.4 contains a summary of the binding energies and the contact surface areas in this set of the simulations.

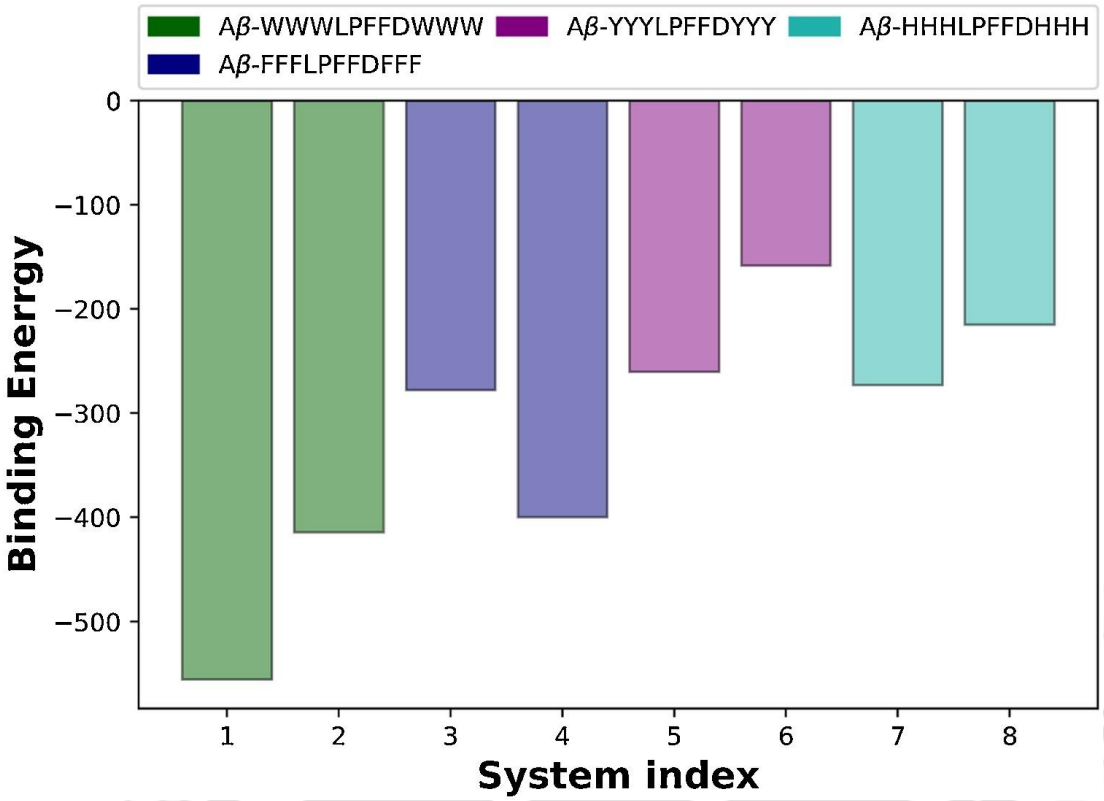


Fig. 5.15. Binding energies in set 2 of the simulations. The ligands in the WWW1 and WWW2 systems had the highest binding affinities.

The WWW1, WWW2 and FFF2 systems had high binding affinities, associated with high contact surface areas. The FFF1 ligand had a lower binding affinity than the FFF2 ligand owing to less favourable electrostatic interactions. The high van der Waals interaction energy terms in these systems corresponded to high contact surface areas, and were comparable in magnitude. The C-terminal of the YYY1 ligand was able to form beta-sheets with the fibrils in the course of the simulations. The YYY2 ligand had a low binding affinity due to unfavourable electrostatic interactions. The HHH2 ligand formed a helix for a brief period in the course of the simulations and was more exposed to the solvent than the HHH1 ligand. As a consequence, it had a low van der Waals interaction energy term and contact surface area.

The model of the fibrils chosen for this study is a disease-relevant form of the Aβ fibrils. In this model, the sidechains of the central hydrophobic core region spanning the residues 16 – 42 were buried in the solvent and inaccessible to the ligands. The binding site of the fibrils in all the

systems considered was a difficult target due to its inherent nature. As observed in experimental studies [47], it is a region which is disordered and highly flexible. It also contains several residues which are charged, aromatic and hydrophobic. The side chains of many residues in this region are exposed to the solvent. The positively charged cluster comprising the residues His 13 to Lys 16 have been reported to participate in the activation of microglia [76], which are involved in the inflammatory response to Aβ observed in Alzheimer's disease. Further, all the modified ligands made hydrophobic contacts with the hydrophobic residues in the C-terminal of the fibrils. This is an important factor to consider in the design of drugs because this region plays an important role in the aggregation of the Aβ peptides and by blocking its interactions the aggregation of the fibrils may be inhibited [87].

In the fibrils, the residues which tended to form the greatest number of contacts in all the systems were Phe 4, Arg 5, His 6, Tyr 10, His 13 and His 14. Arg 5 was able to make could form several cation-pi interactions. Aromatic interactions played an important role in the binding of the ligands to the fibrils by contributing to the van der Waals interaction energy term. Electrostatic repulsions between residues which had similar charge were observed. The LPFFD peptide has a negatively charged Asp which had unfavourable electrostatic interactions with the negatively charged charged Asp 7 and Glu 11 of the fibrils. Increasing the aromatic content of the LPFFD peptide did not always increase its binding affinity due to unfavourable electrostatic complementarity between the ligand and fibril residues and orientation of the ligands with respect to the fibrils. An increased number of sustained contacts did not necessarily translate to enhanced binding in some cases. We note however, that aromatic interactions played an important role in establishing contacts between the ligand and the fibrils. We also found that the secondary structure adopted by the ligands could influence the binding in some cases.

One of the limitations in the current study is the flexible nature of the binding region of the fibrils, which can make the convergence of the MMPBSA problematic. However, we have shown that a good binding affinity is associated with a high number of sustained contacts and a high contact surface area in most of the cases. Further, the control ligand LPFFD was not able to bind to the fibrils beyond 100 ns, but when modified with aromatic amino acids, it was able to bind well to the fibrils. This may be important in the context of this particular model of the fibrils with respect to the development of molecules, which serve as drug candidates for the treatment of Alzheimer's disease or as probes for the fibrils.

An interesting avenue for future research would be the use of more recent force fields. In the present study, the Amber 99SBILDN force field was used which is a popularly used force field in the simulation of amyloid fibrils in several recent studies [93], [94], [95], [96], [97].

Some studies [98], [99] strongly recommended the Amber 99SBILDN force-field in the study of the amyloid fibrils. The use of improved force fields may provide more accurate results. The binding site predicted by the ClusPro algorithm corresponds to the region predicted by the web-servers CASTp 3.0 [100], Bitenet [101] and Active Site Prediction [102]. It may be interesting to replace the charged Asp residue in the LPFFD peptide with an uncharged residue in order to avoid any electrostatic repulsions with the similarly charged Glu 11 residue of the fibrils.

Based on the model used in the present study, it is expected that a free tripeptide of the aromatic amino acids would bind to a similar region of the fibrils. However, due to the small length of the tripeptide, electrostatic interactions may dominate the interactions with respect to their uncapped charged termini, as observed by Viet *et al*. Tryptophan is a bulky amphipathic aromatic acid, phenylalanine is hydrophobic, tyrosine is amphipathic and histidine is polar in nature. Based on these properties, the various combinations of the amino acids would have different amounts of exposure to the solvent, which would have a consequence on the number of sustained contacts formed and buried surface area with respect to the fibrils. A longer ligand can make more contacts with fibril residues, however, this can also mean that if the ligand has a charged residue then it may have electrostatic repulsions with a like-charged residue in the fibrils. The addition of a trimeric amino acid to the termini of a ligand molecule has the advantage of being able to make contacts with the fibrils from both its termini. In the case of the monomer modifications it was seen that only the N-terminal residue of the modified ligand made contacts with the fibrils.

5.4. Conclusions

In our present study, we compared the binding of the aromatically modified LPFFD peptide to the Aβ fibrils. The modified peptides comprised of two sequences: the β-sheet breaker LPFFD and the aromatic residues at the termini which served to bind to the fibrils. The aromatic amino acids enhanced the binding affinity by forming aromatic and hydrophobic contacts. The electrostatic complementarity of the ligands and the fibrils, along with favourable orientation of these ligands played an important role in the binding. The N-terminal region of the contained many aromatic, hydrophobic and charged residues; was flexible, and exposed to the solvent. The ligands could also simultaneously bind to another distinct region: the region spanned by the hydrophobic C-terminal residues. These regions are involved in the aggregation of the fibrils and are important to consider in the design of drug molecules which can inhibit their

aggregation. The tryptophan modified LPFFD peptides had the highest binding affinity. The aromatic contacts between the ligand residues and fibril residues played an important role in initiating and establishing contacts between the ligands and the fibrils. The N-terminal aromatic ligand residues played a dominant role in forming sustained contacts with the fibril residues. Some ligands could adopt transient helical structures or form β-sheets with the fibrils, and this impacted the contact surface area and electrostatic complementarity and therefore the binding to the fibrils. The electrostatic interactions were not always favourable when the ligand tended to form a helix, while the formation of a β-sheet was always associated with good binding and favourable electrostatic interactions. An interesting avenue for further study would be the interaction of these peptides with the central hydrophobic core region of the Aβ fibrils comprising of residues in the positions 16-21 which is another important region to consider in the aggregation of the fibrils. The modification of the LPFFD peptide by the proline trimers is particularly promising, since proline is also known to be a β-sheet breaker residue. The observations based on the protofibrils are likely to extend to mature fibrils. Our study may serve as a basis for future studies of peptide-based drug design involving aromatic clusters of residues and aromatic interactions.

References

1. H.W. Querfurth, F.M. LaFerla, Alzheimer's disease., N. Engl. J. Med. 362 (2010) 329–344.
2. C.A. Ross, M.A. Poirier, Protein aggregation and neurodegenerative disease., Nat. Med. 10 Suppl (2004) S10-7.
3. S.B. Prusiner, Neurodegenerative Diseases and Prions, N. Engl. J. Med. 344 (2001) 1516–1526.
4. D.J. Selkoe, Alzheimer's disease: genes, proteins, and therapy., Physiol. Rev. 81 (2001) 741–766.
5. J. Cummings, G. Lee, A. Ritter, M. Sabbagh, K. Zhong, Alzheimer's disease drug development pipeline: 2019, Alzheimer's Dement. (New York, N. Y.). 5 (2019) 272–293.
6. B.Y.-W. Man, H.-M. Chan, C.-H. Leung, D.S.-H. Chan, L.-P. Bai, Z.-H. Jiang, H.-W. Li, D.-L. Ma, Group 9 metal-based inhibitors of β-amyloid (1–40) fibrillation as potential therapeutic agents for Alzheimer's disease, Chem. Sci. 2 (2011) 917–921.
7. B. Raman, T. Ban, K.-I. Yamaguchi, M. Sakai, T. Kawai, H. Naiki, Y. Goto, Metal ion-dependent effects of clioquinol on the fibril growth of an amyloid β peptide., J. Biol. Chem. 280 (2005) 16157–16162.
8. K.J. Barnham, V.B. Kenche, G.D. Ciccotosto, D.P. Smith, D.J. Tew, X. Liu, K. Perez, G.A. Cranston, T.J. Johanssen, I. Volitakis, A.I. Bush, C.L. Masters, A.R. White, J.P. Smith, R.A. Cherny, R. Cappai, Platinum-based inhibitors of amyloid-beta as

therapeutic agents for Alzheimer's disease., Proc. Natl. Acad. Sci. U. S. A. 105 (2008) 6813–6818.
9. A. Iscen, C.R. Brue, K.F. Roberts, J. Kim, G.C. Schatz, T.J. Meade, Inhibition of Amyloid-β Aggregation by Cobalt (III) Schiff Base Complexes: A Computational and Experimental Approach., J. Am. Chem. Soc. 141 (2019) 16685–16695.
10. L. Lannfelt, K. Blennow, H. Zetterberg, S. Batsman, D. Ames, J. Harrison, C.L. Masters, S. Targum, A.I. Bush, R. Murdoch, J. Wilson, C.W. Ritchie, Safety, efficacy, and biomarker findings of PBT2 in targeting Abeta as a modifying therapy for Alzheimer's disease: a phase IIa, double-blind, randomised, placebo-controlled trial., Lancet. Neurol. 7 (2008) 779–786.
11. Doens, M.E. Valdés-Tresanco, V. Vasquez, M.B. Carreira, Y. De La Guardia, D.E. Stephens, V.D. Nguyen, V.T. Nguyen, J. Gu, M.L. Hegde, O. V Larionov, P.A. Valiente, R. Lleonart, P.L. Fernández, Hexahydropyrrolo[2,3-b] indole Compounds as Potential Therapeutics for Alzheimer's Disease., ACS Chem. Neurosci. 10 (2019) 4250–4263.
12. S. Shi, H. Wang, J. Wang, Y. Wang, X. Xue, Z. Hou, G.-D. Yao, X.-X. Huang, H. Zhao, Q. Liu, S.-J. Song, Semi-synthesis and biological evaluation of flavone hybrids as multifunctional agents for the potential treatment of Alzheimer's disease., Bioorg. Chem. 100 (2020) 103917.
13. A. Kaur, S.S. Narang, A. Kaur, S. Mann, N. Priyadarshi, B. Goyal, N.K. Singhal, D. Goyal, Multifunctional Mono-Triazole Derivatives Inhibit Aβ42 Aggregation and Cu2+-Mediated Aβ42 Aggregation and Protect Against Aβ42-Induced Cytotoxicity, Chem. Res. Toxicol. 32 (2019) 1824–1839
14. D.E. Ehrnhoefer, J. Bieschke, A. Boeddrich, M. Herbst, L. Masino, R. Lurz, S. Engemann, A. Pastore, E.E. Wanker, EGCG redirects amyloidogenic polypeptides into unstructured, off-pathway oligomers., Nat. Struct. Mol. Biol. 15 (2008) 558–566.
15. A.R.A. Ladiwala, M. Bhattacharya, J.M. Perchiacca, P. Cao, D.P. Raleigh, A. Abedini, A.M. Schmidt, J. Varkey, R. Langen, P.M. Tessier, Rational design of potent domain antibody inhibitors of amyloid fibril assembly., Proc. Natl. Acad. Sci. U. S. A. 109 (2012) 19965–19970.
16. S. Salloway, R. Sperling, N.C. Fox, K. Blennow, W. Klunk, M. Raskind, M. Sabbagh, L.S. Honig, A.P. Porsteinsson, S. Ferris, M. Reichert, N. Ketter, B. Nejadnik, V. Guenzler, M. Miloslavsky, D. Wang, Y. Lu, J. Lull, I.C. Tudor, E. Liu, M. Grundman, E. Yuen, R. Black, H.R. Brashear, Two phase 3 trials of bapineuzumab in mild-to-moderate Alzheimer's disease., N. Engl. J. Med. 370 (2014) 322–333.
17. E. Liu, M.E. Schmidt, R. Margolin, R. Sperling, R. Koeppe, N.S. Mason, W.E. Klunk, C.A. Mathis, S. Salloway, N.C. Fox, D.L. Hill, A.S. Les, P. Collins, K.M. Gregg, J. Di, Y. Lu, I.C. Tudor, B.T. Wyman, K. Booth, S. Broome, E. Yuen, M. Grundman, H.R. Brashear, Amyloid-β 11C-PiB-PET imaging results from 2 randomized bapineuzumab phase 3 AD trials., Neurology. 85 (2015) 692–700.
18. R.S. Doody, R.G. Thomas, M. Farlow, T. Iwatsubo, B. Vellas, S. Joffe, K. Kieburtz, R. Raman, X. Sun, P.S. Aisen, E. Siemers, H. Liu-Seifert, R. Mohs, Phase 3 trials of solanezumab for mild-to-moderate Alzheimer's disease., N. Engl. J. Med. 370 (2014) 311–321.
19. F. Panza, M. Lozupone, G. Logroscino, B.P. Imbimbo, A critical appraisal of amyloid-β-targeting therapies for Alzheimer disease., Nat. Rev. Neurol. 15 (2019) 73–88.
20. D. Goyal, S. Shuaib, S. Mann, B. Goyal, Rationally Designed Peptides and Peptidomimetics as Inhibitors of Amyloid-β (Aβ) Aggregation: Potential Therapeutics of Alzheimer's Disease., ACS Comb. Sci. 19 (2017) 55–80.
21. D.J. Craik, D.P. Fairlie, S. Liras, D. Price, The future of peptide-based drugs., Chem. Biol. Drug Des. 81 (2013) 136–147.

22. L.O. Tjernberg, J. Näslund, F. Lindqvist, J. Johansson, A.R. Karlström, J. Thyberg, L. Terenius, C. Nordstedt, Arrest of beta-amyloid fibril formation by a pentapeptide ligand., J. Biol. Chem. 271 (1996) 8545–8548.
23. C. Soto, E.M. Sigurdsson, L. Morelli, R.A. Kumar, E.M. Castaño, B. Frangione, Beta-sheet breaker peptides inhibit fibrillogenesis in a rat brain model of amyloidosis: implications for Alzheimer's therapy., Nat. Med. 4 (1998) 822–826.
24. M.A. Chacón, M.I. Barría, C. Soto, N.C. Inestrosa, β-sheet breaker peptide prevents Aβ-induced spatial memory impairments with partial reduction of amyloid deposits, Mol. Psychiatry. 9 (2004) 953–961
25. C. Soto, M.S. Kindy, M. Baumann, B. Frangione, Inhibition of Alzheimer's Amyloidosis by Peptides That Prevent β-Sheet Conformation, Biochem. Biophys. Res. Commun. 226 (1996) 672–680.
26. C. Soto, L. Estrada, Amyloid inhibitors and beta-sheet breakers, Subcell. Biochem., 38 (2005), 351-364
27. Z. Datki, R. Papp, D. Zádori, K. Soós, L. Fülöp, A. Juhász, G. Laskay, C. Hetényi, E. Mihalik, M. Zarándi, B. Penke, In vitro model of neurotoxicity of Aβ 1–42 and neuroprotection by a pentapeptide: irreversible events during the first hour, Neurobiol. Dis. 17 (2004) 507–515.
28. I. Granic, M.F. Masman, C. Kees Mulder, I.M. Nijholt, P.J.W. Naude, A. de Haan, E. Borbély, B. Penke, P.G.M. Luiten, U.L.M. Eisel, LPYFDa neutralizes amyloid-beta-induced memory impairment and toxicity, J. Alzheimers Dis., 19 (2010), 991-1005
29. V. Minicozzi, R. Chiaraluce, V. Consalvi, C. Giordano, C. Narcisi, P. Punzi, G.C. Rossi, S. Morante, Computational and experimental studies on β-sheet breakers targeting Aβ1-40 fibrils
1. J. Biol. Chem., 289 (2014), 11242-11252
30. J.A. Loureiro, R. Crespo, H. Börner, P.M. Martins, F.A. Rocha, M. Coelho, M.C. Pereira, S. Rocha, Fluorinated beta-sheet breaker peptides, J. Mater. Chem. B, 2 (2014), 2259-2264
31. M.A. Findeis, G.M. Musso, C.C. Arico-Muendel, H.W. Benjamin, A.M. Hundal, J.-J. Lee, J. Chin, M. Kelley, J. Wakefield, N.J. Hayward, S.M. Molineaux, Modified-peptide inhibitors of amyloid β-Peptide polymerization, Biochemistry, 38 (1999), 6791-6800
32. J.F. Poduslo, G.L. Curran, A. Kumar, B. Frangione, C. Soto, β-sheet breaker peptide inhibitor of Alzheimer's amyloidogenesis with increased blood–brain barrier permeability and resistance to proteolytic degradation in plasma, J. Neurobiol., 39 (1999) 371-382.
33. P. De Bona, M. Laura Giuffrida, F. Caraci, A. Copani, B. Pignataro, F. Attanasio, S. Cataldo, G. Pappalardo, E. Rizzarelli, Design and synthesis of new trehalose-conjugated pentapeptides as inhibitors of Aβ(1–42) fibrillogenesis and toxicity, J. Pept. Sci. 15 (2009) 220–228
34. M.H. Viet, K. Siposova, Z. Bednarikova, A. Antosova, T.T. Nguyen, Z. Gazova, M.S. Li,
2. In Silico and in vitro study of binding affinity of Tripeptides to amyloid β fibrils: implications for Alzheimer's disease, J. Phys. Chem. B, 119 (2015) 5145-5155
35. E. Gazit, A possible role for pi-stacking in the self-assembly of amyloid fibrils, FASEB J., 16 (2002), 77-83
36. E. Gazit, Mechanistic Studies of the Process of Amyloid Fibrils Formation by the Use of Peptide Fragments and Analogues: Implications for the Design of Fibrillization Inhibitors, Curr. Med. Chem. 9 (2002) 1725–1735

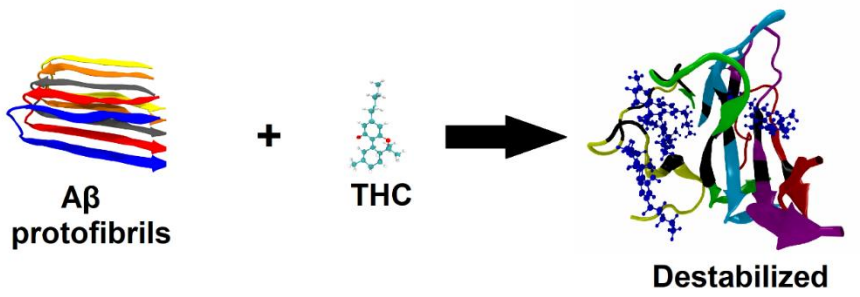
37. G.G. Tartaglia, A. Cavalli, R. Pellarin, A. Caflisch, The role of aromaticity, exposed surface, and dipole moment in determining protein aggregation rates, Protein Sci., 13 (2004) 1939-1941
38. A. Aggeli, M. Bell, N. Boden, J.N. Keen, P.F. Knowles, T.C.B. McLeish, M. Pitkeathly, S.E. Radford, Responsive gels formed by the spontaneous self-assembly of peptides into polymeric β-sheet tapes, Nature, 386 (1997) 259-262
39. S.K. Burley, G.A. Petsko, Aromatic-aromatic interaction: a mechanism of protein structure stabilization, Science (80) 229 (1985) 23 LP – 28
40. C.G. Claessens, J.F. Stoddart, π–π INTERACTIONS IN SELF-ASSEMBLY, J. Phys. Org. Chem. 10 (1997) 254–272
41. W.G. Turnell, J.T. Finch, Binding of the dye congo red to the amyloid protein pig insulin reveals a novel homology amongst amyloid-forming peptide sequences, J. Mol. Biol., 227 (1992). 1205-1223
42. Y. Porat, Y. Mazor, S. Efrat, E. Gazit, Inhibition of islet amyloid polypeptide fibril formation: a potential role for heteroaromatic interactions, Biochemistry, 43 (2004) 14454-14462
43. S. Taniguchi, N. Suzuki, M. Masuda, S. Hisanaga, T. Iwatsubo, M. Goedert, M. Hasegawa
3. Inhibition of heparin-induced tau filament formation by phenothiazines, polyphenols, and porphyrins, J. Biol. Chem., 280 (2005) 7614-7623
44. L. Xie, Y. Luo, G. Wei, Aβ(16–22) Peptides Can Assemble into Ordered β-Barrels and Bilayer β-Sheets, while Substitution of Phenylalanine 19 by Tryptophan Increases the Population of Disordered Aggregates, J. Phys. Chem. B. 117 (2013) 10149–10160
45. M. Tanaka, H. Kiyohara, A. Yoshino, A. Nakano, F. Takata, S. Dohgu, Y. Kataoka, T. Matsui, Brain-transportable soy dipeptide, Tyr-Pro, attenuates amyloid β peptide25-35-induced memory impairment in mice, Npj Sci. Food, 4 (2020) 7
46. E.F. Marondedze, K.K. Govender, P.P. Govender, Computational investigation of the binding characteristics of β-amyloid fibrils, Biophys. Chem., 256 (2020), 106281
47. M.A. Wälti, F. Ravotti, H. Arai, C.G. Glabe, J.S. Wall, A. Böckmann, P. Güntert, B.H. Meier, R. Riek, Atomic-resolution structure of a disease-relevant Aβ(1–42) amyloid fibril, Proc. Natl. Acad. Sci. 113 (2016) E4976 LP-E4984
48. A. Acharya, J. Stockmann, L. Beyer, T. Rudack, A. Nabers, J.C. Gumbart, K. Gerwert, V.S. Batista, The Effect of Epigallocatechin-3-Gallate on the Amyloid Secondary Structure, Biophys. J. 119 (2020) 349–359.
49. H.L. Nguyen, P. Krupa, N.M. Hai, H.Q. Linh, M.S. Li, Structure and physicochemical properties of the Aβ42 tetramer: multiscale molecular dynamics simulations, J. Phys. Chem. B, 123 (2019) 7253-7269
50. A.B. Poma, H.V. Guzman, M.S. Li, P.E. Theodorakis, Mechanical and thermodynamic properties of Aβ(42), Aβ(40), and α-synuclein fibrils: a coarse-grained method to complement experimental studies, Beilstein J. Nanotechnol., 10 (2019), 500-513
51. M. Kouza, N.T. Co, M.S. Li, S. Kmiecik, A. Kolinski, A. Kloczkowski, I.A. Buhimschi, Kinetics and mechanical stability of the fibril state control fibril formation time of polypeptide chains: a computational study, J. Chem. Phys., 148 (2018) 215106
52. A. Balupuri, K.-E. Choi, N.S. Kang, Aggregation Mechanism of Alzheimer's Amyloid β-Peptide Mediated by α-Strand/α-Sheet Structure, Int. J. Mol. Sci. 21 (2020)
53. E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera--a visualization system for exploratory research and analysis, J. Comput. Chem., 25 (2004) 1605-1612
54. D. Kozakov, D.R. Hall, B. Xia, K.A. Porter, D. Padhorny, C. Yueh, D. Beglov, S. Vajda, The ClusPro web server for protein-protein docking, Nat. Protoc., 12 (2017) 255-278

55. D. Kozakov, D. Beglov, T. Bohnuud, S.E. Mottarella, B. Xia, D.R. Hall, S. Vajda, How good is automated protein docking? Proteins, 81 (2013) 2159-2166
56. D. Kozakov, R. Brenke, S.R. Comeau, S. Vajda, PIPER: an FFT-based protein docking program with pairwise potentials, Proteins, 65 (2006) 392-406
57. Stephen R. Comeau, D.W. Gatchell, S. Vajda, C.J. Camacho, ClusPro: an automated docking and discrimination method for the prediction of protein complexes, Bioinformatics, 20 (2004) 45-50
58. Stephen R. Comeau, D.W. Gatchell, S. Vajda, C.J. Camacho, ClusPro: a fully automated algorithm for protein-protein docking, Nucleic Acids Res., 32 (2004) W96-W99
59. S. Nosé, A unified formulation of the constant temperature molecular dynamics methods, J. Chem. Phys. 81 (1984) 511–519.
60. W.G. Hoover, Canonical dynamics: Equilibrium phase-space distributions., Phys. Rev. A, Gen. Phys. 31 (1985) 1695–1697.
61. M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method, J. Appl. Phys. 52 (1981) 7182–7190.
62. S. Nosé, M.L. Klein, Constant pressure molecular dynamics for molecular systems, Mol. Phys. 50 (1983) 1055–1076.
63. M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, E. Lindahl, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX. 1–2 (2015) 19–25.
64. B. Hess, P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation., J. Chem. Theory Comput. 4 (2008) 116–122.
65. U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen, A smooth particle mesh Ewald method, J. Chem. Phys. 103 (1995) 8577–8593.
66. W.G. Touw, C. Baakman, J. Black, T.A.H. te Beek, E. Krieger, R.P. Joosten, G. Vriend, A series of PDB-related databanks for everyday needs., Nucleic Acids Res. 43 (2015) D364-8.
67. W. Kabsch, C. Sander, Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features., Biopolymers. 22 (1983) 2577–2637.
68. D. van der Spoel, P.J. van Maaren, P. Larsson, N. Tîmneanu, Thermodynamics of hydrogen bonding in hydrophilic and hydrophobic media, J. Phys. Chem. B, 110 (2006) 4393-4398
69. W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, J. Mol. Graph., 14 (1996) 33-38
70. S.J. Wodak, J. Janin, Analytical approximation to the accessible surface area of proteins
4. Proc. Natl. Acad. Sci. U. S. A., 77 (1980) 1736-1740
71. J. Srinivasan, J. Miller, P.A. Kollman, D.A. Case, Continuum solvent studies of the stability of RNA hairpin loops and helices., J. Biomol. Struct. Dyn. 16 (1998) 671–682.
72. P.A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D.A. Case, T.E. 3rd Cheatham, Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models., Acc. Chem. Res. 33 (2000) 889–897
73. R. Kumari, R. Kumar, A. Lynn, g_mmpbsa--a GROMACS tool for high-throughput MM-PBSA calculations, J. Chem. Inf. Model. 54 (2014) 1951–1962.
74. N.A. Baker, D. Sept, S. Joseph, M.J. Holst, J.A. McCammon, Electrostatics of nanosystems: application to microtubules and the ribosome., Proc. Natl. Acad. Sci. U. S. A. 98 (2001) 10037–10041.
75. X. Daura, K. Gademann, B. Jaun, D. Seebach, W.F. van Gunsteren, A.E. Mark, Peptide Folding: When Simulation Meets Experiment, Angew. Chemie Int. Ed. 38 (1999) 236–240.

76. W. Garzon-Rodriguez, A. Vega, M. Sepulveda-Becerra, S. Milton, D.A. Johnson, A.K. Yatsimirsky, C.G. Glabe, A Conformation Change in the Carboxyl Terminus of Alzheimer's Amyloid-beta Accompanies the Transition from Dimer to Fibril as Revealed by Fluorescence Quenching Analysis, J. Biol. Chem. 275 (2000) 22645–22649
77. J.M. Jakubowski, A.A. Orr, D.A. Le, P. Tamamis, Interactions between curcumin derivatives and Amyloid-β fibrils: insights from molecular dynamics simulations, J. Chem. Inf. Model., 60 (2020), 289-305
78. D. Goyal, S. Shuaib, S. Mann, B. Goyal, Rationally designed peptides and Peptidomimetics as inhibitors of Amyloid-β (Aβ) aggregation: potential therapeutics of Alzheimer's disease, ACS Comb. Sci., 19 (2017) 55-80
79. K. Przygońska, J. Poznański, U.H. Mistarz, K.D. Rand, M. Dadlez, Side-chain moieties from the N-terminal region of Aβ are involved in an oligomer-stabilizing network of interactions, PLoS One, 13 (2018), 0201761
80. B. Murray, B. Sharma, G. Belfort, N-terminal hypothesis for Alzheimer's disease
5. ACS Chem. Neurosci., 8 (2017), 432-434
81. O. Coskuner, V.N. Uversky, Tyrosine Regulates β-Sheet Structure Formation in Amyloid-β42: A New Clustering Algorithm for Disordered Proteins, J. Chem. Inf. Model. 57 (2017) 1342–1358
82. D.G. Smith, G.D. Ciccotosto, D.J. Tew, K. Perez, C.C. Curtain, J.F. Boas, C.L. Masters, R. Cappai, K.J. Barnham, Histidine 14 modulates membrane binding and neurotoxicity of the Alzheimer's disease amyloid-beta peptide, J. Alzheimers Dis., 19 (2010), 1387-1400
83. K. Przygońska, M. Pacewicz, W. Sadowska, J. Poznański, W. Bal, M. Dadlez, His6, His13, and His14 residues in Aβ 1-40 peptide significantly and specifically affect oligomeric equilibria, Sci. Rep., 9 (2019) 9449
84. A.N. Istrate, S.A. Kozin, S.S. Zhokhov, A.B. Mantsyzov, O.I. Kechko, A. Pastore, A.A. Makarov, V.I. Polshakov, Interplay of histidine residues of the Alzheimer's disease Aβ peptide governs its Zn-induced oligomerization, Sci. Rep. 6 (2016) 21734
85. M.H. Viet, S.T. Ngo, N.S. Lam, M.S. Li, Inhibition of Aggregation of Amyloid Peptides by Beta-Sheet Breaker Peptides and Their Binding Affinity, J. Phys. Chem. B. 115 (2011) 7433–7446
86. F. Liu, W. Du, Y. Sun, J. Zheng, X. Dong, Atomistic characterization of binding modes and affinity of peptide inhibitors to amyloid-β protein, Front. Chem. Sci. Eng., 8 (2014) 433-444
87. R. Tycko, Alzheimer's disease: structure of aggregates revealed, Nature, 537 (2016) 492-493
88. S. Ghosh, T. Bierig, S. Lee, S. Jana, A. Löhle, G. Schnapp, C.S. Tautermann, N. Vaidehi, Engineering Salt Bridge Networks between Transmembrane Helices Confers Thermostability in G-Protein-Coupled Receptors, J. Chem. Theory Comput. 14 (2018) 6574–6585
89. C.A. Hunter, A.J. Kirby, D.H. Williams, Aromatic interactions in proteins, DNA and synthetic receptors, Philos. Trans. R. Soc. London. Ser. A Phys. Eng. Sci. 345 (1993) 77–85.
90. E. Gazit, Self Assembly of Short Aromatic Peptides into Amyloid Fibrils and Related Nanostructures, Prion. 1 (2007) 32–35.
91. M. Nocker, S. Handschuh, C. Tautermann, K.R. Liedl, Theoretical prediction of hydrogen bond strength for use in molecular modelling, J. Chem. Inf. Model., 49 (2009) 2067-2076

92. C.R. Martinez, B.L. Iverson, Rethinking the term “pi-stacking.”, Chem. Sci., 3 (2012) 2191-2201
93. Y. Gong, C. Zhan, Y. Zou, Z. Qian, G. Wei, Q. Zhang, Serotonin and Melatonin Show Different Modes of Action on Aβ42 Protofibril Destabilization, ACS Chem. Neurosci. 12 (2021) 799–809
94. S. Banerjee, M. Hashemi, K. Zagorski, Y.L. Lyubchenko, Cholesterol in membranes facilitates aggregation of amyloid β protein at physiologically relevant concentrations, ACS Chem. Neurosci., 12 (2021) 506-516
95. L. Dorosh, M. Wu, M. Stepanova, Interaction of Aβ1−42 chains and fibrillary seeds studied by all-atom molecular dynamics simulations, Comput. Math. Methods. (2020) e1138
96. J. Wan, Y. Gong, Z. Xu, X. Dong, G. Wei, Q. Zhang, Molecular dynamics simulations reveal the destabilization mechanism of Alzheimer’s disease-related tau R3-R4 Protofilament by norepinephrine., Biophys. Chem. 271 (2021) 106541
97. J. González-Sanmiguel, C.F. Burgos, D. Bascuñán, E.J. Fernández-Pérez, N. Riffo-Lepe, S. Boopathi, A. Fernández-Pérez, C. Bobadilla-Azócar, W. González, M. Figueroa, B. Vicente, L.G. Aguayo, Gabapentin Inhibits Multiple Steps in the Amyloid Beta Toxicity Cascade., ACS Chem. Neurosci. 11 (2020) 3064–3076.
98. V.H. Man, X. He, P. Derreumaux, B. Ji, X.-Q. Xie, P.H. Nguyen, J. Wang, Effects of all-atom molecular mechanics force fields on amyloid peptide assembly: the case of Aβ16–22 dimer, J. Chem. Theory Comput., 15 (2019) 1440-1452
99. A.K. Somavarapu, K.P. Kepp, The dependence of Amyloid-β dynamics on protein force fields and water models, ChemPhysChem, 16 (2015) 3278-3289
100. W. Tian, C. Chen, X. Lei, J. Zhao, J. Liang, CASTp 3.0: computed atlas of surface topography of proteins, Nucleic Acids Res., 46 (2018) W363-W367
101. I. Kozlovskii, P. Popov, Spatiotemporal identification of druggable binding sites using deep learning, Commun. Biol., 3 (2020) 618
102. T. Singh, D. Biswas, B. Jayaram, AADS--an automated active site identification, docking, and scoring protocol for protein targets based on physicochemical descriptors, J. Chem. Inf. Model., 51 (2011) 2515-2527

Chapter 6. Destabilization of the Alzheimer's Amyloid-β Protofibrils by THC: A Molecular Dynamics Simulation study



6.1. INTRODUCTION

Alzheimer's disease is a dangerous neurodegenerative disease which causes dementia in the elderly population, leading to memory loss and death.[1,2] The hallmarks of this disease are the presence of extracellular aggregates of amyloid fibrils and neurofibrillary tangles in the form of plaques which are found in the brains of the patients[2,3]. The plaques are composed of the amyloid fibrils which comprise of the self-assembled amyloid peptides which are produced by the cleavage of the amyloid precursor protein by β- and γ-secretases.[4,5,6] The neuritic plaques contain extracellular deposits of the amyloid-β peptide which is both in the

fibrillar form and the non-fibrillar form, and is surrounded by degenerating axons and dendrites.[7] The 42 residue fibrillar form of this peptide $A\beta_{42}$ is neurotoxic and is found in higher concentrations in the plaques of the brains associated with Alzheimer's disease than the 40 residues long $A\beta_{40}$ variant.[7]

Molecules, which interact directly with these amyloid-beta peptides have attracted attention as potential therapeutic drugs for Alzheimer's disease.[2] Despite several efforts, there has been no drug molecule, which has been approved since 2003.[8] Several molecules have been studied for their ability to inhibit the aggregation of the amyloid-beta peptides and to destabilize them. As of February 2019, some of the various compounds which are in Phase 3 of clinical trials are Gantenerumab, CAD106, Solanezumab, Aducanumab, Crenezumab, AXS-05, Mirtazapine, Methylphenidate, Octohydroaminoacridine succinate, CNPS 20, Masitinib, and Zolpidem, amongst several o

thers.[8] The mechanism of action of these drugs may be broadly classified as disease-modifying therapies, treatment of psychiatric symptoms and symptomatic cognitive enhancers.[8] A comprehensive review of the various drugs in different stages of clinical trials is provided by Cummins *et al*.[8] Apart from these, historically there were many studies done to investigate the therapeutic potential of a variety of molecules both from natural sources and synthesized: Congo red, myricetin, melatonin, anthraquinone, sugar trehalose, Thioflavin T, N-methylated peptides, polyphenol epigallocatechin-3-gallate (EGCG), ibuprofen, naproxen, morin[9], wine-related polyphenolic compounds[10], piceid [11], fullerene [12], a fullerene derivative [13] , the peptide ffvlk [14], wgx-50 [15], humanin.[16] Several peptide-based drugs, some of which are based on the sequence of the amyloid beta peptide, have also been studied, and a comprehensive review of these is provided by Goyal *et al*.[17]

Cannabinoids are known to exert a neuroprotective influence and this property has attracted interest in their possible therapeutic potential to treat neurodegenerative diseases.[18] An

alteration in the cannabinoid system in the brains of the patients of Alzheimer's disease was reported, and this is an indication that the cannabinoid system may be linked to the disease. [18] Cannabinoids can protect neurons from the harmful effects of the amyloid-β peptide, can reduce tau phosphorylation, oxidative stress in the brain.[18] The two major cannabinoids and those best known for their therapeutic potentials are CBD (Cannabidiol) and THC (Δ-tetrahydrocannabidiol).[19] Several in vitro studies investigated the potential of the cannabinoids CBD and THC in treating and preventing Alzheimer's disease.[20] THC has anti-inflammatory, analgesic, appetite-stimulant, and antiemetic properties.[21] CBD has antipsychotic, neuroprotective, anticancer, antidiabetic properties, and other positive effects, such as the ability to reduce tobacco addiction.[21],[22],[23] THC can inhibit the aggregation of the amyloid-β peptide, as well as inhibit the enzyme acetylcholinesterase.[24] THC can therefore directly impact Alzheimer's disease pathology.[24] In a study by Eubanks *et al.* it was shown that THC binds in the peripheral anionic site of AChE, the critical region involved in amyloidgenesis.[24] It was shown that THC was a considerably more effective inhibitor of AChE-induced Aβ deposition than the approved drugs for Alzheimer's disease treatment donepezil and tacrine.[24] A study by Cao *et al.* showed that THC directly interacts with Aβ peptide, and inhibits aggregation.[25] It was also able to enhance mitochondria function.[25] Using a proteotoxicity model based on the inducible expression of Aβ in a human central nervous system nerve cell line, Currais *et al.* showed that THC reduced the levels of the amyloid-β and eliminated the inflammatory response from the nerve cells.[26] Zimmer *et al.* showed that a low dose of THC reversed the age-related decline in cognitive performance of mice aged 12 and 18 months.[27] A significant decrease in soluble amyloid β peptide levels and a change in plaques composition were observed in AβPP/PS1 transgenic mice when they were treated with THC and CBD.[28]

The molecular mechanism of the interactions between the THC molecules and Aβ fibrils has not yet been studied. It is also known that the THC molecule can directly interact with the fibrils.[25] In our present study, by employing all atom molecular dynamics simulations, we studied the interaction between the Aβ protofibrils and the THC molecules. We characterized the binding sites where the THC molecules interacted with the protofibrils. The THC molecules were able to form hydrophobic contacts with the residues in the C-terminal region of the protofibrils. The binding was driven by strong hydrophobic interactions between the THC molecules and the protofibrils. These strong hydrophobic interactions between the THC molecules and the protofibrils weakened the native hydrophobic packing in the protofibrils, leading to its destabilization. Hydrophobic interactions which stabilize salt bridges were disrupted, leading to the loss of important native salt-bridges between the protofibril chains. Interchain hydrogen bonds were disrupted, leading to the loss of the β-sheet structure of the protofibrils. There was an increase in the coil content of the protofibrils. We identified regions of the protofibrils where the disruption in conformation was the greatest. The tight interatomic packing of the protofibrils was made loose by increased exposure to the solvent. Our study provides the mechanistic insights into the destabilization of the protofibrils by THC molecules.

6.2. METHODOLOGY

A pentapeptide segment of the Aβ protofilament which was determined by solid-state NMR was chosen to represent the Aβ fibrils.[29] This was obtained from the Protein Data Bank (PDB ID: 2BEG).[29] This structure contains the residues 17 to 42 of the Aβ sequence, and the first 16 N-terminal residues are missing. This structure is referred to as a protofibril, as it is a small structure and a repeat unit of the mature fibrils.[30] The residues are numbered according to their sequence in the full length Aβ peptide for convenience. The $A\beta_{17-42}$ pentapeptide

structure has 5 chains which are labeled A, B, C, D, and E for convenience as shown in Figure 6.1A. The residues in the sequence 17-42 of the Aβ peptide form its core region and contribute the most to the stability of the mature fibrils.[30] [31] The first 16 residues are disordered, and contribute little to the overall stability of the fibrils. Although the first 16 residues are missing, observations using this model are likely to extend to the full-length fibrils.[30] Pentameric oligomers are potent on-pathway intermediates that eventually form U-shaped fibrillar structures.[32][33][34] This model has been used to model the Aβ fibrils in several studies. [35][36][37][38][39][40][41][42][43] This structure consists of two β strands which are connected by a bend region. The residues 18-26 form the first β strand, labeled as β-1 for convenience while the residues 31-42 form the second β strand, labeled as β-2. The bend region is comprised of the residues 27-30. These regions are shown in Figure 6.1A.

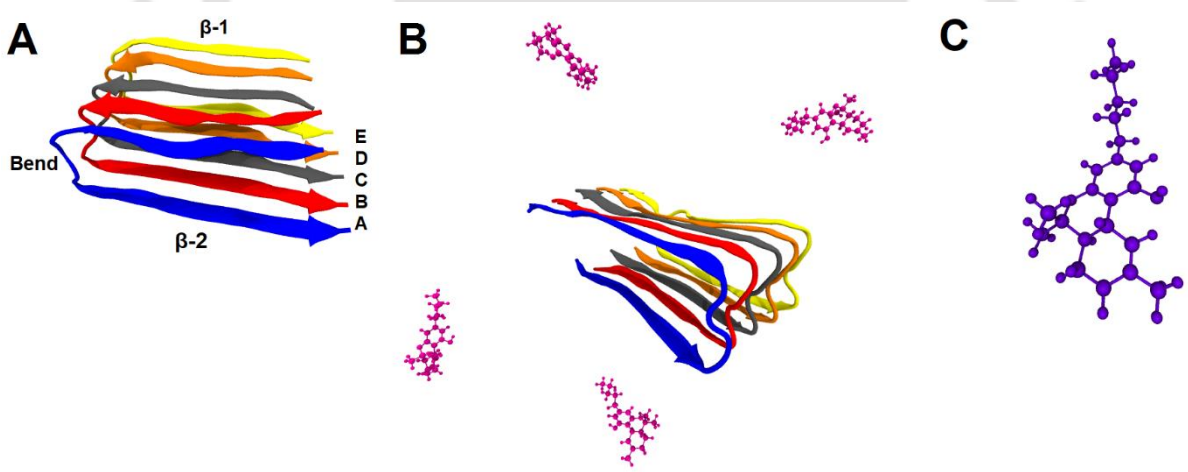


Figure 6.1. The model of the $A\beta_{17-42}$ protofibrils used in this study. This model contains five chains of the $A\beta_{17-42}$ peptide which are labeled A-E for convenience. The protofibrils consist of three regions: an N-terminal β-strand labeled β-1, a bend region and a C-terminal β-strand labeled as β-2. **B.** Initial configuration of the system of four THC molecules placed randomly around the Aβ protofibrils. The THC molecules are shown in the CPK representation and the Aβ protofibrils are shown in the cartoon representation. **C.** Structure of the THC molecule.

The structure of the THC molecule was obtained from the ZINC database [44]. The structural parameters for the THC molecule were obtained from the Swiss Param server.[45] The topology of the THC molecule is provided in Table 6.1. Simulations employed the

CHARMM36m forcefield [46]. In the initial starting structure, four THC molecules were placed randomly around the protofibrils. The initial configuration is shown in Figure 6.1B. The THC molecules and the protofibrils were sufficiently far apart from each other such that there were no interactions between them in the starting structure. Three replica runs were performed for this configuration of the THC molecules and the Aβ protofibrils. Three control simulations of the protofibrils in the absence of the THC molecules were also performed in order to compare the effect of the THC molecules on the structure of the protofibrils.

Each system was placed in a cubic box of TIP3P water. Details about the system size and the number of water molecules are provided in Table 6.2. Five sodium ions were added in order to neutralize the charge of the systems. An energy minimization run was performed using the steepest descents algorithm following which there were two runs of equilibration in which position restraints were applied to the heavy atoms. First a 200ps NVT equilibration was carried out at 300K by coupling the system to a Nose-Hoover temperature bath [47], [48]. The next stage of equilibration was an NPT run employing the Parinello-Rahman barostat [49] [50]. A molecular dynamics production run was done for 900 ns using a timestep of 2fs. The simulations were done using GROMACS 5.1.4. [51] Periodic boundary conditions were applied in all the directions. The constraint algorithm used for bond lengths was P-LINCS. [52] The neighbor search cut-off was at approximately 1.0 nm. The fast smooth particle mesh Ewald summation method was used to calculate the long-range electrostatic interactions and the Fourier grid spacing was 0.16 nm. [53] The minimum distance of a protein system to its periodic image was measured to check for artifacts. The secondary structure was calculated using the do_dssp [54], [55] tool in GROMACS. Hydrogen bonds were calculated using the *gmx hbond* [56] tool in GROMACS. Salt bridges were calculated using VMD [57]. Sustained contacts were measured using the VMD tcl script contactFreq.tcl. Sustained contacts are defined as those contacts which lasted for 40% of the time interval considered, using the method of Vaidehi *et al*. [58] Graphs

were plotted using Python 3.7.3. Solvent accessible surface areas (SASA) were calculated using the GROMACS tool sasa. [59] The root mean squared deviation (RMSD), the radius of gyration, and distances between residues were calculated using the gromacs tools *gmx rms, gmx gyrate* and *gmx mindist* respectively.

Table 6.1. System Information

System	**Box size (nm)**	**Number of Water molecules**	**Total number of atoms**	**Number of Na^+ counter ions**	**Duration of simulation (ns)**
Controls	6.78853	10115	32220	5	900
Runs 1, 2, and 3	9.25196	26298	80981	5	900

The MM/PBSA method was used to obtain the binding free energies of the THC molecules with the amyloid protofibrils. In the MM/PBSA method, the binding energy is given by the following expression: [60][61]

$$\Delta G^{bind} = \Delta EMM + \Delta G^{psolv} + \Delta G^{npsolv} - T\Delta S \quad \text{(Equation 6.1)}$$

In this expression, ΔEMM is the molecular mechanics contribution to the binding free energy in vacuum, ΔG^{psolv} is the polar contribution to the solvation energy and ΔG^{npsolv} is the non-polar contribution to the solvation energy. The polar solvation energy term was calculated by solving the Poisson-Boltzmann equation and the non-polar solvation energy term was calculated by using the solvent-accessible surface area (SASA) model. The entropic terms were excluded in the calculations in order to reduce the computational cost and the statistical uncertainty involved. The g_mmpbsa tool [62][63] was used to calculate the binding free energies using the single trajectory protocol. Snapshots from the last 20ns of the production run extracted every 100ps were used for the analysis. The energetic components are expressed as an average with standard deviations. The solute and solvent dielectric constants were taken as

4 and 80 respectively. The non-linear Poisson-Boltzmann equation was solved to calculate the polar solvation energy.

6.3. RESULTS AND DISCUSSION

6.3.1 Binding Modes of THC Molecules

The duration of the simulations was sufficient for the THC molecules to deposit on the surface of the protofibrils and maintain stable interactions with the protofibrils such that their positions did not fluctuate much in the last 200 ns of the simulations. Thus the binding of the THC molecules to the protofibrils could be observed in the time frame of the simulations. Previous experimental studies have shown that the THC molecule is highly insoluble in water.[24] The THC molecules were able to bind to the hydrophobic regions of the protofibrils: this can be explained by the insolubility of the THC molecules in water and the presence of a hydrophobic surface of the protofibrils formed by hydrophobic residues in the $A\beta_{17-42}$ sequence. Figure 6.2 shows the residues in the final snapshot with which these THC molecules were able to make sustained contacts with in the last 200 ns of the simulations.

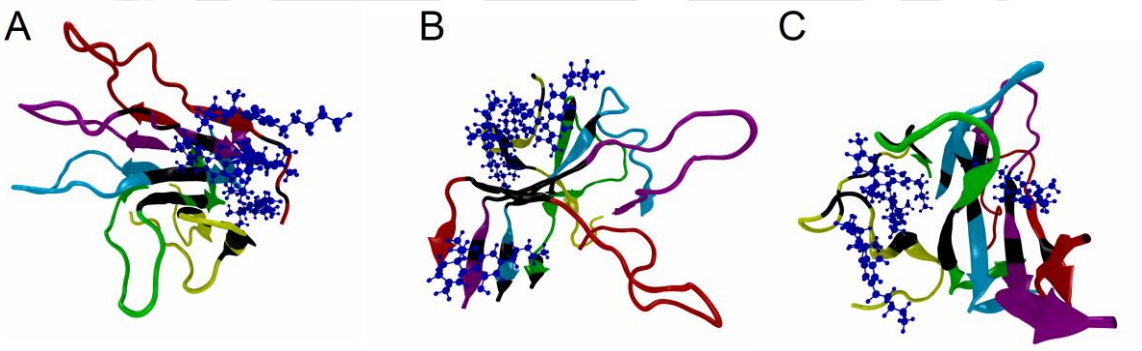


Figure 6.2. Final snapshot of the THC- Aβ protofibrils systems in : **A.** run 1, **B.** run 2, and **C.** run 3. The regions in the protofibrils which are highlighted in black are the regions with which the THC molecules made sustained contacts with in the last 200ns. The THC molecules are shown in the CPK representation.

For convenience, we refer to the four THC ligands as ligand-1, ligand-2, ligand-3, and ligand-4. In the last 200 ns, the THC ligands made several sustained contacts with the hydrophobic residues in the β-2 regions of the protofibrils and a few contacts with the residues in the bend regions. These are shown in Figure 6.2. The MMPBSA calculations confirmed the role of the hydrophobic interactions. The dominant term in the binding energy equation was the van der Waals interactions energy term on account of these hydrophobic interactions. The MMPBSA calculations are summarized in Table 6.2.

Table 6.2. MMPBSA calculations. All terms are in kJ mol^{-1}.

Run	**Ligand**	**van der Waals energy**	**Electrostatic energy**	**Polar Solvation energy**	**SASA Energy**	**Binding Energy**
Run 1	Ligand - 1	-90.8 ± 1.5	-5.2 ± 0.6	40.0 ± 0.9	-11.9 ± 0.1	-67.8 ± 1.5
	Ligand - 2	-103.7 ± 0.80	-13.4 ± 1.0	76.6 ± 1.1	-15.3 ± 0.1	-55.9 ± 1.2
	Ligand - 3	-87.9 ± 1.1	-9.2 ± 0.5	41.5 ± 2.0	-12.0 ± 0.1	-67.7 ± 2.0
	Ligand - 4	-93.1 ± 1.3	-11.1 ± 0.7	30.4 ± 1.8	-10.5 ± 0.1	-84.2 ± 1.9
Run 2	Ligand - 1	-90.8 ± 1.5	-5.2 ± 0.5	40.1 ± 0.9	-11.9 ± 0.1	-67.8 ± 1.5
	Ligand - 2	-103.8 ± 0.8	-13.2 ± 1.0	76.7 ± 1.2	-15.3 ± 0.1	-55.9 ± 1.2
	Ligand - 3	-87.9 ± 1.2	-9.2 ± 0.5	41.5 ± 1.9	-12.0 ± 0.1	-67.6 ± 1.9
	Ligand - 4	-93.1 ± 1.2	-11.0 ± 0.7	30.5 ± 1.8	-10.5 ± 0.1	-84.1 ± 1.9
Run 3	Ligand - 1	-67.4 +/- 1.0	-2.8 +/- 0.6	23.9 +/- 0.7	-9.8 +/- 0.1	-56.2 +/- 0.9
	Ligand - 2	-108.1 +/- 1.1	-64.5 +/- 2.5	96.6 +/- 1.5	-15.6 +/- 0.1	-91.6 +/- 1.8
	Ligand - 3	-128.8 +/- 0.8	-16.8 +/- 1.7	78.8 +/- 1.5	-17.1 +/- 0.1	-83.9 +/- 1.5
	Ligand - 4	-138.6 +/- 0.8	-66.8 +/- 0.6	66.6 +/- 1.0	-16.7 +/- 0.1	-155.6 +/- 1.0

In run 1, the THC ligands bound to the hydrophobic surface of the protofibrils formed by the residues in the region Ile 32 - Gly 38, and did not insert into the hydrophobic core of the protofibrils. Similarly, in run 2, the THC ligands bound to the surface of the protofibrils. The ligands 1, 3 and 4 attached to the hydrophobic surface formed by the β-2 residues. Ligand – 2 made contacts with the C-terminal residues Val 40 and Ala 42 of chains B, C, and D. Due to the effect of the destabilization of the protofibrils, the N-terminal residues Leu 17E and Phe 19E made contacts with ligand - 3 as a consequence of the formation of a β-hairpin by the residues in this region with the bend region residues. In run 3, ligand 1 bound to the surface of the protofibrils by making sustained contacts with Ile 32D, Met 35D, Leu 34E. Ligands 2 and 3 bound to a cavity which was formed by the β-2 residues of chains D and E, while ligand 4 bound to a cavity formed by the β-2 residues of chains B and C. This is shown in Figure 6.3.

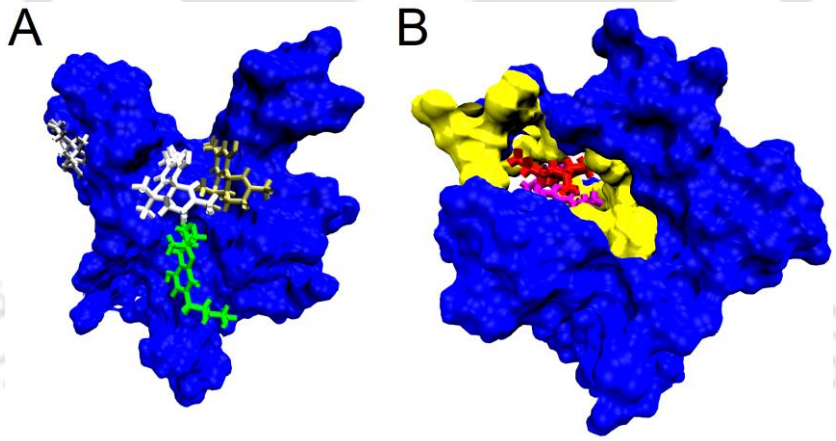


Figure 6.3. The difference in binding modes of the THC molecules in runs 2 and 3. Figure **A** shows the THC molecules bound to the surface of the protofibrils. Figure **B** shows the THC molecules bound to a cavity formed in the protofibrils.

In runs 1 and 2, the THC molecules bound to the surface of the protofibrils, while in run 3 they were also able to bind to cavities formed in the protofibrils. Compared to the other runs, the THC molecules in run 2 made the maximum number of contacts with residues in chains A, B, and C; while in run 3 the maximum number of contacts were made with the residues in chains

D and E. We have summarized the number of sustained contacts by the THC molecules in the three runs in Figure 6.4. The binding energy calculations show that the driving force for the binding of the THC molecules to the protofibrils was the van der Waals interactions due to hydrophobic contacts which were formed between the THC molecules and the protofibril residues in the β-2 strand. Previous studies have shown that the Aβ protofibrils were disrupted by drug candidate molecules which bind to the residues in the C-terminal region. In studies by Kuang *et al.* [43], and Wu *et al*. [64] Thioflavin – T could bind to the region around Met 35. Ile 32, Leu 34, and Val 36 were able to bind to the RR-AFC molecules [65]. Residues in this region were also able to interact with a resveratrol and clioquinol hybrid to destabilize the fibrils [66]. The residues Leu 34 and Val 36 in the hydrophobic region were able to interact with EGCG [67]. A drug candidate wgx-50 was able to destabilize the protofibrils by packing against the side chains of Ile 32 and Leu 34 [68]. Designed pentapeptides PPFFE and PPFYE by Goyal *et al.* interacted with the C-terminus residues to disrupt the protofibril structures [69].

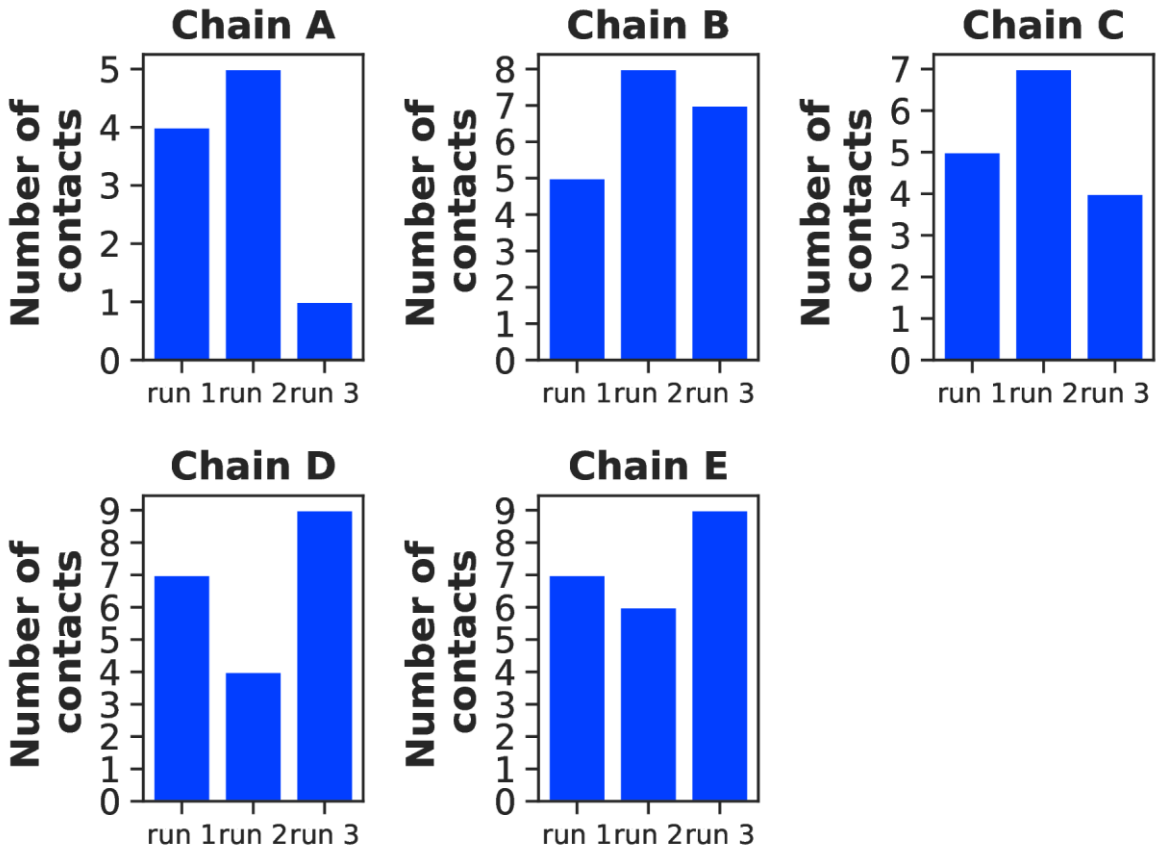


Figure 6.4. Number of sustained contacts made by the THC molecules in different chains in the three runs.

6.3.2 Destabilization of Aβ Protofibrils

In the presence of the THC molecules, the Aβ protofibrils were destabilized. The representative snapshots from the most occupied conformation cluster in the last 200 ns for each system were obtained from clustering analysis and are shown in Figure 6.5. The structural stability of the Aβ protofibrils in the absence of the THC molecules was examined using the C-α root-mean squared deviation (RMSD). The protofibrils were stable after about 200ns. The average RMSD values of the protofibrils in the last 200 ns in the three control systems were 0.87 ± 0.02 nm, 0.92 ± 0.03 nm, and 0.89 ± 0.02 nm, respectively. The RMSD values of the control systems are shown in Figure 6.6A. As observed in previous studies, there were twists in the peripheral chains A and E relative to their neighboring chains, due to their side chain atoms being free [30][68][70]. As a result, these chains contributed the most to the RMSD. As

these chains were at the periphery of the protofibrils, they tended to cause conformational changes until they maximized their contacts with the other chains [30]. In a study by Kuang *et al.* it was observed that when position restraints were applied to chains A and E, the structural of deviation of the protofibrils was minimized [43]. Overall, the characteristic "U-shape" formed by the β-1 – bend – β-2 regions was preserved.

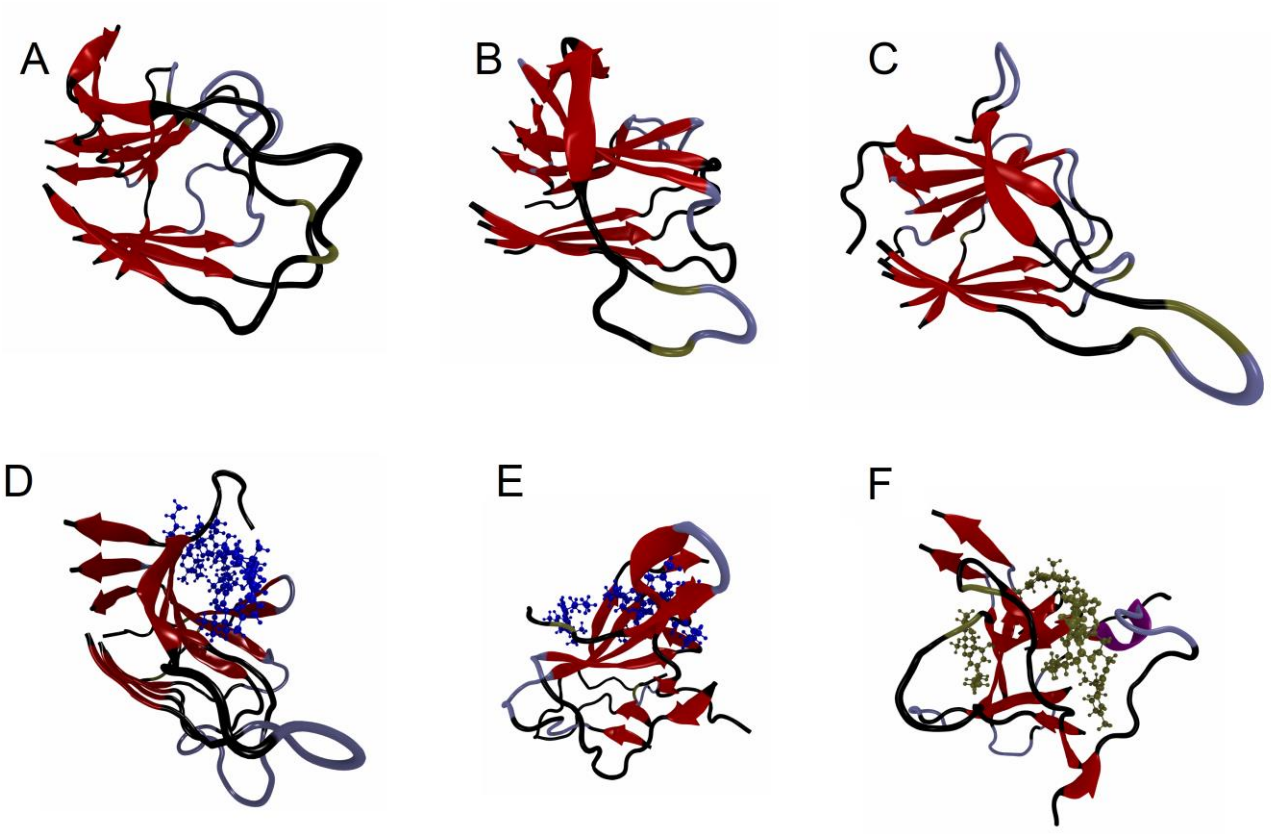


Figure 6.5. Representative snapshots from the most occupied conformation cluster for each system from the last 200ns. A. control 1. B. control 2. C. control 3. D. run 1. E. run 2. F. run 3.

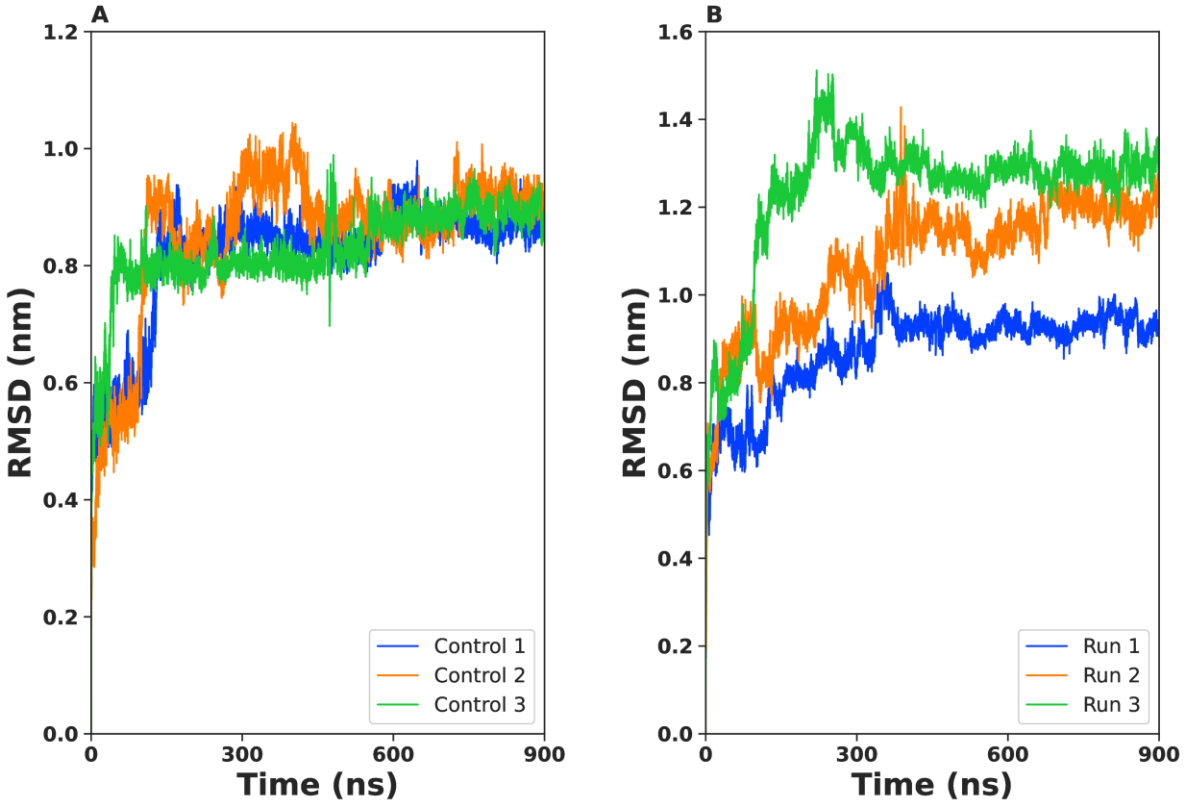


Figure 6.6. A. RMSD values of the protofibrils in the control Aβ protofibril systems. The major contribution to the RMSD was from the peripheral chains which were more exposed to the solvent than the interior chains. **B.** RMSD values of the THC-Aβ protofibril systems. Of the three runs, the protofibrils in run 3 had the highest RMSD values indicating the maximum disruption in the protofibril structure.

We compared the RMSD values of the Aβ protofibrils in the control systems with the RMSD values of the protofibrils in the presence of the THC molecule in the three runs. An increase in the RMSD value of the protofibrils in the presence of the inhibitors is associated with their destabilization, as shown in previous studies [30][66][67][68][70][71][72][73]. In the presence of the THC molecules, these values increased. This is indicative of the destabilization of the protofibrillar structure. Figure 6.6B shows the RMSD values in the three runs. The average RMSD values of the three runs in the last 200 ns were 0.93 ± 0.02 nm, 1.20 ± 0.03 nm and 1.29 ± 0.03 nm respectively. The RMSD values of the three runs were higher than the control systems, indicating that the THC molecules have successfully destabilized the protofibril structure. Run 3 had the highest RMSD of the three runs, on account of the THC

molecules binding to cavities in the protofibrils which moved chain E further away from chain D.

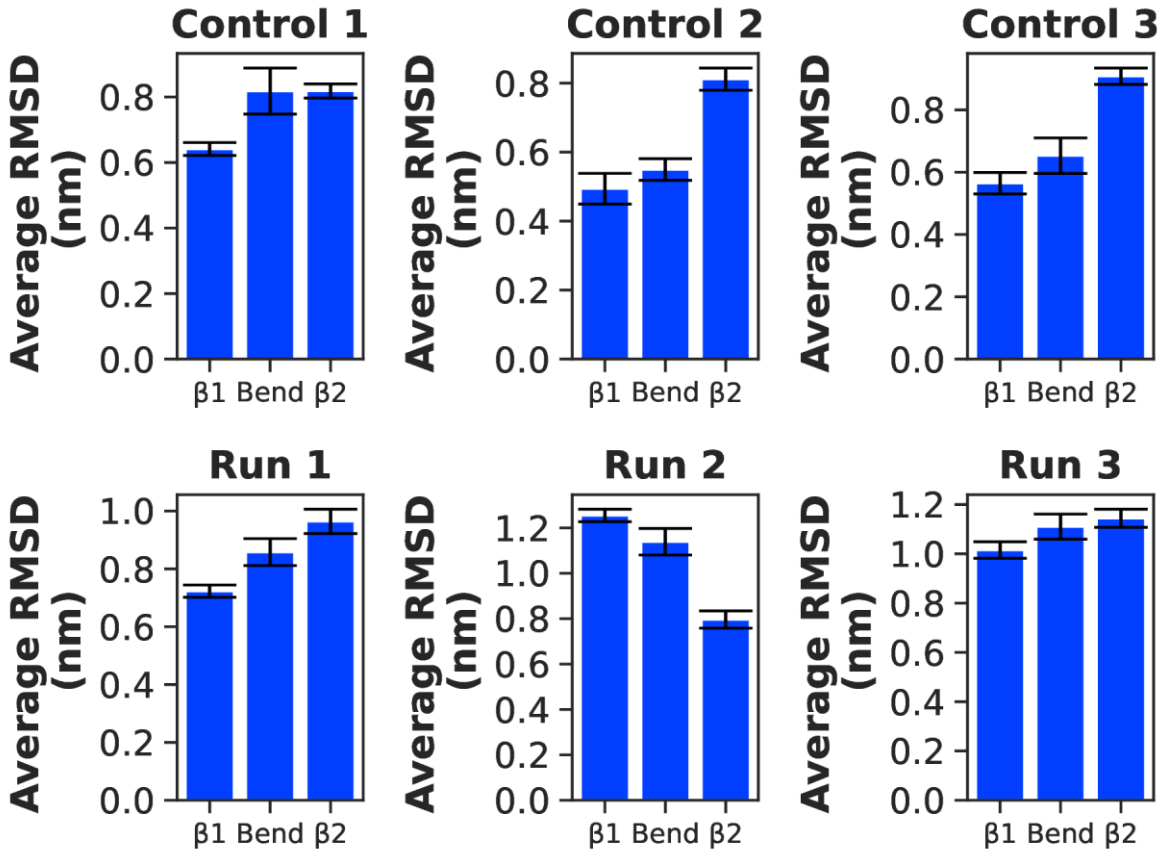


Figure 6.7. RMSD values in different regions of the protofibrils in the control systems and in the presence of the THC molecules. There was an increase in RMSD values in all regions, most notably in the β-1 and bend regions.

The THC molecules destabilized the protofibrils primarily in the β-1 and bend regions. The RMSD values of the three regions of the protofibrils, with the associated standard deviations represented by error bars are shown in Figure 6.7. The average RMSD of the β-1 region in run 2 was the highest compared to the other runs and the control systems. In the bend region, runs 2 and 3 had comparable values of the average RMSD, and were higher than run 1 and the control systems. The average RMSD value of the β-2 regions of the controls and runs 1 and 2 were comparable. Run 3 had an increased average RMSD of the β-2 region.

In the presence of the THC molecules, the average RMSD values of the β-1 and bend regions in the three runs and in the β-2 region of run 3 increased, indicating the destabilizing

effect of the THC molecules. The error bars in Figure 6.7 show the increased destabilization of the β-1 and bend regions in the three runs compared to the control systems 2 and 3. Thus, in the systems with the THC molecules, the protofibrils were destabilized and the destabilization of the β-1 and bend regions were more prominent.

In order to study the destabilization effect of the THC molecules on the protofibrils, we investigated the effect of these molecules on the secondary structure of the protofibrils. In the absence of the THC molecule, the average percentage of residues forming a β-sheet in the last 200ns in the Aβ protofibrils was 38.67%, 46.69%, and 44.44% in the three control systems respectively. In the presence of the THC molecules, these values decreased to 37.44%, 33.71% and 35.90% in the three runs, indicating that in these runs the THC molecule had destabilized the Aβ protofibrils. Figure 6.8A shows the β-sheet content of the control systems and the three runs.

In the control systems, the protofibril residues which preserved their β-sheet structure were in the region spanned by the residues Val 18 – Asp 23, Ile 31 – Val 36 and Val 39 – Ile 41. In the presence of the THC molecules, the loss of the β-sheet structure was observed in the residues spanning the region Ala 21 – Asp 23 and Ile 31 - Gly 33. The Glu 22 and Asp 23 residues participate in the formation of salt bridges with Lys 28, while the Ile 32 and Gly 33 residues play an important role in the hydrophobic packing of the protofibrils, as will be discussed in the next section. The loss of secondary structure was comparable in the three runs, with the maximum loss observed in chain E of runs 3, followed by chain E in run 2. Thus, the loss of secondary structure in these residues contributed to the destabilization of the protofibrils. Figures 6.9-6.14 show a plot of the time evolution of the secondary structures of the residues in which the regions which lose their β-sheet content can be identified.

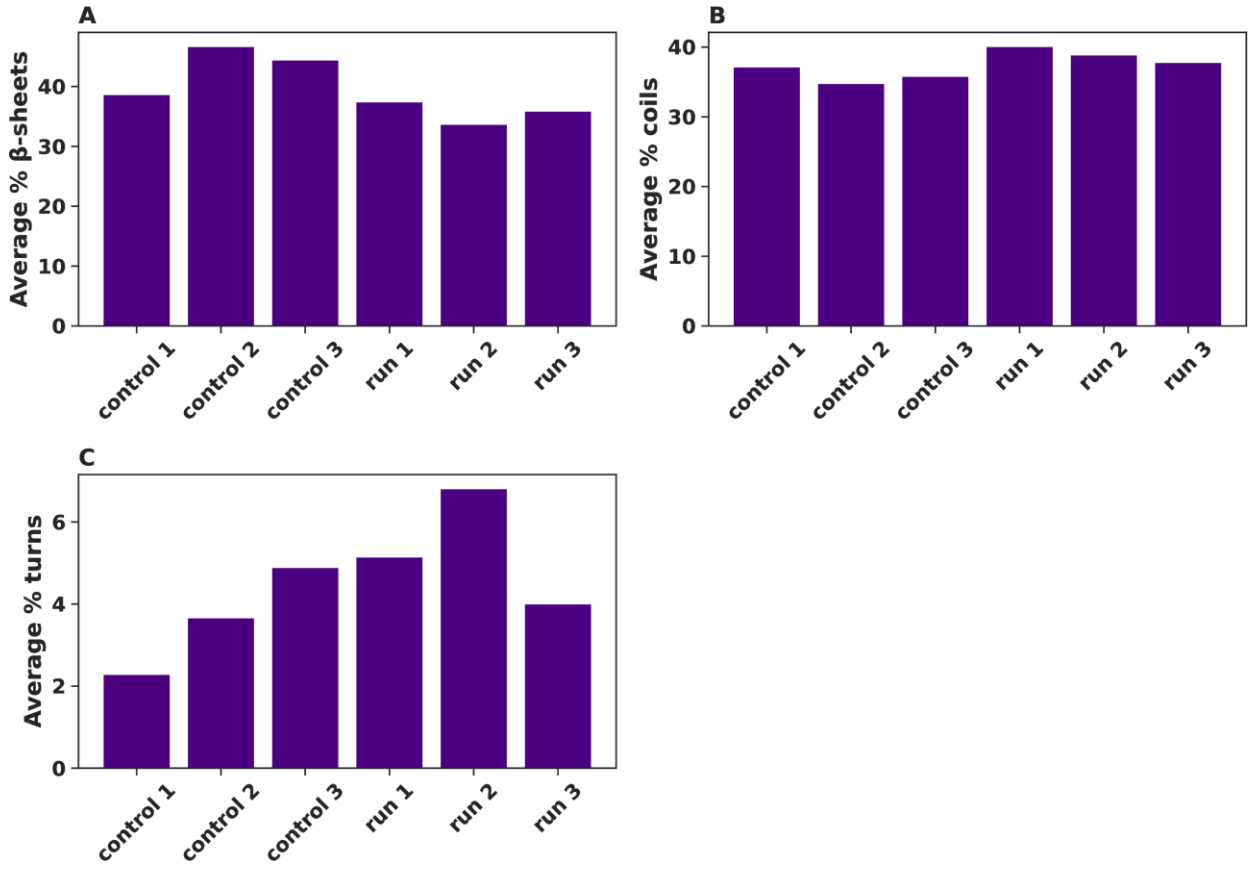


Figure 6.8. A. A plot of the average number of residues forming a β-sheet in the control Aβ protofibril systems and the three THC – Aβ protofibril systems in the last 200ns. It can be seen that in run 2 and run 3 there is a considerable reduced content of β-sheets. **B.** Percentage of residues forming coils in the last 200ns. The coil content of the protofibrils in runs 1 and 2 was significantly higher than the control systems. **C.** The percentage of protofibril residues which formed turns in the last 200ns. This value was higher in runs 2 and 3.

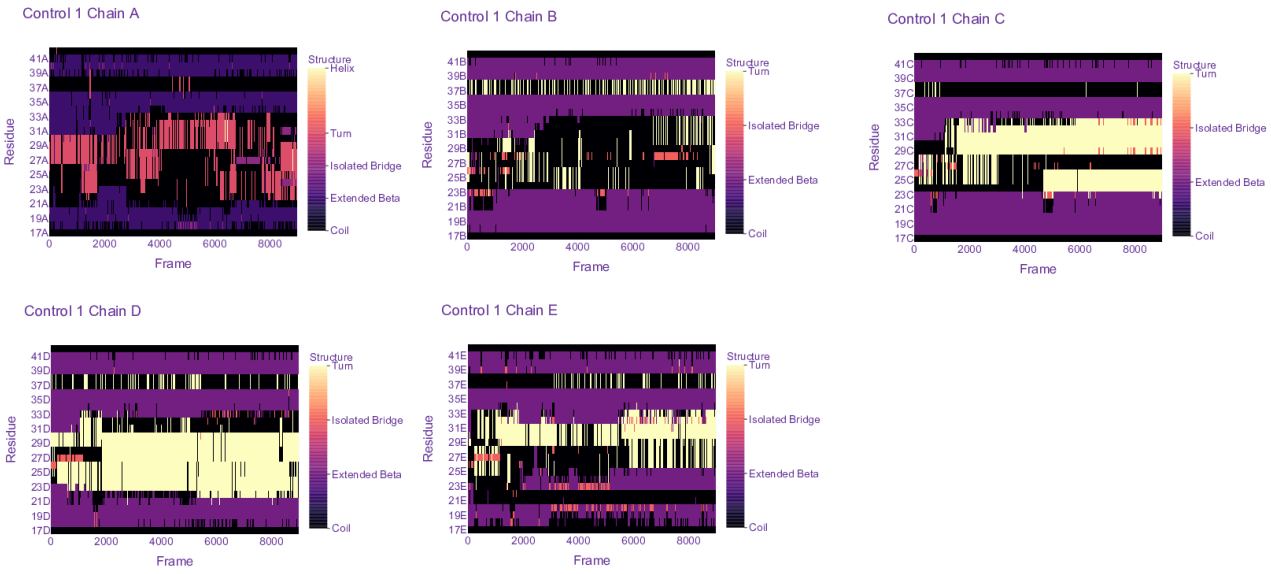


Figure 6.9. Time evolution of the secondary structure of control – 1

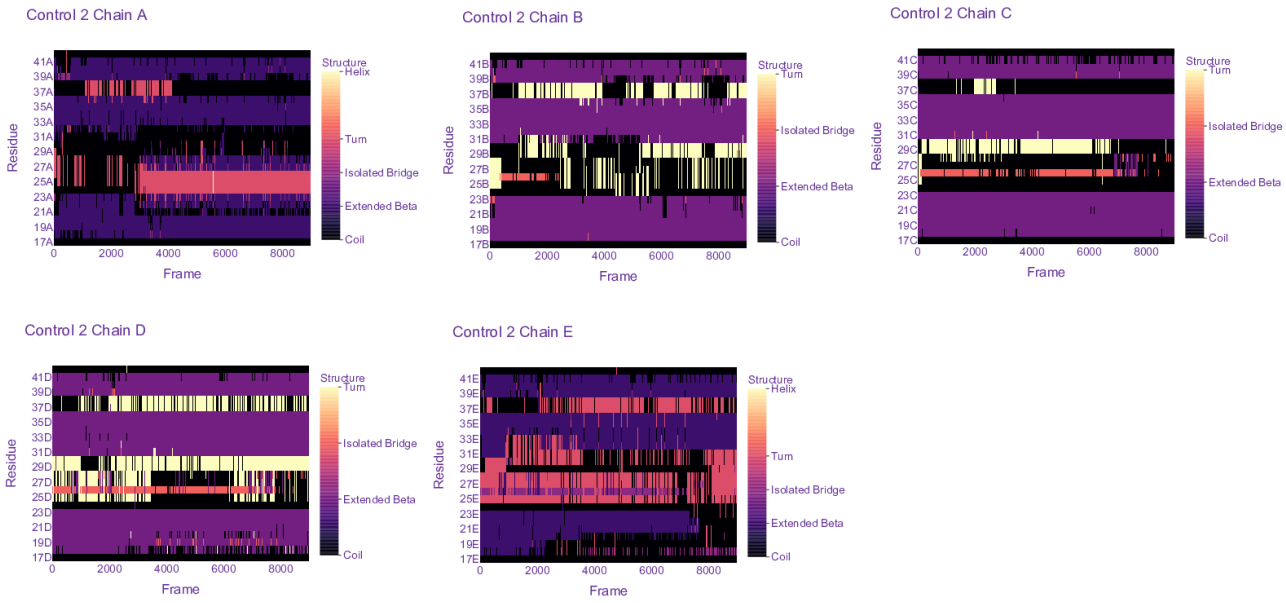


Figure 6.10. Time evolution of the secondary structure of control – 2

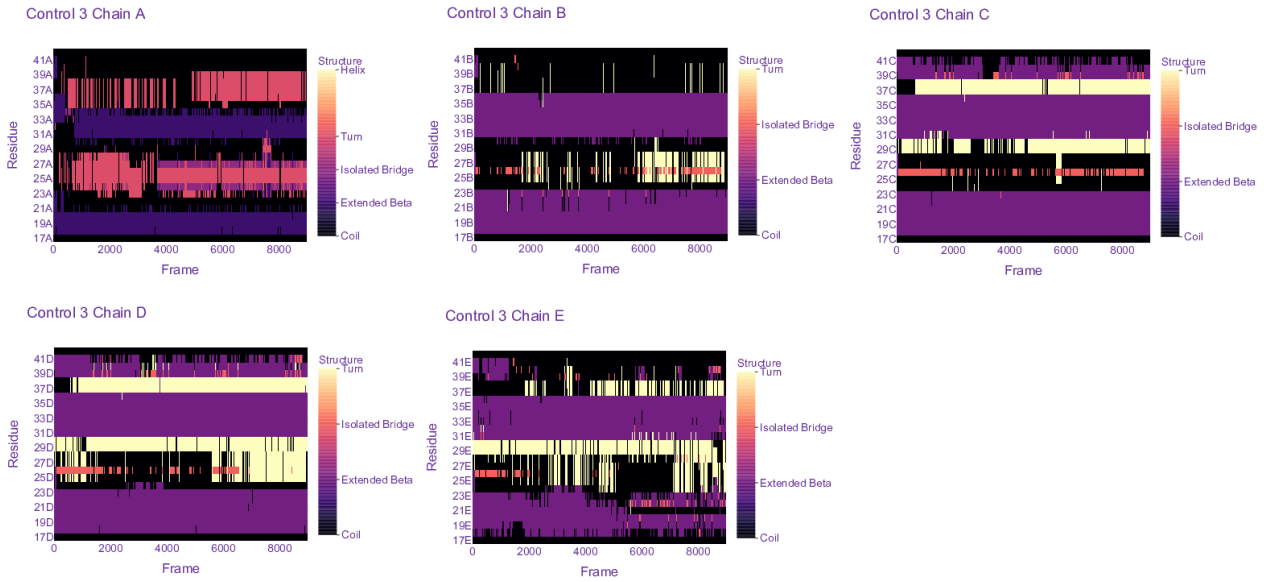


Figure 6.11. Time evolution of the secondary structure of control – 3

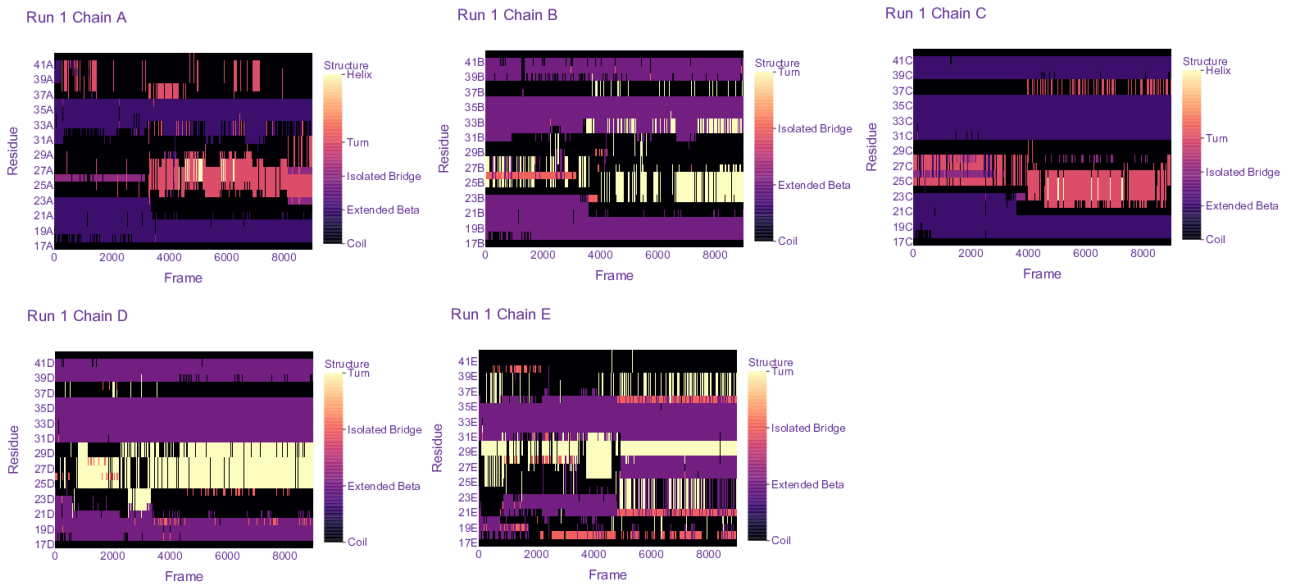


Figure 6.12. Time evolution of the secondary structure of run – 1

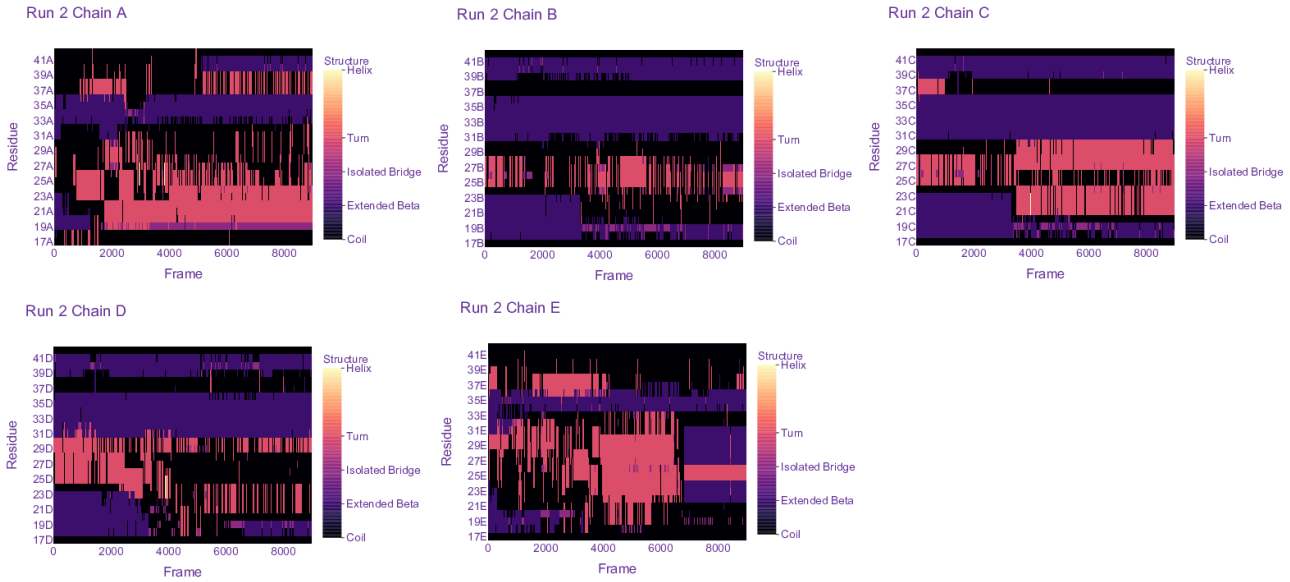


Figure 6.13. Time evolution of the secondary structure of run – 2

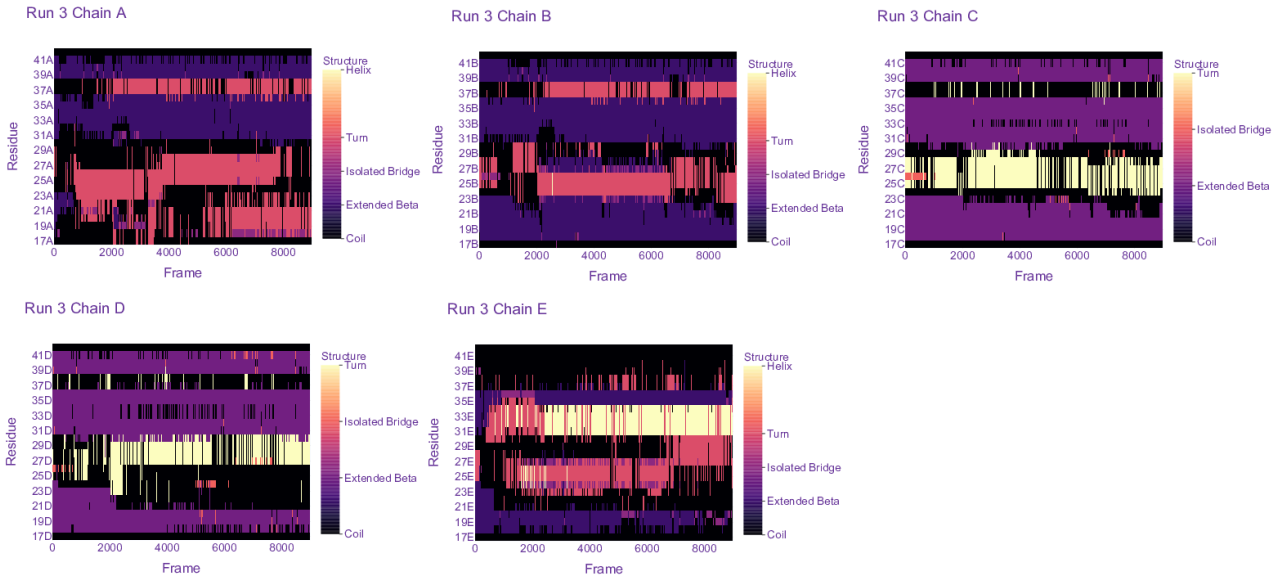


Figure 6.14. Time evolution of the secondary structure of run – 3

In the presence of the THC molecule, the number of residues which formed coils also increased. The average percentage of residues which formed coils in the three control systems was 37.18%, 34.80% and 35.83%. In the presence of the THC molecules, these values were 40.10%, 38.90%, and 37.82% in the three runs, indicating an increase in the coil content of the residues. This is shown in Figure 6.8B. The number of residues which formed turns also increased in runs 1 and 2, as shown in Figure 6.8C. Thus, in the presence of the THC molecule, the β-sheet content of the protofibrils reduced and the number of residues forming coils and turns increased, indicating the destabilization effect of the THC molecules on the secondary structure of the protofibrils. Table 6.4 contains a summary of the average number of residues in the amyloid protofibrils which formed beta-sheets, coils and turns.

Table 6.3. Average number of residues which formed various secondary structures in the last 200ns.

System	Average number of residues forming beta-sheets	Average number of residues forming coils	Average number of residues forming turns
Control – 1	48 ± 3	46 ± 3	3 ± 2
Control – 2	58 ± 4	43 ± 3	5 ± 2
Control – 3	55 ± 3	44 ± 3	6 ± 2
Run – 1	47 ± 4	50 ± 3	7 ± 2
Run – 2	42 ± 3	49 ± 3	9 ± 2
Run – 3	45 ± 4	48 ± 4	5 ± 2

In the presence of the THC molecules, the salt bridges in the protofibrils were disrupted leading to the destabilization of the protofibrils. The average number of salt bridges in the three control systems in the last 200ns of the simulations was 2.79 ± 0.85, 12.81 ± 2.0, and 8.06 ± 1.27. In the three runs in which the protofibrils were in the presence of the THC molecules, the average number of salt bridges was 5.08 ± 1.28, 4.51 ± 1.75, and 5.65 ± 1.56 respectively, as shown in Figure 6.15. The average number of salt bridges in the presence of the THC molecules was reduced compared to controls 2 and 3. We consider the low average number of salt bridges in control – 1 as an anomaly.

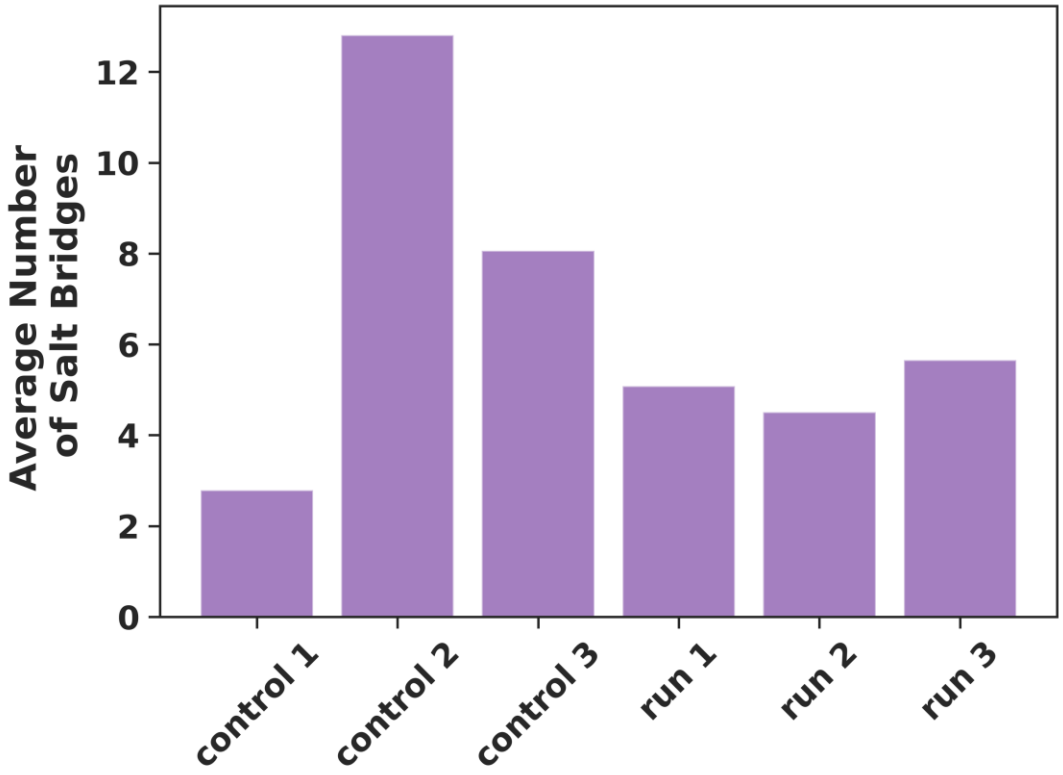


Figure 6.15. Average number of salt bridges in the systems with the Aβ protofibrils alone and in the systems with the Aβ protofibrils in the presence of the THC molecules.

We considered the destabilizing impact of the THC molecules on the protofibrils in disrupting the interchain salt bridges, which were important for the stability of these protofibrils. In the protofibrils, the salt bridges between the residues Asp 23 - Lys 28 were more prominent than the Glu 22 – Lys 28 salt bridges. Also, the interchain salt-bridges which existed were those which were between the residues Asp 23 of a chain and the Lys 28 of the next corresponding neighboring chain, which were, Asp 23A – Lys 28B, Asp 23B – Lys 28C, Asp 23C- Lys 28D, and Asp 23D – Lys 28E. The salt bridges in the reverse order did not exist, for example, Asp 23B – Lys 28A and Asp 23C – Lys 28B. Hence we report only the prominent interchain salt bridges. In order to characterize the probability of a salt bridge, we measured the percentage of duration of the total simulation time in which these Asp 23 and Lys 28 residues were within 4Å of each other for more than 40% of the total simulation time.

In control system 1, the only interchain salt bridge which was preserved was the one between Asp 23B – Lys 28C. In control 2, all the interchain salt-bridges were preserved, while in control 3, the salt bridges Asp 23B – Lys 28C and Asp 23C – Lys 28D were preserved. The salt bridges in controls 2 and 3 form the basis for the comparison with the interchain salt-bridges in the three runs in which the protofibrils were in the presence of the THC molecules. In the presence of the THC molecules, all the interchain salt-bridges were disrupted. Similarly, intrachain salt bridges were evaluated on the basis of sustained contacts between the Asp 23 and Lys 28 residues within the same chain. In controls 2 and 3 all the intrachain salt-bridges were preserved except the Asp 23A – Lys 28A salt bridge in control 2 and the Asp 23E – Lys 28E salt bridge in control 3. In the presence of the THC molecules, the intrachain salt bridges which were disrupted were those between Asp 23A – Lys 28A and Asp 23D – Lys 28D. In run 1, the Asp 23E – Lys 28E salt bridge was also lost. Due to the destabilization effect of the THC molecules in each of the three runs, important salt bridges which contribute to the stability of the protofibrils were disrupted and weakened. The interchain and intrachain salt bridges are summarized in Table 6.5.

Table 6.4. Sustained contacts between the salt-bridge forming residue pairs, expressed as a percentage of the time in which they within the cut-off distance for the formation of the salt bridge.

Salt Bridge Pair	Control 1	Control 2	Control 3	Run 1	Run 2	Run 3
Asp 23A – Lys 28B	-	65%	-	-	-	-
Asp 23B – Lys 28C	100%	100%	100%	-	-	-
Asp 23C – Lys 28D	-	99%	98%	-	-	-
Asp 23D – Lys 28E	-	60%	-	-	-	-
Asp 23A – Lys 28A	-	-	98%	-	-	-
Asp 23B – Lys 28B	-	100%	99%	100%	92%	51%
Asp 23C – Lys 28C	69%	100%	100%	95%	64%	51%
Asp 23D – Lys 28D	-	100%	95%	-	-	-
Asp 23E – Lys 28E	-	97%	-	-	64%	51%

Hydrogen bonds play an important role in the stability of the protofibrils. In the presence of the THC molecules, important hydrogen bonds in the protofibrils were disrupted, leading to the destabilization of the protofibrils. We calculated the average number of hydrogen bonds between neighboring chains. These are summarized in Table 6.6. The magnitude of the reduction of number hydrogen bonds in runs 2 and 3 were comparable, while run 1 had slightly more interchain hydrogen bonds. The loss of the interchain hydrogen bonds between chains B – C was more prominent in run 3, and chains C – D in run 2. These interchain hydrogen bonds were more preserved in run 1 comparatively, and hence the extent of destabilization was lower in run 1.

Table 6.5. Average inter-chain hydrogen bonds in the last 200ns

System	Chain A-B	Chain B-C	Chain C-D	Chain D-E
Control 1	13 ± 2	15 ± 1	11 ± 1	11 ± 1
Control 2	14 ± 1	13 ± 1	20 ± 1	11 ± 1
Control 3	8 ± 1	16 ± 1	19 ± 1	10 ± 1
Run 1	10 ± 1	13 ± 1	16 ± 2	8 ± 1
Run 2	9 ± 1	13 ± 1	11 ± 1	6 ± 1
Run 3	12 ± 2	7 ± 1	14 ± 2	6 ± 1

One of the ways of measuring the compactness of the fibril structure is by measuring the interchain Ala 21 – Val 36 distance (A21 – V36) [65][66]. In the presence of the THC molecules, the protofibril structure was made looser as the stabilizing hydrophobic contacts between Ala 21 – Val 36 of neighboring were disrupted. These are summarized in Table 6.7. The disruption of hydrophobic contacts in the protofibrils which are important for their stability by the THC molecules shows that the protofibrils lost their native structure and were destabilized. Another measure of measuring the tightness of the interatomic packing of the protofibrils is by measuring the solvent-accessible surface areas (SASA) of the protofibrils. In the presence of the THC molecules, the inter-atomic packing of the protofibrils was made looser and more exposed to the solvent. This indicates the destabilization of the protofibrils. These values are summarized in Table 6.7. The solvent plays a crucial role in stabilizing proteins [74]. An increase in the overall SASA of the protofibrils indicates that the protofibrils are exposed to more water and thus there is a weakening of the hydrophobic packing which makes them looser and hence less stable. The compactness of the protofibril structure is also characterized by the radius of gyration. In the presence of the THC molecules, the radius of gyration was higher in

the protofibrils, indicating that the protofibrils were destabilized. The average values of the radius of gyration in the last 200ns are shown in Table 6.7. Figure 6.16 shows the contact maps of all the systems from the last frame of the trajectory (900ns).

Table 6.6. Average distances between interchain Ala 21-Val 36 residues, SASA values and radius of gyration (RG) values, in the last 200ns.

System	A21A – V36B distance (nm)	A21B – V36C distance (nm)	A21C – V36D distance (nm)	A21D – V36E distance (nm)	Average SASA (nm^2)	Average RG (nm)
Control 1	0.74 ± 0.1	0.53 ± 0.1	0.34 ± 0.1	0.28	75.35 ± 2.1	1.40
Control 2	1.44 ± 0.1	1.01 ± 0.1	0.56 ± 0.1	0.24 ± 0.1	72.13 ± 1.5	1.38
Control 3	1.62 ± 0.1	1.12 ± 0.1	0.73	0.38 ± 0.1	76.85 ± 1.6	1.41
Run 1	1.45	1.27	1.08 ± 0.1	0.79 ± 0.1	76.39 ± 1.6	1.41
Run 2	1.78 ± 0.1	1.36 ± 0.1	1.24 ± 0.1	0.58 ± 0.1	84.67 ± 2.2	1.50
Run 3	1.60 ± 0.1	1.61 ± 0.2	1.41 ± 0.1	1.15 ± 0.1	87.53 ± 2.9	1.48

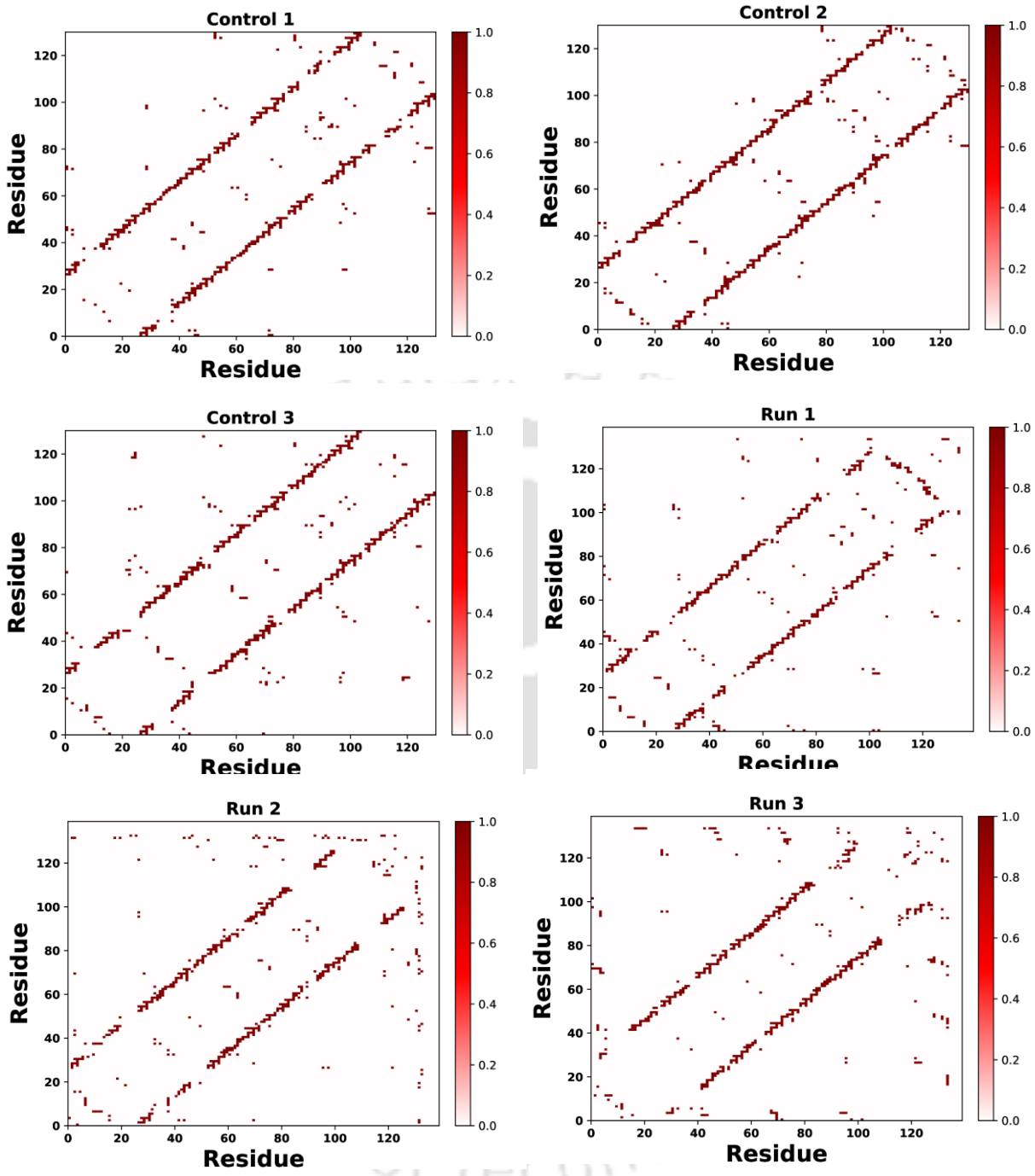


Figure 6.16. Contact maps of the systems at the last frame (900ns).

A tight binding of the THC molecules with the protofibrils, via strong hydrophobic interactions made the native packing in the hydrophobic core loose, as a result of which the protofibrils were exposed to more solvent molecules. A previous study showed that the packing between Ile 32 and Leu 34 is important for maintaining the hydration of the the Asp 23 – Lys

28 salt bridge [30]. In our simulations, this packing between the Ile 32 and Leu 34 residues was disrupted due to the THC molecules making strong hydrophobic contacts with the residues in this region. This led to the disruption of the salt bridges. The loss of the native hydrophobic contacts in the protofibrils further led to the decrease in stability, leading to the loss of hydrogen bonds. Thus, the THC molecules destabilized the protofibrils via hydrophobic binding.

An interesting avenue for future research would be the interaction of the THC molecules with protofibrils of increased lengths. A study by Dong *et al*. showed that an 11-monomer long single-fold protofibril structure is a more realistic representation of the fibrils [75]. In context of the U-shaped protofibrils used in the present study, Zhao *et al*. showed that depending on the pH, 12 to 15 layers of the monomers were more representative of the protofibrils [76]. In the present study, we observed the formation of helices in the process of destabilization. The transition from the beta sheet to helix leading to the disassembly of the protofibrils can be modulated by varying the pH, as shown in a study by Zhao *et al.* [77]. The disassembly of the S-shaped and U-shaped protofibrils follow different mechanisms as shown in a study by Xing *et al.* [78]. Another interesting avenue for future research would be the investigation of the interaction between the S-shaped protofibrils and the THC molecules. Neurotoxic oligomers exist in a variety of sizes, and in experimental and computational studies on the assembly of the protofibrils, pentamers and hexamers were observed to form paranuclei, which eventually formed larger oligomeric structures [79][80].

6.4 CONCLUSIONS

In our present study, we used all atom molecular dynamics simulations to investigate the interactions between THC molecules and Aβ protofibrils in order to characterize the binding sites and to see the impact of these molecules on the protofibril structure. We obtained insights into the destabilization of the protofibril structure by the THC molecules. Our results show that

the THC molecules bind to the hydrophobic surface formed by hydrophobic residues of the C-terminal of the protofibrils. Upon binding, the THC molecules formed strong hydrophobic contacts with the protofibril residues and weakened interchain interactions of the protofibrils. Hydrophobic interactions were the driving force for the binding of the THC molecules to the protofibrils. By competing for these hydrophobic contacts with other protofibril residues, the THC molecules disrupted the native contacts of the protofibrils and made them less stable. The tight inter-atomic packing of the protofibrils was made loose, and as a result interactions which were important for their stability were disrupted. The interchain hydrogen bonding network was disrupted, leading to the loss of β-sheet structure and the increase in the content of coils, turns, and helices. Salt bridge interactions which were critical for protofibril stability were disrupted, as a result of which the protofibrils were further destabilized. The native hydrophobic contacts which keep the salt bridges intact were disrupted. We have shown that when more than one THC molecule interacted with the protofibrils, the extent of destabilization was greater. We have identified the residues in the protofibrils which interacted with the THC molecules, these were in the hydrophobic region of the C-terminal region of the protofibrils. The binding regions were similar in the three replica simulations. While previous experimental work indicated the therapeutic effect of THC, the present study has explained the molecular mechanism of interactions. Our study shows that the THC molecule may be considered as a therapeutic drug candidate for the treatment of Alzheimer's disease.

REFERENCES

1. H.W. Querfurth, F.M. LaFerla, Alzheimer's disease, N Engl J Med. 362 (2010) 329–344.
2. C.A. Ross, M.A. Poirier, Protein aggregation and neurodegenerative disease, Nat Med. 10 (2004) S10–S17.
3. D.J. Selkoe, Alzheimer's disease: Genes, Proteins, and Therapy, Physiological Reviews. 81 (2001) 741–766.
4. Y. Ling, K. Morgan, N. Kalsheker, Amyloid precursor protein (APP) and the biology of proteolytic processing: relevance to Alzheimer's disease, The International Journal of Biochemistry & Cell Biology. 35 (2003) 1505–1535.

5. T. Kawasaki, K. Onodera, S. Kamijo, Identification of Novel Short Peptide Inhibitors of Soluble 37/48 kDa Oligomers of Amyloid β42, Bioscience, Biotechnology, and Biochemistry. 75 (2011) 1496–1501.
6. G.P. Gellermann, H. Byrnes, A. Striebinger, K. Ullrich, R. Mueller, H. Hillen, S. Barghorn, Aβ-globulomers are formed independently of the fibril pathway, Neurobiology of Disease. 30 (2008) 212–220.
7. D.J. Selkoe, Translating cell biology into therapeutic advances in Alzheimer's disease, Nature. 399 (1999) A23–A31. .
8. J. Cummings, G. Lee, A. Ritter, M. Sabbagh, K. Zhong, Alzheimer's disease drug development pipeline: 2019, Alzheimer's & Dementia: Translational Research & Clinical Interventions. 5 (2019) 272–293.
9. J.A. Lemkul, D.R. Bevan, The Role of Molecular Simulations in the Development of Inhibitors of Amyloid β-Peptide Aggregation for the Treatment of Alzheimer's Disease, ACS Chem. Neurosci. 3 (2012) 845–856.
10. K. Ono, Y. Yoshiike, A. Takashima, K. Hasegawa, H. Naiki, M. Yamada, Potent anti-amyloidogenic and fibril-destabilizing effects of polyphenols in vitro: implications for the prevention and therapeutics of Alzheimer's disease: Anti-amyloidogenic effects of polyphenols, Journal of Neurochemistry. 87 (2003) 172–181.
11. C. Rivière, J.-C. Delaunay, F. Immel, C. Cullin, J.-P. Monti, The Polyphenol Piceid Destabilizes Preformed Amyloid Fibrils and Oligomers In Vitro: Hypothesis on Possible Molecular Mechanisms, Neurochem Res. 34 (2009) 1120–1128.
12. J.E. Kim, M. Lee, Fullerene inhibits β-amyloid peptide aggregation, Biochemical and Biophysical Research Communications. 303 (2003) 576–579.
1. 13.X. Zhou, W. Xi, Y. Luo, S. Cao, G. Wei, Interactions of a Water-Soluble Fullerene Derivative with Amyloid-β Protofibrils: Dynamics, Binding Mechanism, and the Resulting Salt-Bridge Disruption, J. Phys. Chem. B. 118 (2014) 6733–6741.
2. 14.G. Zhang, M.J. Leibowitz, P.J. Sinko, S. Stein, Multiple-Peptide Conjugates for Binding β-Amyloid Plaques of Alzheimer's Disease, Bioconjugate Chem. 14 (2003) 86–92.
3. 15. R.-X. Gu, H. Gu, Z.-Y. Xie, J.-F. Wang, H. R. Arias, D.-Q. Wei and K.-C. Chou, Possible Drug Candidates for Alzheimers Disease Deduced from Studying their Binding Interactions with α7 Nicotinic Acetylcholine Receptor, Med. Chem. 5 (2009) 250
4. 16. M. Matsuoka, Humanin; A Defender Against Alzheimers Disease? RPCN. 4 (2009) 37–42.
5. 17. D. Goyal, S. Shuaib, S. Mann, B. Goyal, Rationally Designed Peptides and Peptidomimetics as Inhibitors of Amyloid-β (Aβ) Aggregation: Potential Therapeutics of Alzheimer's Disease, ACS Comb. Sci. 19 (2017) 55–80.
6. 18.V.A. Campbell, A. Gowran, Alzheimer's disease; taking the edge off with cannabinoids? Cannabinoids and Alzheimer's disease, British Journal of Pharmacology. 152 (2007) 655–662.
7. 19. O. Aizpurua-Olaizola, U. Soydaner, E. Öztürk, D. Schibano, Y. Simsir, P. Navarro, N. Etxebarria, A. Usobiaga, Evolution of the Cannabinoid and Terpene Content during the Growth of Cannabis sativa Plants from Different Chemotypes, J. Nat. Prod. 79 (2016) 324–331.
8. 20.S.H. Kim, J.W. Yang, K.H. Kim, J.U. Kim, T.H. Yook, A Review on Studies of Marijuana for Alzheimer's Disease – Focusing on CBD, THC, J. Pharmacopuncture 22 (2019) 225-230
9. 21. O. Aizpurua-Olaizola, U. Soydaner, E. Öztürk, D. Schibano, Y. Simsir, P. Navarro, N. Etxebarria, A. Usobiaga, Evolution of the Cannabinoid and Terpene Content during

the Growth of Cannabis sativa Plants from Different Chemotypes, J. Nat. Prod. 79 (2016) 324–331.
10. 22. A.A. Izzo, F. Borrelli, R. Capasso, V. Di Marzo, R. Mechoulam, Non-psychotropic plant cannabinoids: new therapeutic opportunities from an ancient herb, Trends in Pharmacological Sciences. 30 (2009) 515–527.
11. 23. R. Mechoulam, L.A. Parker, R. Gallily, Cannabidiol: An Overview of Some Pharmacological Aspects, The Journal of Clinical Pharmacology. 42 (2002) 11S-19S.
12. 24.L.M. Eubanks, C.J. Rogers, Beuscher, G.F. Koob, A.J. Olson, T.J. Dickerson, K.D. Janda, A Molecular Link between the Active Component of Marijuana and Alzheimer's Disease Pathology, Mol. Pharmaceutics. 3 (2006) 773–777.
13. 25.C. Cao, Y. Li, H. Liu, G. Bai, J. Mayl, X. Lin, K. Sutherland, N. Nabar, J. Cai, The Potential Therapeutic Effects of THC on Alzheimer's Disease, JAD. 42 (2014) 973–984.
14. 26. A. Currais, O. Quehenberger, A. M Armando, D. Daugherty, P. Maher, D. Schubert, Amyloid proteotoxicity initiates an inflammatory response blocked by cannabinoids, Npj Aging Mech Dis. 2 (2016) 16012.
15. 27. A. Bilkei-Gorzo, O. Albayram, A. Draffehn, K. Michel, A. Piyanova, H. Oppenheimer, M. Dvir-Ginzberg, I. Rácz, T. Ulas, S. Imbeault, I. Bab, J.L. Schultze, A. Zimmer, A chronic low dose of Δ9-tetrahydrocannabinol (THC) restores cognitive function in old mice, Nat Med. 23 (2017) 782–787.
29. E. Aso, A. Sánchez-Pla, E. Vegas-Lozano, R. Maldonado, I. Ferrer, Cannabis-Based Medicine Reduces Multiple Pathological Processes in AβPP/PS1 Mice, JAD. 43 (2014) 977–991.
30. T. Luhrs, C. Ritter, M. Adrian, D. Riek-Loher, B. Bohrmann, H. Dobeli, D. Schubert, R. Riek, 3D structure of Alzheimer's amyloid- (1-42) fibrils, Proceedings of the National Academy of Sciences. 102 (2005) 17342–17347.
31. J.A. Lemkul, D.R. Bevan, Assessing the Stability of Alzheimer's Amyloid Protofibrils Using Molecular Dynamics, J. Phys. Chem. B. 114 (2010) 1652–1660.
32. M.F. Masman, U.L.M. Eisel, I.G. Csizmadia, B. Penke, R.D. Enriz, S.J. Marrink, P.G.M. Luiten, In Silico Study of Full-Length Amyloid β 1−42 Tri- and Penta-Oligomers in Solution, J. Phys. Chem. B. 113 (2009) 11710–11719.
33. G. Bitan, M.D. Kirkitadze, A. Lomakin, S.S. Vollers, G.B. Benedek, D.B. Teplow, Amyloid -protein (A) assembly: A 40 and A 42 oligomerize through distinct pathways, Proceedings of the National Academy of Sciences. 100 (2003) 330–335.
34. M. Cheon, M. Kang, I. Chang, Polymorphism of fibrillar structures depending on the size of assembled Aβ17-42 peptides, Sci Rep. 6 (2016) 38196.
35. A. Kahler, H. Sticht, A.H.C. Horn, Conformational Stability of Fibrillar Amyloid-Beta Oligomers via Protofilament Pair Formation – A Systematic Computational Study, PLoS ONE. 8 (2013) e70521.
36. A. Aloisi, A. Barca, A. Romano, S. Guerrieri, C. Storelli, R. Rinaldi, T. Verri, Anti-Aggregating Effect of the Naturally Occurring Dipeptide Carnosine on Aβ1-42 Fibril Formation, PLoS ONE. 8 (2013) e68159.
37. I. Autiero, E. Langella, M. Saviano, Insights into the mechanism of interaction between trehalose-conjugated beta-sheet breaker peptides and Aβ(1–42) fibrils by molecular dynamics simulations, Mol. BioSyst. 9 (2013) 2835.
38. L. Cui, Y. Zhang, H. Cao, Y. Wang, T. Teng, G. Ma, Y. Li, K. Li, Y. Zhang, Ferulic Acid Inhibits the Transition of Amyloid-β42 Monomers to Oligomers but Accelerates the Transition from Oligomers to Fibrils, JAD. 37 (2013) 19–28.
39. P. Das, S. Kang, S. Temple, G. Belfort, Interaction of Amyloid Inhibitor Proteins with Amyloid Beta Peptides: Insight from Molecular Dynamics Simulations, PLoS ONE. 9 (2014) e113041.

40. W.-J. Du, J.-J. Guo, M.-T. Gao, S.-Q. Hu, X.-Y. Dong, Y.-F. Han, F.-F. Liu, S. Jiang, Y. Sun, Brazilin inhibits amyloid β-protein fibrillogenesis, remodels amyloid fibrils and reduces amyloid cytotoxicity, Sci. Rep. 5 (2015) 7992.
41. A. Espargaró, T. Ginex, M. del M. Vadell, M.A. Busquets, J. Estelrich, D. Muñoz-Torrero, F.J. Luque, R. Sabate, Combined in Vitro Cell-Based/in Silico Screening of Naturally Occurring Flavonoids and Phenolic Compounds as Potential Anti-Alzheimer Drugs, J. Nat. Prod. 80 (2017) 278–289.
42. H. He, J. Xu, D.-Y. Cheng, L. Fu, Y.-S. Ge, F.-L. Jiang, Y. Liu, Identification of Binding Modes for Amino Naphthalene 2-Cyanoacrylate (ANCA) Probes to Amyloid Fibrils from Molecular Dynamics Simulations, J. Phys. Chem. B. 121 (2017) 1211–1221.
43. M. Hernández-Rodríguez, J. Correa-Basurto, M.I. Nicolás-Vázquez, R. Miranda-Ruvalcaba, C.G. Benítez-Cardoza, A.A. Reséndiz-Albor, J.V. Méndez-Méndez, M.C. Rosales-Hernández, Virtual and In Vitro Screens Reveal a Potential Pharmacophore that Avoids the Fibrillization of Aβ1–42, PLoS ONE. 10 (2015) e0130263.
44. G. Kuang, N.A. Murugan, Y. Tu, A. Nordberg, H. Ågren, Investigation of the Binding Profiles of AZD2184 and Thioflavin T with Amyloid-β(1–42) Fibril by Molecular Docking and Molecular Dynamics Methods, J. Phys. Chem. B. 119 (2015) 11560–11567.
45. T. Sterling, J.J. Irwin, ZINC 15 – Ligand Discovery for Everyone, J. Chem. Inf. Model. 55 (2015) 2324–2337.
46. V. Zoete, M.A. Cuendet, A. Grosdidier, O. Michielin, SwissParam: A fast force field generation tool for small organic molecules, J. Comput. Chem. 32 (2011) 2359–2368.
47. J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B.L. de Groot, H. Grubmüller, A.D. MacKerell, CHARMM36m: an improved force field for folded and intrinsically disordered proteins, Nat Methods. 14 (2017) 71–73.
48. S. Nosé, A unified formulation of the constant temperature molecular dynamics methods, The Journal of Chemical Physics. 81 (1984) 511–519.
49. W.G. Hoover, Canonical dynamics: Equilibrium phase-space distributions, Phys. Rev. A. 31 (1985) 1695–1697.
50. M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method, Journal of Applied Physics. 52 (1981) 7182–7190.
51. S. Nosé, M.L. Klein, Constant pressure molecular dynamics for molecular systems, Molecular Physics. 50 (1983) 1055–1076.
52. M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, E. Lindahl, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX. 1–2 (2015) 19–25.
53. B. Hess, P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation, J. Chem. Theory Comput. 4 (2008) 116–122.
54. U. Essmann, L. Perera, M. Berkowitz, T. Darden, H. Lee, L. Pedersen, A smooth
16. particle mesh ewald method, J. Chem. Phys. 103 (1995) 8577-8593.
55. W.G. Touw, C. Baakman, J. Black, T.A. te Beek, E. Krieger, R.P. Joosten, G. Vriend, A series of PDB related databases for everyday needs, Nucleic Acids Res. 43 (2015) D364-D368.
56. W. Kabsch, C. Sander, Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features, Biopolymers. 22 (1983) 2577–2637.
57. D. van der Spoel, P.J. van Maaren, P. Larsson, N. Tîmneanu, Thermodynamics of Hydrogen Bonding in Hydrophilic and Hydrophobic Media, J. Phys. Chem. B. 110 (2006) 4393–4398.
58. W. Humphrey, A. Dalke, K. Schulten, VMD: Visual molecular dynamics, Journal of Molecular Graphics. 14 (1996) 33–38.

59. S. Ghosh, T. Bierig, S. Lee, S. Jana, A. Löhle, G. Schnapp, C.S. Tautermann, N. Vaidehi, Engineering Salt Bridge Networks between Transmembrane Helices Confers Thermostability in G-Protein-Coupled Receptors, J. Chem. Theory Comput. 14 (2018) 6574–6585.
60. F. Eisenhaber, P. Lijnzaad, P. Argos, C. Sander, M. Scharf, The double cubic lattice method: Efficient approaches to numerical integration of surface area and volume and to dot surface contouring of molecular assemblies, J. Comput. Chem. 16 (1995) 273–284.
61. J. Srinivasan, T.E. Cheatham, P. Cieplak, P.A. Kollman, D.A. Case, Continuum Solvent Studies of the Stability of DNA, RNA, and Phosphoramidate-DNA Helices, J. Am. Chem. Soc. 120 (1998), 9401–9409
62. P.A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D.A. Case, T.E. Cheatham, Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models, Acc. Chem. Res. 33 (2000) 889–897.
63. R. Kumari, R. Kumar, Open Source Drug Discovery Consortium, A. Lynn, g_mmpbsa —A GROMACS Tool for High-Throughput MM-PBSA Calculations, J. Chem. Inf. Model. 54 (2014) 1951–1962.
64. N.A. Baker, D. Sept, S. Joseph, M.J. Holst, J.A. McCammon, Electrostatics of nanosystems: Application to microtubules and the ribosome, Proceedings of the National Academy of Sciences. 98 (2001) 10037–10041.
65. C. Wu, M.T. Bowers, J.-E. Shea, On the Origin of the Stronger Binding of PIB over Thioflavin T to Protofibrils of the Alzheimer Amyloid-β Peptide: A Molecular Dynamics Study, Biophysical Journal. 100 (2011) 1316–1324.
66. S.S. Barale, R.S. Parulekar, P.M. Fandilolu, M.J. Dhanavade, K.D. Sonawane, Molecular Insights into Destabilization of Alzheimer's Aβ Protofibril by Arginine Containing Short Peptides: A Molecular Modeling Approach, ACS Omega. 4 (2019) 892–903.
67. R.K. Saini, S. Shuaib, D. Goyal, B. Goyal, Insights into the inhibitory mechanism of a resveratrol and clioquinol hybrid against Aβ 42 aggregation and protofibril destabilization: A molecular dynamics simulation study, Journal of Biomolecular Structure and Dynamics. 37 (2019) 3183–3197.
68. C. Zhan, Y. Chen, Y. Tang, G. Wei, Green Tea Extracts EGCG and EGC Display Distinct Mechanisms in Disrupting Aβ 42 Protofibril, ACS Chem. Neurosci. 11 (2020) 1841–1851.
69. H.-M. Fan, R.-X. Gu, Y.-J. Wang, Y.-L. Pi, Y.-H. Zhang, Q. Xu, D.-Q. Wei, Destabilization of Alzheimer's Aβ42 Protofibrils with a Novel Drug Candidate wgx-50 by Molecular Dynamics Simulations, J. Phys. Chem. B. 119 (2015) 11196–11202.
70. S. Shuaib, S.S. Narang, D. Goyal, B. Goyal, Computational design and evaluation of β-sheet breaker peptides for destabilizing Alzheimer's amyloid-β 42 protofibrils, J Cell Biochem. 120 (2019) 17935–17950.
71. L. Tran, J. Kaffy, S. Ongeri, T. Ha-Duong, Binding Modes of a Glycopeptidomimetic Molecule on Aβ Protofibrils: Implication for Its Inhibition Mechanism, ACS Chem. Neurosci. 9 (2018) 2859–2869.
17. 71.Y. Mo, J. Lei, Y. Sun, Q. Zhang, G. Wei, Conformational Ensemble of hIAPP Dimer: Insight into the Molecular Mechanism by which a Green Tea Extract inhibits hIAPP Aggregation, Sci Rep. 6 (2016) 33076.
72. J.A. Lemkul, D.R. Bevan, Destabilizing Alzheimer's Aβ 42 Protofibrils with Morin: Mechanistic Insights from Molecular Dynamics Simulations, Biochemistry. 49 (2010) 3935–3946.

73. D. Radziuk, H. Möhwald, Ultrasonically treated liquid interfaces for progress in cleaning and separation processes, Phys. Chem. Chem. Phys. 18 (2016) 21–46.
74. S. Ali, Md. Hassan, A. Islam, F. Ahmad, A Review of Methods Available to Estimate Solvent-Accessible Surface Areas of Soluble Proteins in the Folded and Unfolded States, CPPS. 15 (2014) 456–476.
75. M. Dong, T.J. Paul, Z. Hoffmann, K. Chan, D. Hu, H. Ai, R. Prabhakar, Structural and Material Properties of Amyloid Aβ40/42 Fibrils, ChemPhysChem. 17 (2016) 2558–2566.
76. W. Zhao, X. Xing, B. Kang, X. Zhu, H. Ai, Positive effect of strong acidity on the twist of Aβ42 fibrils and the counteraction of Aβ42 N-terminus, Journal of Molecular Graphics and Modelling. 82 (2018) 59–66.
77. W. Zhao, H. Ai, Effect of pH on Aβ42 Monomer and Fibril-like Oligomers—Decoding in Silico of the Roles of pK Values of Charged Residues, ChemPhysChem. 19 (2018) 1103–1116.
78. X. Xing, C. Liu, A. Ali, B. Kang, P. Li, H. Ai, Novel Disassembly Mechanisms of Sigmoid Aβ42 Protofibrils by Introduced Neutral and Charged Drug Molecules, ACS Chem. Neurosci. 11 (2020) 45–56.
79. G. Bitan, M.D. Kirkitadze, A. Lomakin, S.S. Vollers, G.B. Benedek, D.B. Teplow, Amyloid β-protein (Aβ) assembly: Aβ40 and Aβ42 oligomerize through distinct pathways, Proc Natl Acad Sci USA. 100 (2003) 330.
80. M. Cheon, M. Kang, I. Chang, Polymorphism of fibrillar structures depending on the size of assembled Aβ17-42 peptides, Scientific Reports. 6 (2016) 38196.

Chapter 7. Summary and Future Scope of Research

7.1. Summary

There is no cure for Alzheimer's disease at present. In the present thesis, the interaction of the amyloid protofibrils with novel designed peptide sequences and the molecule THC was investigated. In Chapter 3, we found that that oligoproline chains of various lengths could break the beta-sheet structure of the protofibrils and induce the formation of random coils. Critical interactions which are important for protofibril stability such as hydrogen bonds and salt bridges were disrupted. Proline was able to bind strongly to the protofibrils by strong electrostatic interactions, making proline an important amino acid to consider in the design of novel peptide-based drugs.

In Chapter 4, we build on our observations that proline can bind well to the protofibrils and disrupt their structure. We explored the possibility of combining a pentamer proline sequence along with a well-known beta-sheet breaker peptide KLVFF to design a sequence $KLVFFP_5$ which was designed to exploit the properties of the self-recognition sequence of the amyloid-β peptide KLVFF and the β-sheet breaker amino acid proline. This peptide could destabilize the amyloid protofibrils to a greater extent than the KLVFF peptide. In the presence of the $KLVFFP_5$ peptide the protofibrils lost their beta-sheet structure leading to the formation of coils and helices. The hydrogen bonding network of the protofibrils and the salt bridges are critical for protofibril stability. The $KLVFFP_5$ peptide disrupted the hydrogen bonding network and the salt bridges in the protofibrils to a greater extent than the KLVFF peptide. The tight interatomic packing of the protofibrils was made looser by the $KLVFFP_5$ peptide.

In Chapter 5, we considered a model of the protofibrils which is known to be a particularly difficult target for drugs. The increased presence of aromatic amino acids can enhance the binding of a known beta-sheet breaker peptide LPFFD, when it was modified by these aromatic amino acids. The aromatic amino acids enhanced the binding affinity by forming aromatic and hydrophobic contacts. The electrostatic complementarity of the ligands and the protofibrils, along with favourable orientation of these ligands played an important role in the binding. The tryptophan modified LPFFD peptides had the highest binding affinity.

In Chapter 6, we showed that THC molecules could destabilize the amyloid-beta protofibrils. We obtained insights into the destabilization of the protofibril structure by the THC molecules. Our results show that the THC molecules bind to the hydrophobic surface formed by hydrophobic residues of the C-terminal of the protofibrils. Upon binding, the THC molecules formed strong hydrophobic contacts with the protofibril residues and weakened interchain interactions of the protofibrils. Hydrophobic interactions were the driving force for the binding of the THC molecules to the protofibrils. By competing for these hydrophobic contacts with other protofibril residues, the THC molecules disrupted the native contacts of the protofibrils and made them less stable.

7.2 Future Scope of Research

In Chapter 3, we had shown that proline can disrupt the beta-sheet structure of pre-formed amyloid-beta protofibrils. An exciting avenue for future research would be to investigate if proline can interefere in the aggregation of the amyloid protofibrils. Based on the observations made in Chapter 4, it would be interesting to functionalize the nanoparticles with the designed peptide which may be able to cross the blood-brain barrier.In Chapter 5, we considered various aromatic modifications of the beta-sheet breaker peptide LPFFD. It would be interesting to modify the KLVFF sequence of the amyloid protofibrils with aromatic amino acids and study their ability to form beta-sheets with each other. Such a study can determine which aromatic modification can form beta-sheets in the shortest period of time, which can be useful in rapidly preventing the further aggregation of these protofibrils. Another interesting avenue for future research would be the investigation of the aggregation of the amyloid protofibrils in the presence of different concentrations of the THC molecules. While we had shown that the THC molecules could disrupt the protofibril structure, its role in the aggregation process has not been investigated in computer simulations so far. The results of the present study may be applicable to mature fibrils, since they are repeat units of the protofibrils. Better statistics may be obtained with more replicates, and by running longer simulations. For a more realistic simulation of the aggregation process, we can incorporate the effects of crowding: the presence of other molecules in the vicinity of the systems studied.

Research Output

1. P.K. Kanchi, A.K. Dasmahapatra, Polyproline chains destabilize the Alzheimer's amyloid-β protofibrils: A molecular dynamics simulation study, J. Mol. Graph. Model. 93 (2019) 107456.

2. P.K. Kanchi, A.K. Dasmahapatra, Enhancing the binding of the β-sheet breaker peptide LPFFD to the amyloid-β fibrils by aromatic modifications: A Molecular Dynamics Simulation study, Comput. Biol. Chem. (2021) 107471.

3. P.K. Kanchi, A.K. Dasmahapatra, Destabilization of the Alzheimer's Amyloid-β Protofibrils by THC: A Molecular Dynamics Simulation study, J. Mol. Graph. Model., 105, 107889

4. P.K. Kanchi, A.K. Dasmahapatra, Destabilization of the Alzheimer's Amyloid-β Peptide by a Proline-rich β-sheet Breaker Peptide: A Molecular Dynamics Simulation study **(submitted)**

Conference Poster

1. Interaction of proline chains with the Alzheimer's amyloid-beta protofibrils: proline induces alpha and 3_{10} helix formation. COMPLU 2018, IIT Roorkee

APPENDIX A.

ITP file (topology) of the THC molecule

nr	type	resnr	resid	atom	cgnr	charge	mass
1	CB	1	LIG	C6	1	-0.1435	12.011
2	CB	1	LIG	C7	2	-0.15	12.011
3	CB	1	LIG	C8	3	0.0825	12.011
4	CB	1	LIG	C10	4	-0.1435	12.011
5	CB	1	LIG	C11	5	0.0825	12.011
6	CB	1	LIG	C12	6	-0.15	12.011
7	CR	1	LIG	C1	7	0	12.011
8	CR	1	LIG	C2	8	0	12.011
9	CR	1	LIG	C3	9	0	12.011
10	CR	1	LIG	C4	10	0	12.011
11	CR	1	LIG	C5	11	0.1435	12.011
12	OR	1	LIG	O9	12	-0.5325	15.9994
13	OR	1	LIG	O13	13	-0.3625	15.9994
14	CR	1	LIG	C14	14	0.28	12.011
15	CR	1	LIG	C15	15	0	12.011
16	CR	1	LIG	C16	16	0	12.011
17	CR	1	LIG	C17	17	0	12.011
18	HCMM	1	LIG	H18	18	0	1.0079
19	CR	1	LIG	C19	19	0	12.011
20	CR	1	LIG	C20	20	0.1382	12.011
21	C=C	1	LIG	C21	21	-0.2764	12.011
22	CR	1	LIG	C22	22	0.1382	12.011
23	C=C	1	LIG	C23	23	-0.2882	12.011
24	CR	1	LIG	C24	24	0.2817	12.011
25	HCMM	1	LIG	H25	25	0	1.0079
26	HCMM	1	LIG	H26	26	0	1.0079
27	HCMM	1	LIG	H27	27	0	1.0079
28	HCMM	1	LIG	H28	28	0	1.0079
29	HCMM	1	LIG	H29	29	0	1.0079
30	HCMM	1	LIG	H30	30	0	1.0079
31	HCMM	1	LIG	H31	31	0	1.0079
32	HCMM	1	LIG	H32	32	0	1.0079
33	HCMM	1	LIG	H33	33	0	1.0079
34	HCMM	1	LIG	H34	34	0	1.0079
35	HCMM	1	LIG	H35	35	0	1.0079
36	HCMM	1	LIG	H36	36	0	1.0079
37	HCMM	1	LIG	H37	37	0.15	1.0079
38	HOCC	1	LIG	H38	38	0.45	1.0079
39	HCMM	1	LIG	H39	39	0.15	1.0079

40	HCMM	1	LIG	H40	40	0	1.0079
41	HCMM	1	LIG	H41	41	0	1.0079
42	HCMM	1	LIG	H42	42	0	1.0079
43	HCMM	1	LIG	H43	43	0	1.0079
44	HCMM	1	LIG	H44	44	0	1.0079
45	HCMM	1	LIG	H45	45	0	1.0079
46	HCMM	1	LIG	H46	46	0	1.0079
47	HCMM	1	LIG	H47	47	0	1.0079
48	HCMM	1	LIG	H48	48	0	1.0079
49	HCMM	1	LIG	H49	49	0	1.0079
50	HCMM	1	LIG	H50	50	0	1.0079
51	HCMM	1	LIG	H51	51	0	1.0079
52	HCMM	1	LIG	H52	52	0	1.0079
53	HCMM	1	LIG	H53	53	0.15	1.0079

APPENDIX B.

1. **Secondary structure analysis**: The secondary structure of each time frame is computed by using the GROMACS command gmx do_dssp. The command used is:

gmx_mpi do_dssp -f xtcfile.xtc -s tprfile.tpr -o outputfile.xpm -ver 2

Here ver 2 refers to the version of the dssp program used. The xpm file produced contains the secondary structure assignment for each residue and time and can be converted to an eps file using the command:

gmx xpm2ps -f dssp.xpm –di file.m2p -o output.eps

An m2p file may be provided as input which contains parameters for settings for the x and y axes such as font names and size of ticks. The data in the xpm file can also be used to plot the total number of residues in a particular secondary structure as a function of time.

2. **Hydrogen Bonds:** The number of hydrogen bonds as a function of time may be calculated using the folling GROMACS command:

gmx hbond -f xtcfile.xtc -s tprfile.tpr -n indexfile.ndx –num outputfile.xvg

Here the indexfile.ndx is an optional index file supplied to calculate the hydrogen bonds between two groups of interest, such as two chains. GROMACS also displays the average number of hydrogen bonds over the entire simulation period, although this is not written to a separate file.

3. **Salt Bridge analysis**: The number of salt bridges as a function of time can be calculated by using VMD's Timeline plugin. These salt bridges can also be visualized and we can identify intrapeptide and interpeptide salt bridges. A data file may be written which gives an output of the existence of a particular salt bridge for a given frame. We can add the total number of salt

bridges for every time frame and make a plot of the total number of salt bridges as a function of time.

4. **Solvent accessible surface area**: The following GROMACS command calculates the solvent accessible surface area as a function of time:

gmx sasa -f xtcfile.xtc -s tprfile.tpr -n indexfile.ndx -o outputfile.xvg

5. **Hydrophobic interactions:** Hydrophobic contacts can be calculated by the Protein Interactions Calculator, a web-server located at http://pic.mbu.iisc.ernet.in/. A pdb file is supplied as input. Output consists of information of interactions such as disulphide bonds, hydrophobic interactions, ionic interactions, hydrogen bonds, aromatic- aromatic interactions, aromatic-sulphur interactions and cation - π interactions within a protein or between proteins in a complex, the accessible surface area as well as the distance of a residue from the surface of the protein. Hydrophobic contacts may be visualized by the software LigPlot+. A pdb file is supplied as input and a 2D representation is the output. A 3D output may be visualized in Rasmol and Pymol.

6. Sustained Contacts: The following modified script can calculate sustained contacts and write them to an output text file.

```
proc contactFreq { {sel1} {sel2} {percent 0} {outFile stdout} {mol
top} } {
  if { $outFile != "stdout" } {
     set outFile [open $outFile w]
  }
  puts $outFile "[clock format [clock scan now]] Search started."

  set percent 40
  set allAtoms {}
  set allCount {}
  set numberOfFrames [molinfo $mol get numframes]
  for { set i 0 } { $i < $numberOfFrames } { incr i } {
      molinfo $mol set frame $i

      set frameCount {}
      set frameAtoms [atomselect $mol "$sel1 and noh and within 4 of
($sel2 and noh)"]
```

```
      foreach {segid} [$frameAtoms get segid] {resname} [$frameAtoms
get resname] {resid} [$frameAtoms get resid] {name} [$frameAtoms get
name] {
        set atom [list $resid $resname $segid ]
        if {[lsearch $frameCount $atom] != -1} continue
        lappend frameCount $atom
        set loc [lsearch $allAtoms $atom]
           if { $loc == -1 } {
              lappend allAtoms $atom
              lappend allCount 1
           } else {
              lset allCount $loc [expr { [lindex $allCount $loc] + 1 } ]
           }
     }
     $frameAtoms delete
  }

  puts $outFile "[clock format [clock scan now]] Search finished."

  puts $outFile "Find interactions:"
  puts $outFile "Residue \t\tfraction"
  #print count after sorting
  set outData {}
  foreach { a } $allAtoms { c } $allCount {
      lappend outData [concat $c $a]
  }
  foreach { data } [lsort -integer -index 1 $outData] {
      set c [lindex $data 0]
      set fraction [expr { 100*$c/($numberOfFrames+0.0) }]
      if { $fraction >= $percent } {
       puts $sel1
          puts $outFile [ format "%s-%s-%s \t\t %.2f%%" [lindex $data
3] [lindex $data 2] [lindex $data 1] $fraction]
          set chan [open my.log a]
          puts $chan $sel1
       puts $chan $sel2
       puts  $chan  [format  "%s-%s-%s  \t\t  %.2f%%"  [lindex  $data  3]
[lindex $data 2] [lindex $data 1] $fraction]
          close $chan

      }
      #set beta according to the fraction, this is optional
      set atom [atomselect $mol "segid [lindex $data 3] and resname
[lindex $data 2] and resid [lindex $data 1]"]
      $atom set beta $fraction
      $atom delete
  }

  if { $outFile != "stdout" } {
```

```
      close $outFile
  }

}
```

The following script generates the residues for the sustained contacts analysis:

```
import sys
orig_stdout = sys.stdout
f = open('contacts.tcl', 'w')
sys.stdout = f

d = list(range(0,223))

for i in d:
     for j in range (224,234):
          print                                                    ("puts
{.............................%d.............................
......}" %(i))
          print                                                    ("puts
{                                                                    }")
          print ('puts { %s = }' %(j) )
          print                                                    ("puts
{                                                                    }")
          print ("catch {contactFreq " '"protein and residue %d"' '
' '"protein and residue %d"}' %(i, j) )
          print ("wait 5")

sys.stdout = orig_stdout
```